Evaluating the Efficacy of Automated Text Summarization and Key Point Extraction in Reading-Driven Learning Systems for Educational Outcomes

Bakytbek Uulu[1] and Ainura Sadykova[2]

[1]Osh Technological University, Faculty of Computer Technologies, Lenin Avenue, Osh, Kyrgyzstan
[2]Kyrgyz State Technical University, Department of Information Systems, Chuy Avenue, Bishkek, Kyrgyzstan

2023

Abstract

Automated text summarization and key point extraction have emerged as vital technologies in the development of reading-driven learning systems intended to enhance educational outcomes. By condensing extensive textual materials into a manageable set of salient points, these methods aim to improve learners' comprehension, retention, and overall cognitive engagement. The underlying computational techniques often involve natural language processing algorithms, such as extractive or abstractive strategies, which can reduce content to smaller yet meaningfully representative subsets. The integration of such systems into formal and informal education settings has opened possibilities for more personalized, efficient, and adaptive learning experiences, enabling students to interact with texts that accommodate their individual reading proficiencies. Despite promising results reported in various experiments, there remains a need for systematic investigations into the precise impact of automation-driven summaries on knowledge acquisition, recall, and metacognition. Researchers are also exploring the application of advanced optimization models and symbolic logic frameworks to improve accuracy, consistency, and interpretability. This paper provides an in-depth analysis of how these techniques can potentially reshape the educational landscape by facilitating learners' navigation through extensive textual resources. Emphasis is placed on the underlying methodologies, the subtleties of system design, and the empirical evidence that evaluates effectiveness across diverse learning contexts and subject areas.

1 Introduction

The integration of automated text summarization and key point extraction into reading-driven learning systems is a crucial point of intersection between computational linguistics, cognitive psychology, and pedagogical science [1]. The technological landscape for text analysis has evolved significantly over the past decade, driven by new algorithmic paradigms and by an increasing need to handle vast quantities of educational data [2]. Contemporary digital learning platforms aim to provide concise, high-level overviews of educational materials, enabling students to grasp essential points rapidly, while also affording opportunities to investigate finer details at will [3]. This dual-layer approach leverages algorithmic summarization and textual analysis to optimize both the macro-level and micro-level reading experience. Traditional reading models typically require extensive manual annotation and summarization by instructors or students; however, as data-driven methods gain prominence, the manual overhead can be mitigated through automated approaches that reduce textual corpora to more manageable formats. [4]

The impetus for studying the efficacy of these systems arises from the recognition that reading comprehension and critical thinking are pillars of learning [5]. When learners engage with content, whether in science, literature, or specialized technical fields, the capacity to distill the essential arguments and logical frameworks under discussion is paramount [6]. By deploying automated summarization, key statements that convey fundamental conceptual relationships can be highlighted, potentially acting as anchors for deeper exploration. This capability aligns well with modern pedagogical approaches that emphasize active learning, metacognition, and high-order thinking skills [7], [8]. However, the extant body of research remains fragmented, with varied results reported under different contexts, data sets, and algorithmic settings [9]. It is essential to engage in a granular exploration of how well these summarization algorithms transfer across domains and learner types.

In many implementations, reading-driven learning systems incorporate extractive summarization methods, wherein the original sentences or phrases are selected in accordance with an importance score [10]. An importance score can be derived from advanced linguistic features, including term frequencies, part-of-speech relationships, and syntactic dependencies [11], [12]. Yet, some limitations persist, such as the risk of generating summaries that lack cohesion or do not encapsulate the overall logical flow of the original text [13]. To address these deficiencies, abstractive summarization approaches have gained traction. These approaches typically involve encoding sentences into vector spaces, applying transformations to these representations through neural architectures, and subsequently decoding them into new sentences [14]. The process can often be formalized through the function $f : X \to Y$, where X is a set of original sentences in a high-dimensional embedding space, and Y is a set of newly generated sentences that preserve semantic fidelity to the original text [15]. However, ensuring that f yields interpretable and factually accurate outputs remains an ongoing challenge.

The educational context of these systems introduces further complexity [16]. As a simplified logical statement, one might express an educational objective as $\forall l \in L, \exists s \in S : \text{Comprehension}(l, s) \geq \delta$, where L is a set of learners, S is a set of possible summaries, and δ is a predefined threshold denoting acceptable comprehension. While this statement simplifies the intricacies of learner-centered design, it highlights the importance of guaranteeing that each learner can access at least one summary that meets a comprehension standard [17]. The challenge, however, is ensuring that this threshold is maintained across diverse linguistic backgrounds, academic disciplines, and cognitive levels.

Another dimension lies in understanding how reading-driven learning systems can gather and interpret user interactions to improve both summarization quality and instructional design. For instance, many modern systems track the time spent by learners on specific sections of text, their interactions with highlighted key points, and their requests for additional elaboration [18]. Such user engagement metrics can be mapped to an updated weighting scheme for the summarization algorithm [19]. Suppose w_t represents a dynamic weight that adjusts the importance of certain textual segments based on cumulative user engagement up to time t [20]. The updated summarization pipeline might then integrate w_t directly into the scoring function g_s used to select or generate representative sentences. Through iterative optimization, the system can adapt summaries to better match evolving learner needs, effectively merging collaborative filtering techniques with text analysis models. [21]

Advancements in machine learning, particularly in transformer-based architectures, have further catalyzed the capabilities of text summarization and key point extraction [22]. The self-attention mechanism inherent in these models enables the handling of longer textual sequences without relying solely on recurrent connections. For large texts, one might represent entire paragraphs as sequences of tokens, forming a matrix $M \in \mathbb{R}^{n \times d}$, where n is the number of tokens and d is the embedding dimension. A multi-head attention layer then applies a set of learned weight matrices $\{W_Q, W_K, W_V\}$ to generate query, key, and value representations. The final summary representation can be formalized by a function [23]

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V,$$

where $Q, K, V \in \mathbb{R}^{n \times d_k}$, ensuring that relevant parts of the input text receive higher weighting in the summary generation process.

By examining the complex interplay of these theoretical perspectives, algorithmic methods, and practical implementations, it becomes apparent that rigorous and methodologically sound evaluations of automated summarization's pedagogical utility are warranted [24]. Questions persist regarding the balance between automation and human oversight, especially when the summarized material is nuanced and demands critical interpretive skills [25]. Moreover, the adaptability of these systems to domain-specific lexicons (for instance, advanced scientific texts containing specialized terminology) introduces additional layers of complexity. The ongoing development of reading-driven learning systems shows considerable promise for enhancing educational outcomes, but it also calls for precision in design, quality assurance, and alignment with instructional objectives. [26]

2 Literature and Theoretical Underpinnings

The earliest conceptualizations of automated text summarization trace back to mid-20th-century attempts at keyword extraction and straightforward frequency-based analysis [27]. These approaches laid the groundwork for more sophisticated strategies that incorporate notions of syntactic parsing, semantic role labeling, and rhetorical structure theory. Rhetorical structure theory suggests that natural language texts possess a discernible organizational framework, guiding the arrangement of sentences and paragraphs in meaningful ways [28]. In this theoretical context, summaries seek not only to compress content but also to maintain the rhetorical integrity of the text, capturing essential argumentative relations [29]. The subsequent era witnessed the emergence of vector-based approaches in which textual elements were projected into continuous vector spaces via embeddings that rely on context-free or context-sensitive transformations. [30]

Contemporary advances have been fueled by deep learning methods that capitalize on large-scale corpora to learn generalized language representations. By modeling co-occurrence patterns of tokens in corpora of billions of words, these techniques provide robust embeddings that can be fine-tuned for specialized domains, such as legal or medical texts [31], [32]. In education research, this ability to adapt to specialized domains is critical, as texts for learners often span different levels of complexity and topics [33]. The requirement for domain specificity can be represented through an adaptation function $\phi : \mathbb{R}^d \to \mathbb{R}^d$, where d denotes the embedding dimension, transforming general-purpose embeddings into domain-appropriate variants that prioritize relevant terminology and phrases. A typical example might be $\phi(v) = W_\phi v$, where W_ϕ is a learned matrix that shifts the embedding space to emphasize domain-specific semantic attributes. [34]

Another strand of theoretical investigation resides in logic-based or symbolic representations. In certain research threads, textual content is parsed into semantic graphs or conceptual maps, which capture relationships between entities and their attributes [35]. These structures can be encapsulated in an adjacency matrix $A \in \{0,1\}^{n\times n}$, where n represents the set of unique concepts extracted from the text, and $A_{ij} = 1$ indicates a conceptual linkage or relationship between concept i and concept j. Summarization in this framework becomes a problem of selecting the most salient subgraph $G' \subseteq G$ that preserves the dominant logical connections [36]. Selecting such a subgraph may involve solving a combinatorial optimization problem, with constraints such as $|G'| \leq k$, where k represents a limit on the size of the subgraph. This perspective is particularly appealing in instructional contexts that require coherence in domain-specific relationships, such as linking chemical reactants and products in a chemistry curriculum or tracing historical events in a chronological timeline. [37]

Grounded in these theoretical perspectives, the integration of automated summarization into reading-driven systems for improved educational outcomes has raised questions about the influence of summarization on knowledge recall, metacognitive awareness, and learner motivation [38]. Studies have proposed that short, high-level overviews of a text might facilitate the formation of mental schemata, thereby scaffolding the subsequent assimilation of more complex material [39]. Further inquiries have used quasi-experimental or randomized control designs in which one group receives machine-generated summaries and another receives either no summaries or instructor-produced summaries. The evaluation metrics for these studies often include comprehension tests, time-on-task measurements, and subjective measures of perceived difficulty [40]. A typical example might be to define a comprehension function $\mathrm{Comp}(S)$ that evaluates a summary S through a combination of lexical overlap with instructor key points and learner accuracy in a post-reading quiz. If $\mathrm{Comp}(S) \geq \alpha$, for some threshold α, one might consider the summary effective in facilitating learning.

Additionally, from the standpoint of theoretical modeling, the complexities of language acquisition and concept formation suggest that the mapping between textual materials and cognitive representations is highly non-linear [41]. A function such as $\mathrm{Learn}(x)$ might denote the educational value derived from text x. When engaging with automated summaries, the learner's gain could be represented by $\mathrm{Learn}(f(x))$, where f is the summarization operator. A central research question is to ascertain how well $\mathrm{Learn}(f(x))$ approximates $\mathrm{Learn}(x)$ in various contexts. If the summarization preserves enough salient information, then $\mathrm{Learn}(f(x)) \approx \mathrm{Learn}(x)$. However, in certain domains characterized by nuanced, context-heavy texts, a large divergence might occur. Researchers have attempted to minimize this divergence by aligning automated summaries more closely with the hierarchical structure of the domain, employing advanced inference engines that recognize and highlight the most pedagogically valuable segments. [42]

In recent years, Bayesian perspectives have also been proposed to model uncertainties associated with automated summarization [43]. For instance, consider a probabilistic model in which the latent variable z represents the crucial subset of information in a text [44]. The system attempts to infer $p(z \mid x)$ using a posterior distribution derived from both textual cues and user feedback. The summary produced by the system is then a function of z, designed to maximize some notion of expected informativeness relative to the entire text [45]. In educational contexts, the feedback loop can be particularly informative because learners' performance on quizzes or their navigational paths through the text can serve as empirical evidence to refine the posterior belief about which aspects of the text are truly pivotal [46]. Such Bayesian learning frameworks illustrate the interplay between user data, summary production, and iterative refinement, revealing a rich space for ongoing research into how these models adapt to individual learner profiles and to diverse textual genres. [47]

3 Proposed Methodological Frameworks for Automated Summaries in Education

Examining the potential of automated text summarization within reading-driven learning systems requires a robust methodological structure that accounts for the heterogeneity of learners, instructional objectives, and text genres. One approach is to conceptualize the summarization task as a constrained optimization problem in which an objective function $J(S)$ aims to balance brevity, coverage, coherence, and pedagogical value [48]. For instance, one might define [49]

$$J(S) = \lambda_1 \cdot \text{Brevity}(S) + \lambda_2 \cdot \text{Coverage}(S) + \lambda_3 \cdot \text{Coherence}(S) + \lambda_4 \cdot \text{PedValue}(S),$$

subject to constraints that the summary S must contain no more than k sentences or words. The coefficients λ_i govern the relative importance of each term in the objective [50]. Brevity focuses on ensuring that the summary remains succinct, coverage addresses the fraction of critical content included, coherence mandates logical flow, and pedagogical value emphasizes alignment with learning objectives. [51]

Implementing this optimization framework may employ algorithms inspired by linear programming or integer linear programming, although the problem is typically NP-hard due to its combinatorial nature [52]. Modern heuristic or metaheuristic methods, such as simulated annealing or genetic algorithms, can be adapted to navigate the large search space efficiently. Alternatively, neural optimization techniques could be integrated, utilizing gradient-based methods if the objective and constraints can be differentiably approximated [53]. For example, coverage could be estimated through a function $C(S)$ that measures the embedding similarity between the summary S and the original text x, or between S and a set of reference key concepts [54]. Coherence might be assessed through a specialized neural discriminator trained to assign higher scores to well-structured textual outputs.

Parallel to these algorithmic developments, key point extraction has often been treated as a narrower task focused on distilling statements that encapsulate a text's core arguments or factual assertions [55]. A logic-driven approach might parse the text into propositional or predicate forms, grouping related predicates [56], [57]. A given text might be represented by $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_m\}$, where each γ_i is a logical proposition. Key point extraction then seeks a minimal subset $\Gamma' \subseteq \Gamma$ such that Γ' retains the core semantic relationships, fulfilling a condition akin to $\Gamma' \models \Gamma_{\text{core}}$, where Γ_{core} denotes the essential propositions. Methods that rely on first-order logic resolution, knowledge graph constraints, or even extended formalisms that incorporate modalities and temporal relations could be used, especially for texts that contain intricate cause-and-effect sequences. [58]

In the educational realm, these frameworks often incorporate user modeling to personalize summaries or key points. By tracking each learner's mastery over different concepts, the system might modulate summarization to emphasize underdeveloped competencies [59]. For example, let $u \in \mathbb{R}^m$ be a vector representing a learner's mastery profile over m concepts, and let $v \in \mathbb{R}^m$ represent the concept coverage of a text to be summarized. One could define a personalized weighting function $W_{\text{pers}} = u \odot v$ (where $\odot$ denotes element-wise multiplication) to highlight those aspects of the text that align with the learner's weaker areas. The resulting summary would concentrate on bridging knowledge gaps, effectively functioning as a targeted scaffold [60]. Machine learning models trained on interactions across a population of learners can refine these personalizations, eventually converging toward robust summarization strategies that are sensitive to varied reading levels and concept mastery patterns. [61]

Evaluation of these frameworks must also be addressed. Standard metrics like ROUGE or BLEU, frequently used in summarization research, quantify the lexical overlap between system-generated outputs and reference summaries [62]. However, in educational contexts, the question of whether automated summaries genuinely facilitate learning transcends simple lexical similarity [63]. Thus, multi-dimensional assessments that integrate domain expert reviews, learner performance analytics, and subjective user satisfaction are crucial. A scenario might involve administering a comprehension quiz both before and after exposure to the system's summaries [64]. The difference in performance, correlated with variables such as the user's reading time and interactive queries, can provide empirical evidence for or against the efficacy of the summarization approach [65]. Advanced statistical modeling, such as hierarchical linear modeling, can discern whether these benefits scale across different educational levels or subject areas. [66]

Another challenge lies in balancing universal design principles against domain-specific optimizations. A summarization system configured for language arts might emphasize narrative coherence and thematic structure, while one tailored for scientific texts might prioritize clarity in the relationships between hypotheses, experimental results, and conclusions [67]. To handle these divergent needs, modular architectures can be employed [68]. One could develop a base model for general-purpose summarization and augment it with domain adapters that specialize in domain-relevant vocabulary, rhetorical frames, and logic patterns. In many cases, these adapters involve fine-tuning neural layers or adjusting domain-specific parameters in a multi-task learning setting [69]. The system can thus adapt to various educational domains without forfeiting the comprehensive insights derived from large-scale pretraining. [70]

All of the above considerations form part of an ever-evolving methodological landscape that challenges researchers and practitioners alike [71]. The fluidity of language, combined with the diverse objectives of learning, underscores the need for flexible frameworks that can adapt to a range of contexts. A thorough understanding of the interplay between text complexity, learner characteristics, and algorithmic design choices is essential for driving progress [72]. As the field advances, it is envisioned that these frameworks will expand to incorporate multimodal data—spanning images, videos, and interactive simulations—further enriching the potential for robust summaries that function as comprehensive educational tools. [73]

4 Empirical Evaluations and Data Analyses

The empirical dimension of evaluating automated summarization within educational contexts encompasses multiple study designs, data sets, and performance metrics [74]. One prevalent design involves comparing a control group, exposed to unmodified textual materials, with an experimental group receiving automated summaries or key point extractions. Dependent variables might include reading comprehension scores, measured through domain-specific quizzes, and physiological or behavioral indicators of cognitive load, such as eye-tracking metrics or keystroke logging data [75]. For instance, if T denotes the time spent reading a segment and C denotes a comprehension metric, investigators may examine correlations or even define a performance index $\Pi = \frac{C}{T}$ to capture reading efficiency improvements attributed to summarization.

In many studies, data are drawn from educational platforms hosting tens of thousands of learners, offering a broad variety of backgrounds and language proficiencies [76]. This diversity in data sets can reveal the potential biases and limitations of summarization algorithms when deployed in large-scale educational contexts. One line of inquiry focuses on the difference in system performance between native and non-native language readers, or between learners at distinct levels of prior domain knowledge [77]. Researchers may model these interactions via a hierarchical Bayesian framework, letting θ represent system parameters, g denote group-level factors such as language background or proficiency, and y the observed performance outcomes [78]. A generative model might be structured as: [79]

$$y \mid \theta, g \sim \mathcal{N}(\mu_g(\theta), \sigma^2),$$

where $\mu_g(\theta)$ is the group-specific mean outcome. By fitting such models, one can disentangle algorithmic performance from group-level variations, identifying whether automated summarization confers uniform advantages or disproportionately benefits certain subsets of learners. [80]

A parallel thread of research investigates how learners engage with the system's outputs, examining whether the presence of automatically highlighted key points shifts the locus of attention in beneficial or detrimental ways [81]. Log data might reveal frequent toggling between summary segments and the original text, possibly indicating confusion or the search for missing details. Alternatively, a smooth reading path might suggest that the provided summaries align well with learner expectations and informational needs [82]. Advanced analytics can capture these interactions through Markov Chain models, wherein each state represents a mode of engagement (e.g., reading summary, reading full text, consulting an external resource), and transitions occur based on user actions [83], [84]. The transition probabilities can reveal patterns in how automated summarization either streamlines or complicates the reading process. [85]

In certain large-scale empirical initiatives, research labs or educational technology companies have instrumented their platforms to run randomized A/B tests on live user populations. System A might offer single-sentence extracts as key points, while System B generates more fluid abstractive summaries [86]. Analyses of engagement rates, quiz performance, and even longer-term retention or course completion can yield valuable insights into the real-world impact of summarization [87]. Yet, these commercial data sets often pose additional challenges, including limited transparency into proprietary algorithms or platform-driven confounds [88]. Thus, academic collaborations and open data initiatives become paramount for ensuring reproducibility and for facilitating cross-institutional meta-analyses that aggregate findings across varied educational scenarios.

Another dimension of empirical evaluation concerns system reliability and error patterns [89]. Automated summarizers may occasionally introduce factual distortions, especially in abstractive modes where new sentences are generated [90]. Identifying, quantifying, and rectifying these errors is essential, as misinformation in an educational context can yield detrimental effects on learning. Error analysis often involves comparing each system-generated statement with reference materials or expert annotations, assigning penalties if the statement contradicts known facts or introduces inaccurate inferences [91]. Symbolic or logic-based approaches can employ consistency checks to ensure that the generated summary does not produce logical contradictions with a validated knowledge base [92]. For example, if the summarization inadvertently produces a statement p such that $p \land \neg p$ is derivable in the reference ontology, the system reveals a contradiction indicating a critical error in text interpretation. [93]

To further analyze the effect of summarization on educational gains, sophisticated statistical modeling can be employed, accounting for both the item-level difficulty of test questions and the inherent variability in student ability. Item Response Theory (IRT) models, for instance, can measure whether summarization shifts the difficulty parameter of certain items [94]. If a set of items is strongly aligned with the summarized content, that subset might effectively become easier for the group receiving summaries [95]. A typical IRT formulation might set the probability of a correct response on an item i for a learner j at time t as:

$$P(y_{i,j,t} = 1) = \sigma(\alpha_j - \beta_i + \delta_t),$$

where σ is a logistic function, α_j is the learner's latent ability, β_i is the item difficulty, and δ_t captures potential shifts associated with an intervention (in this case, automated summarization) [96], [97]. Analyzing changes in δ_t

across conditions provides empirical grounding for concluding whether summaries fundamentally alter the difficulty of assessment items and thus facilitate improved performance. [98]

As the body of empirical studies grows, meta-analyses that compile effect sizes, confidence intervals, and context-specific moderators will be essential for synthesizing the broader outcomes [99]. Only through cumulative evidence across diverse learning environments—ranging from language arts classrooms to professional development modules—can one ascertain the external validity and practical significance of automated summarization. Researchers can then build more refined theories or produce consensus guidelines on how to optimally integrate these technologies into reading-driven learning systems. [100]

5 Conclusion

This paper has explored the technical, theoretical, and empirical dimensions of automated text summarization and key point extraction as they intersect with reading-driven learning systems [101]. From the foundations of frequency-based methods to the advent of complex neural architectures, the field has seen dramatic improvements in its capacity to produce concise yet informative textual outputs [102]. These capabilities hold particular promise for educational environments, where learners grapple with extensive reading materials that demand high degrees of comprehension, retention, and analytical thought. Summaries and key points, automatically generated through extractive or abstractive methods, can serve as cognitive aids that streamline content navigation and highlight vital conceptual linkages. [103]

The theoretical constructs discussed—ranging from rhetorical structure theory to logical formalisms and graph-based methods—illustrate the diverse underpinnings that guide the design of summarization algorithms [104]. The inclusion of personalized user models enables these systems to align more closely with individual learning trajectories, potentially enhancing mastery over challenging subject matter. Methodologically, the proposition of objective functions that balance brevity, coverage, coherence, and pedagogical value underscores the complexity and multidimensionality of educational summarization tasks [105]. Equally important is the inclusion of rigorous error analysis and validation processes, ensuring that any automated system maintains factual accuracy and logical consistency. [106]

Empirical evaluations reveal that automated summaries can improve reading efficiency and, in some contexts, comprehension performance, but outcomes are not universally uniform [107]. Discrepancies emerge based on domain complexity, learners' language proficiency, and the intrinsic difficulty of the source text. Data-driven analyses, including hierarchical Bayesian models and Item Response Theory, lend granularity to understanding how different user groups respond to system-generated content [108]. By dissecting patterns of user interaction, reading duration, and post-reading assessments, researchers gain a nuanced view of both the benefits and possible pitfalls of automated summarization in educational workflows. [109]

Despite the promise inherent in this technology, significant challenges persist. Abstractive methods risk introducing inaccuracies or inconsistencies, especially in domain-specific texts where specialized terminology and precise logical relations matter [110]. Maintaining a balance between succinctness and depth remains a critical issue, particularly when students require not only distilled facts but also contextual understanding that fosters higher-order thinking skills [111]. Moreover, the integration of multimodal materials—ranging from interactive simulations to audiovisual content—poses fresh challenges for how best to represent and summarize information [112], [113]. Advances in symbolic reasoning, domain adaptation, and user modeling continue to drive research toward more robust and context-aware solutions.

In sum, automated text summarization and key point extraction hold the potential to serve as integral components of modern reading-driven learning systems, offering scalable and adaptive support to learners across varying stages of education [114]. The theoretical frameworks, algorithmic advances, and empirical findings presented underscore the multifaceted nature of this field, pointing to a future in which educational technologies are replete with tools that intelligently parse and highlight text for optimized learning experiences [115]. Ongoing research will undoubtedly refine these methods, incorporating ever more sophisticated approaches to represent knowledge, manage domain-specific constraints, and respond to the diverse needs of learners. As the technology continues to evolve, it is paramount that interdisciplinary collaborations among computer scientists, cognitive psychologists, instructional designers, and subject-matter experts remain at the forefront, ensuring that the convergence of automated summarization and education achieves its fullest potential in advancing learning outcomes. [116]

References

[1] M. L. Bernardi, M. Cimitile, and F. M. Maggi, "Data-aware process discovery for malware detection: An empirical study," *Machine Learning*, vol. 112, no. 4, pp. 1171–1199, Mar. 31, 2022. DOI: 10.1007/s10994-022-06154-3.

[2] S. Faieq, R. Saidi, H. E. Ghazi, A. Front, and M. D. Rahmani, "Building adaptive context-aware service-based smart systems," *Service Oriented Computing and Applications*, vol. 15, no. 1, pp. 21–42, Jan. 6, 2021. DOI: 10.1007/s11761-020-00310-0.

[3] Y. Li, M. Soliman, and P. Avgeriou, "Identifying self-admitted technical debt in issue tracking systems using machine learning," *Empirical Software Engineering*, vol. 27, no. 6, Jul. 10, 2022. DOI: 10.1007/s10664-022-10128-3.

[4] T. Diwan and J. V. Tembhurne, "Sentiment analysis: A convolutional neural networks perspective," *Multimedia Tools and Applications*, vol. 81, no. 30, pp. 44 405–44 429, Jun. 3, 2022. DOI: 10.1007/s11042-021-11759-2.

[5] C. McColeman, J. J. Thompson, N. Anvari, *et al.*, "Digit eyes: Learning-related changes in information access in a computer game parallel those of oculomotor attention in laboratory studies," *Attention, perception & psychophysics*, vol. 82, no. 5, pp. 2434–2447, Apr. 24, 2020. DOI: 10.3758/s13414-020-02019-w.

[6] R. vanOostveen, F. Desjardins, and S. M. Bullock, "Professional development learning environments (pdles) embedded in a collaborative online learning environment (cole): Moving towards a new conception of online professional learning," *Education and Information Technologies*, vol. 24, no. 2, pp. 1863–1900, Feb. 7, 2018. DOI: 10.1007/s10639-018-9686-6.

[7] S. Iqbal, S.-U. Hassan, N. R. Aljohani, S. Alelyani, R. Nawaz, and L. Bornmann, "A decade of in-text citation analysis based on natural language processing and machine learning techniques: An overview of empirical studies," *Scientometrics*, vol. 126, no. 8, pp. 6551–6599, Jun. 23, 2021. DOI: 10.1007/s11192-021-04055-1.

[8] M. Kansara, "A comparative analysis of security algorithms and mechanisms for protecting data, applications, and services during cloud migration," *International Journal of Information and Cybersecurity*, vol. 6, no. 1, pp. 164–197, 2022. DOI: 10.17613/9ps72-wfe91.

[9] Y. Zhang, M. Wang, M. Saberi, and E. Chang, "Knowledge fusion through academic articles: A survey of definitions, techniques, applications and challenges," *Scientometrics*, vol. 125, no. 3, pp. 2637–2666, Aug. 28, 2020. DOI: 10.1007/s11192-020-03683-3.

[10] A. Hora, "Excluding code from test coverage: Practices, motivations, and impact," *Empirical Software Engineering*, vol. 28, no. 1, Dec. 7, 2022. DOI: 10.1007/s10664-022-10259-7.

[11] N. Kalinnik, R. Kiesel, T. Rauber, M. Richter, and G. Rünger, "A performance- and energy-oriented extended tuning process for time-step-based scientific applications," *The Journal of Supercomputing*, vol. 77, no. 4, pp. 3484–3515, Aug. 25, 2020. DOI: 10.1007/s11227-020-03402-y.

[12] K. D. Forbus, C. Riesbeck, L. Birnbaum, K. Livingston, A. Sharma, and L. Ureel, "Integrating natural language, knowledge representation and reasoning, and analogical processing to learn by reading," in *Proceedings of the National Conference on Artificial Intelligence*, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, vol. 22, 2007, p. 1542.

[13] H. Tebbi and M. Hamadouche, "Multi-agent based arabic speech synthesis," *International Journal of Speech Technology*, vol. 27, no. 1, pp. 1–17, May 17, 2022. DOI: 10.1007/s10772-022-09975-8.

[14] K. Mandal, K. Kalita, and S. Chakraborty, "Gene expression programming for parametric optimization of an electrochemical machining process," *International Journal on Interactive Design and Manufacturing (IJIDeM)*, vol. 17, no. 2, pp. 649–666, Aug. 12, 2022. DOI: 10.1007/s12008-022-00989-9.

[15] A. Lyutov, Y. Uygun, and M.-T. Hütt, "Machine learning misclassification of academic publications reveals non-trivial interdependencies of scientific disciplines," *Scientometrics*, vol. 126, no. 2, pp. 1173–1186, Nov. 27, 2020. DOI: 10.1007/s11192-020-03789-8.

[16] P. A. P. Júnior and R. A. D. Penteado, "Obascid(-tool): An ontologically based approach for concern identification and classification and its computational support," *Journal of the Brazilian Computer Society*, vol. 24, no. 1, pp. 3–, Jan. 11, 2018. DOI: 10.1186/s13173-017-0067-6.

[17] L. Carvalho, R. Martinez-Maldonado, and P. Goodyear, "Instrumental genesis in the design studio," *International Journal of Computer-Supported Collaborative Learning*, vol. 14, no. 1, pp. 77–107, Feb. 28, 2019. DOI: 10.1007/s11412-019-09294-2.

[18] C. A. Collazos, F. L. Gutiérrez, J. Gallardo, M. Ortega, H. M. Fardoun, and A. I. Molina, "Descriptive theory of awareness for groupware development," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 12, pp. 4789–4818, Dec. 14, 2018. DOI: 10.1007/s12652-018-1165-9.

[19] H. Xu, Y. Ma, H. Liu, *et al.*, "Adversarial attacks and defenses in images, graphs and text: A review," *International Journal of Automation and Computing*, vol. 17, no. 2, pp. 151–178, Mar. 27, 2020. DOI: 10.1007/s11633-019-1211-x.

[20] J. Figueiredo and F. J. García-Peñalvo, "Design science research applied to difficulties of teaching and learning initial programming.," *Universal access in the information society*, vol. 23, no. 3, pp. 1–1161, Nov. 7, 2022. DOI: 10.1007/s10209-022-00941-4.

[21] W. Chen, J. Zhou, J. Zhu, G. Wu, and J. Wei, "Semi-supervised learning based tag recommendation for docker repositories," *Journal of Computer Science and Technology*, vol. 34, no. 5, pp. 957–971, Sep. 6, 2019. DOI: 10.1007/s11390-019-1954-4.

[22] L. Wu, Y. Wen, Y. Xiaoxi, B. Yan, S. He, and B. Xiaochen, "Synthetic lethal interactions prediction based on multiple similarity measures fusion," *Journal of Computer Science and Technology*, vol. 36, no. 2, pp. 261–275, Mar. 31, 2021. DOI: 10.1007/s11390-021-0866-2.

[23] N. Dipsis and K. Stathis, "A restful middleware for ai controlled sensors, actuators and smart devices," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 7, pp. 2963–2986, Sep. 12, 2019. DOI: 10.1007/s12652-019-01439-3.

[24] C. H. Hoffmann, "A philosophical view on singularity and strong ai," *AI & SOCIETY*, vol. 38, no. 4, pp. 1697–1714, Jan. 17, 2022. DOI: 10.1007/s00146-021-01327-5.

[25] I. Kuznetcova, T.-J. Lin, and M. Glassman, "Teacher presence in a different light: Authority shift in multi-user virtual environments," *Technology, Knowledge and Learning*, vol. 26, no. 1, pp. 79–103, Mar. 26, 2020. DOI: 10.1007/s10758-020-09438-6.

[26] S. Kemchi, A. Zitouni, and M. Djoudi, "Amace: Agent based multi-criterions adaptation in cloud environment," *Human-centric Computing and Information Sciences*, vol. 8, no. 1, pp. 1–28, Sep. 12, 2018. DOI: 10.1186/s13673-018-0149-2.

[27] I. Belyaeva, B. Gabrielson, Y.-P. Wang, *et al.*, "Multi-subject analysis for brain developmental patterns discovery via tensor decomposition of meg data.," *Neuroinformatics*, vol. 21, no. 1, pp. 115–141, Aug. 24, 2022. DOI: 10.1007/s12021-022-09599-y.

[28] I. Amirtharaj, T. Groot, and B. Dezfouli, "Profiling and improving the duty-cycling performance of linux-based iot devices," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 5, pp. 1967–1995, Jan. 16, 2019. DOI: 10.1007/s12652-019-01197-2.

[29] M. S. Jaliaawala and R. A. Khan, "Can autism be catered with artificial intelligence-assisted intervention technology? a comprehensive survey," *Artificial Intelligence Review*, vol. 53, no. 2, pp. 1039–1069, Feb. 2, 2019. DOI: 10.1007/s10462-019-09686-8.

[30] K. F. F. Jiechieu and N. Tsopze, "Skills prediction based on multi-label resume classification using cnn with model predictions explanation," *Neural Computing and Applications*, vol. 33, no. 10, pp. 5069–5087, Aug. 28, 2020. DOI: 10.1007/s00521-020-05302-x.

[31] M. M. Bersani, F. Marconi, D. A. Tamburri, A. Nodari, and P. Jamshidi, "Verifying big data topologies by-design : A semi-automated approach," *Journal of Big Data*, vol. 6, no. 1, pp. 1–23, May 18, 2019. DOI: 10.1186/s40537-019-0199-y.

[32] K. D. Forbus, C. Riesbeck, L. Birnbaum, K. Livingston, A. B. Sharma, and L. C. Ureel, "A prototype system that learns by reading simplified texts.," in *AAAI Spring Symposium: Machine Reading*, 2007, pp. 49–54.

[33] L. H. Iwaya, M. A. Babar, A. Rashid, and C. Wijayarathna, "On the privacy of mental health apps: An empirical investigation and its implications for app development.," *Empirical software engineering*, vol. 28, no. 1, pp. 2–, Nov. 8, 2022. DOI: 10.1007/s10664-022-10236-0.

[34] A. Ligot and M. Birattari, "Simulation-only experiments to mimic the effects of the reality gap in the automatic design of robot swarms," *Swarm Intelligence*, vol. 14, no. 1, pp. 1–24, Oct. 3, 2019. DOI: 10.1007/s11721-019-00175-w.

[35] C.-I. Hsu, S. P. J. Wu, and C. Chiu, "A hybrid swarm intelligence approach for blog success prediction," *International Journal of Computational Intelligence Systems*, vol. 12, no. 2, pp. 571–579, 2019. DOI: 10.2991/ijcis.d.190423.001.

[36] X. Chen, H. Xie, Z. Li, *et al.*, "Leveraging deep learning for automatic literature screening in intelligent bibliometrics," *International Journal of Machine Learning and Cybernetics*, vol. 14, no. 4, pp. 1483–1525, Dec. 15, 2022. DOI: 10.1007/s13042-022-01710-8.

[37] R. Singh, A. Sonawane, and R. Srivastava, "Recent evolution of modern datasets for human activity recognition: A deep survey," *Multimedia Systems*, vol. 26, no. 2, pp. 83–106, Oct. 14, 2019. DOI: 10.1007/s00530-019-00635-7.

[38] A. Ouadah, L. Zemmouchi-Ghomari, and N. Salhi, "Selecting an appropriate supervised machine learning algorithm for predictive maintenance," *The International Journal of Advanced Manufacturing Technology*, vol. 119, no. 7-8, pp. 4277–4301, Jan. 11, 2022. DOI: 10.1007/s00170-021-08551-9.

[39] E. Gobet, J. G. López-Salas, P. Turkedjiev, and C. Vasquez, "Stratified regression monte-carlo scheme for semilinear pdes and bsdes with large scale parallelization on gpus," *Archives of Computational Methods in Engineering*, vol. 27, no. 3, pp. 889–921, Apr. 4, 2019. DOI: 10.1007/s11831-019-09335-x.

[40] A. Maatoug, G. Belalem, and S. Mahmoudi, "A location-based fog computing optimization of energy management in smart buildings: Devs modeling and design of connected objects," *Frontiers of Computer Science*, vol. 17, no. 2, Aug. 9, 2022. DOI: 10.1007/s11704-021-0375-z.

[41] J. de Sousa e Silva, R. Gonçalves, F. Branco, A. Pereira, M. Au-Yong-Oliveira, and J. Martins, "Accessible software development: A conceptual model proposal," *Universal Access in the Information Society*, vol. 18, no. 3, pp. 703–716, Jul. 28, 2019. DOI: 10.1007/s10209-019-00688-5.

[42] X. Zhang, F. Wu, N. Yang, *et al.*, "In silico methods for identification of potential therapeutic targets," *Interdisciplinary sciences, computational life sciences*, vol. 14, no. 2, pp. 1–26, Nov. 26, 2021. DOI: 10.1007/s12539-021-00491-y.

[43] N. Idrissi and A. Zellou, "A systematic literature review of sparsity issues in recommender systems," *Social Network Analysis and Mining*, vol. 10, no. 1, pp. 1–23, Feb. 11, 2020. DOI: 10.1007/s13278-020-0626-2.

[44] A. L. Yuille and C. Liu, "Deep nets: What have they ever done for vision?" *International Journal of Computer Vision*, vol. 129, no. 3, pp. 781–802, Nov. 27, 2020. DOI: 10.1007/s11263-020-01405-z.

[45] L. Daming, L. Deng, and Z. Cai, "Statistical analysis of tourist flow in tourist spots based on big data platform and da-hkrvm algorithms," *Personal and Ubiquitous Computing*, vol. 24, no. 1, pp. 87–101, Nov. 14, 2019. DOI: 10.1007/s00779-019-01341-x.

[46] L. Wu, Q. Liu, W. Zhou, G. Mao, J. Huang, and H. Huang, "A semantic web-based recommendation framework of educational resources in e-learning," *Technology, Knowledge and Learning*, vol. 25, no. 4, pp. 811–833, Nov. 27, 2018. DOI: 10.1007/s10758-018-9395-7.

[47] H. Jin, Y. Miao, X. Chen, J.-H. Seo, and E.-M. Park, "Motivation research on the participation of multimedia web page users in "share" behavior: A case study of alipay," *Multimedia Tools and Applications*, vol. 80, no. 26, pp. 34 571–34 589, Mar. 13, 2020. DOI: 10.1007/s11042-020-08819-4.

[48] A. Garain, A. Basu, F. Giampaolo, J. D. Velásquez, and R. Sarkar, "Detection of covid-19 from ct scan images: A spiking neural network-based approach.," *Neural computing & applications*, vol. 33, no. 19, pp. 1–14, Apr. 16, 2021. DOI: 10.1007/s00521-021-05910-1.

[49] S. Scalabrino, G. Bavota, M. Linares-Vasquez, V. Piantadosi, M. Lanza, and R. Oliveto, "Api compatibility issues in android: Causes and effectiveness of data-driven detection techniques," *Empirical Software Engineering*, vol. 25, no. 6, pp. 5006–5046, Oct. 7, 2020. DOI: 10.1007/s10664-020-09877-w.

[50] T. R. Rao, P. Mitra, R. Bhatt, and A. Goswami, "The big data system, components, tools, and technologies: A survey," *Knowledge and Information Systems*, vol. 60, no. 3, pp. 1165–1245, Sep. 18, 2018. DOI: 10.1007/s10115-018-1248-0.

[51] A. Sánchez-Caballero, D. Fuentes-Jiménez, and C. Losada-Gutiérrez, "Real-time human action recognition using raw depth video-based recurrent neural networks," *Multimedia Tools and Applications*, vol. 82, no. 11, pp. 16 213–16 235, Oct. 28, 2022. DOI: 10.1007/s11042-022-14075-5.

[52] A. Vangala, A. K. Das, V. Chamola, V. Korotaev, and J. J. P. C. Rodrigues, "Security in iot-enabled smart agriculture: Architecture, security solutions and challenges," *Cluster Computing*, vol. 26, no. 2, pp. 879–902, Apr. 18, 2022. DOI: 10.1007/s10586-022-03566-7.

[53] Y. Hu, H. Li, H. Yang, Y. Sun, L. Sun, and Z. Wang, "Detecting stealthy attacks against industrial control systems based on residual skewness analysis," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 1–14, Mar. 19, 2019. DOI: 10.1186/s13638-019-1389-1.

[54] Y. Nawal, M. Oussalah, B. Fergani, and A. Fleury, "New incremental svm algorithms for human activity recognition in smart homes," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 10, pp. 13 433–13 450, Mar. 24, 2022. DOI: 10.1007/s12652-022-03798-w.

[55] P. Gunawardena, O. Amila, H. Sudarshana, *et al.*, "Real-time automated video highlight generation with dual-stream hierarchical growing self-organizing maps," *Journal of Real-Time Image Processing*, vol. 18, no. 5, pp. 1457–1475, Mar. 18, 2020. DOI: 10.1007/s11554-020-00957-0.

[56] B. N. Subudhi, D. K. Rout, and A. Ghosh, "Big data analytics for video surveillance," *Multimedia Tools and Applications*, vol. 78, no. 18, pp. 26 129–26 162, Jun. 5, 2019. DOI: 10.1007/s11042-019-07793-w.

[57] A. Sharma and K. D. Forbus, "Graph-based reasoning and reinforcement learning for improving q/a performance in large knowledge-based systems," in *2010 AAAI Fall Symposium Series*, 2010.

[58] W. L. Nowinski, "Towards an architecture of a multi-purpose, user-extendable reference human brain atlas.," *Neuroinformatics*, vol. 20, no. 2, pp. 1–22, Nov. 26, 2021. DOI: 10.1007/s12021-021-09555-2.

[59] Z. Liu, X. Xia, D. Lo, and J. Grundy, "Automatic, highly accurate app permission recommendation," *Automated Software Engineering*, vol. 26, no. 2, pp. 241–274, Mar. 19, 2019. DOI: 10.1007/s10515-019-00254-6.

[60] T. Sonnekalb, T. S. Heinze, and P. Mäder, "Deep security analysis of program code," *Empirical Software Engineering*, vol. 27, no. 1, pp. 1–39, Oct. 21, 2021. DOI: 10.1007/s10664-021-10029-x.

[61] C. Ioannou, I. Nurdiani, A. Burattin, and B. Weber, "Mining reading patterns from eye-tracking data: Method and demonstration," *Software and Systems Modeling*, vol. 19, no. 2, pp. 345–369, Oct. 12, 2019. DOI: 10.1007/s10270-019-00759-4.

[62] Z. Lu, S. Qin, X. Li, L. Li, and D. Zhang, "One-shot learning hand gesture recognition based on modified 3d convolutional neural networks," *Machine Vision and Applications*, vol. 30, no. 7, pp. 1157–1180, Aug. 1, 2019. DOI: 10.1007/s00138-019-01043-7.

[63] D. Hu, C. Wang, S. Zheng, and X. Cui, "Investigating the genealogy of the literature on digital pathology: A two-dimensional bibliometric approach," *Scientometrics*, vol. 127, no. 2, pp. 785–801, Jan. 11, 2022. DOI: 10.1007/s11192-021-04224-2.

[64] E. G. Zahran, A. A. Arafa, H. I. Saleh, and M. I. Dessouky, "A self learned invasive weed-mixed biogeography based optimization algorithm for rfid network planning," *Wireless Networks*, vol. 26, no. 6, pp. 4109–4127, Apr. 3, 2020. DOI: 10.1007/s11276-020-02316-0.

[65] D. Dell'Anna, F. Dalpiaz, and M. Dastani, "Requirements-driven evolution of sociotechnical systems via probabilistic reasoning and hill climbing," *Automated Software Engineering*, vol. 26, no. 3, pp. 513–557, Apr. 22, 2019. DOI: 10.1007/s10515-019-00255-5.

[66] T. Parr and K. J. Friston, "Generalised free energy and active inference," *Biological cybernetics*, vol. 113, no. 5, pp. 495–513, Sep. 27, 2019. DOI: 10.1007/s00422-019-00805-w.

[67] J. Ceha, E. Law, D. Kulic, P.-Y. Oudeyer, and D. Roy, "Identifying functions and behaviours of social robots during learning activities: Teachers' perspective," *International Journal of Social Robotics*, vol. 14, no. 3, pp. 1–15, Sep. 16, 2021. DOI: 10.1007/s12369-021-00820-7.

[68] M. Barros, A. Moitinho, and F. M. Couto, "Seen: Sequential enriched datasets for sequence-aware recommendations.," *Scientific data*, vol. 9, no. 1, pp. 478–, Aug. 4, 2022. DOI: 10.1038/s41597-022-01598-7.

[69] X. Qin, H. Guo, C. He, and X. Zhang, "Lightweight human pose estimation: Cvc-net," *Multimedia Tools and Applications*, vol. 81, no. 13, pp. 17 615–17 637, Mar. 7, 2022. DOI: 10.1007/s11042-022-12245-z.

[70] D. Yu and Z. Yan, "Combining machine learning and main path analysis to identify research front: From the perspective of science-technology linkage," *Scientometrics*, vol. 127, no. 7, pp. 4251–4274, Jun. 28, 2022. DOI: 10.1007/s11192-022-04443-1.

[71] J. An, S. Zhao, X. Lu, and N. Liu, "A two-stage multiple-factor aware method for travel product recommendation," *Multimedia Tools and Applications*, vol. 77, no. 21, pp. 28 991–29 012, May 11, 2018. DOI: 10.1007/s11042-018-5992-7.

[72] N. Ding, J. Lv, and L. Hu, "Research on national pattern reuse design and optimization method based on improved shape grammar," *International Journal of Computational Intelligence Systems*, vol. 13, no. 1, pp. 300–309, 2020. DOI: 10.2991/ijcis.d.200310.003.

[73] C. Chen, Z. Xing, and Y. Liu, "What's spain's paris? mining analogical libraries from q&a discussions," *Empirical Software Engineering*, vol. 24, no. 3, pp. 1155–1194, Sep. 18, 2018. DOI: 10.1007/s10664-018-9657-y.

[74] S. Höck and R. Riedl, "Cyby 2 : A strongly typed, purely functional framework for chemical data management," *Journal of cheminformatics*, vol. 11, no. 1, pp. 80–80, Dec. 30, 2019. DOI: 10.1186/s13321-019-0403-2.

[75] S. Chowdhury, R. Holmes, A. Zaidman, and R. Kazman, "Revisiting the debate: Are code metrics useful for measuring maintenance effort?" *Empirical Software Engineering*, vol. 27, no. 6, Aug. 17, 2022. DOI: 10.1007/s10664-022-10193-8.

[76] A. Bhattacharyya, A. Gupta, L. Kuppusamy, *et al.*, "A formal methods approach to predicting new features of the eukaryotic vesicle traffic system," *Acta Informatica*, vol. 58, no. 1, pp. 57–93, Dec. 2, 2019. DOI: 10.1007/s00236-019-00357-3.

[77] A. Tatnall, "Editorial for eait issue 3, 2022.," *Education and information technologies*, vol. 27, no. 3, pp. 2893–2908, Mar. 31, 2022. DOI: 10.1007/s10639-022-11030-7.

[78] J. L. Suárez, S. García, and F. Herrera, "Ordinal regression with explainable distance metric learning based on ordered sequences," *Machine Learning*, vol. 110, no. 10, pp. 2729–2762, Jul. 7, 2021. DOI: 10.1007/s10994-021-06010-w.

[79] M. Araszkiewicz, T. Bench-Capon, E. Francesconi, M. Lauritsen, and A. Rotolo, "Thirty years of artificial intelligence and law: Overviews," *Artificial Intelligence and Law*, vol. 30, no. 4, pp. 593–610, Aug. 6, 2022. DOI: 10.1007/s10506-022-09324-9.

[80] J. Casas-Roma, M. E. Rodríguez, and A. Huertas, "A common frame for formal imagination," *Minds and Machines*, vol. 29, no. 4, pp. 603–634, Oct. 19, 2019. DOI: 10.1007/s11023-019-09510-w.

[81] Z. Sharafi, B. Sharif, Y.-G. Guéhéneuc, A. Begel, R. Bednarik, and M. E. Crosby, "A practical guide on conducting eye tracking studies in software engineering," *Empirical Software Engineering*, vol. 25, no. 5, pp. 3128–3174, Jun. 12, 2020. DOI: 10.1007/s10664-020-09829-4.

[82] M. Dolata and G. Schwabe, "Translation and adoption: Exploring vocabulary work in expert-layperson encounters," *Computer Supported Cooperative Work (CSCW)*, vol. 28, no. 3, pp. 685–722, Jun. 4, 2019. DOI: 10.1007/s10606-019-09358-9.

[83] B. Emo, G. Farquet, M. Kaijima, and C. Hoelscher, "Linking the design of university learning spaces to students' experience," *Cognitive Processing*, vol. 22, no. S1, pp. 39–39, Sep. 6, 2021. DOI: 10.1007/s10339-021-01058-x.

[84] A. Sharma, "Structural and network-based methods for knowledge-based systems," Ph.D. dissertation, Northwestern University, 2011.

[85] F. Cena, S. Likavec, and A. Rapp, "Real world user model: Evolution of user modeling triggered by advances in wearable and ubiquitous computing: State of the art and future directions," *Information Systems Frontiers*, vol. 21, no. 5, pp. 1085–1110, Jan. 2, 2018. DOI: 10.1007/s10796-017-9818-3.

[86] P. Verma and S. K. Sood, "A comprehensive framework for student stress monitoring in fog-cloud iot environment: M-health perspective," *Medical & biological engineering & computing*, vol. 57, no. 1, pp. 231–244, Aug. 6, 2018. DOI: 10.1007/s11517-018-1877-1.

[87] X. Chen, H. Jiang, X. Li, *et al.*, "A systemic framework for crowdsourced test report quality assessment," *Empirical Software Engineering*, vol. 25, no. 2, pp. 1382–1418, Feb. 27, 2020. DOI: 10.1007/s10664-019-09793-8.

[88] J. Reinwein and S. Tassé, "Modality effects examined by means of an online sentence-picture comparison task.," *Journal of psycholinguistic research*, vol. 51, no. 3, pp. 521–542, Mar. 11, 2022. DOI: 10.1007/s10936-022-09849-9.

[89] S. Ranaei, A. Suominen, A. L. Porter, and S. Carley, "Evaluating technological emergence using text analytics: Two case technologies and three approaches," *Scientometrics*, vol. 122, no. 1, pp. 215–247, Nov. 2, 2019. DOI: 10.1007/s11192-019-03275-w.

[90] A. Decan, T. Mens, and P. Grosjean, "An empirical comparison of dependency network evolution in seven software packaging ecosystems," *Empirical Software Engineering*, vol. 24, no. 1, pp. 381–416, Feb. 10, 2018. DOI: 10.1007/s10664-017-9589-y.

[91] A. Prakash, "Continuous user authentication based score level fusion with hybrid optimization," *Cluster Computing*, vol. 22, no. 5, pp. 12 959–12 969, Feb. 10, 2018. DOI: 10.1007/s10586-018-1819-6.

[92] F. Sun, H. Xu, Y. Meng, and Z. Lu, "A bert-based model for coupled biological strategies in biomimetic design," *Neural Computing and Applications*, vol. 35, no. 3, pp. 2827–2843, Sep. 6, 2022. DOI: 10.1007/s00521-022-07734-z.

[93] I. H. Montano, E. P. Lafuente, J. B. Martínez, A. O. Mansilla, I. de la Torre Díez, and M. L. D. Río-Solá, "Systematic review of telemedicine and ehealth systems applied to vascular surgery.," *Journal of medical systems*, vol. 46, no. 12, pp. 104–, Dec. 6, 2022. DOI: 10.1007/s10916-022-01895-z.

[94] M. Al-Hawawreh, N. Moustafa, and J. Slay, "A threat intelligence framework for protecting smart satellite-based healthcare networks," *Neural computing & applications*, vol. 36, no. 1, pp. 1–21, Sep. 8, 2021. DOI: 10.1007/s00521-021-06441-5.

[95] G. Liebel and S. Chakraborty, "Ethical issues in empirical studies using student subjects: Re-visiting practices and perceptions," *Empirical Software Engineering*, vol. 26, no. 3, pp. 1–37, Mar. 20, 2021. DOI: 10.1007/s10664-021-09958-4.

[96] Z. Wang, H. Jin, W. Dai, K.-K. R. Choo, and D. Zou, "Ethereum smart contract security research: Survey and future research opportunities," *Frontiers of Computer Science*, vol. 15, no. 2, pp. 152 802–, Oct. 2, 2020. DOI: 10.1007/s11704-020-9284-9.

[97] A. Abhishek and A. Basu, "A framework for disambiguation in ambiguous iconic environments," in *AI 2004: Advances in Artificial Intelligence: 17th Australian Joint Conference on Artificial Intelligence, Cairns, Australia, December 4-6, 2004. Proceedings 17*, Springer, 2005, pp. 1135–1140.

[98] T. B. Tuli, T. O. Terefe, and M. U. Rashid, "Telepresence mobile robots design and control for social interaction.," *International journal of social robotics*, vol. 13, no. 5, pp. 1–10, Jul. 13, 2020. DOI: 10.1007/s12369-020-00676-3.

[99] T. Hashimoto, "The emergent constructive approach to evolinguistics: Considering hierarchy and intention sharing in linguistic communication," *Journal of Systems Science and Systems Engineering*, vol. 29, no. 6, pp. 675–696, Oct. 29, 2020. DOI: 10.1007/s11518-020-5469-x.

[100] V. Nejkovic, F. Jelenkovic, N. Makris, V. Passas, T. Korakis, and M. Tosic, "Semantic coordination on the edge of heterogeneous ultra dense networks," *Journal of Network and Systems Management*, vol. 29, no. 2, pp. 1–28, Jan. 20, 2021. DOI: 10.1007/s10922-020-09576-3.

[101] F. Leotta, M. Mecella, and D. Sora, "Visual process maps: A visualization tool for discovering habits in smart homes," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 5, pp. 1997–2025, Jan. 30, 2019. DOI: 10.1007/s12652-019-01211-7.

[102] C. Mallick, A. K. Das, J. Nayak, D. Pelusi, and S. Vimal, "Evolutionary algorithm based ensemble extractive summarization for developing smart medical system," *Interdisciplinary sciences, computational life sciences*, vol. 13, no. 2, pp. 229–259, Feb. 12, 2021. DOI: 10.1007/s12539-020-00412-5.

[103] L. K. Singh, null Pooja, H. Garg, and M. Khanna, "Retracted article: An iot based predictive modeling for glaucoma detection in optical coherence tomography images using hybrid genetic algorithm," *Multimedia Tools and Applications*, vol. 81, no. 26, pp. 37 203–37 242, Aug. 12, 2022. DOI: 10.1007/s11042-022-13540-5.

[104] A. Ferrari, G. Gori, B. Rosadini, *et al.*, "Detecting requirements defects with nlp patterns: An industrial experience in the railway domain," *Empirical Software Engineering*, vol. 23, no. 6, pp. 3684–3733, Feb. 20, 2018. DOI: 10.1007/s10664-018-9596-7.

[105] F. A. Niaki, A. Pleta, L. Mears, N. Potthoff, J. A. Bergmann, and P. Wiederkehr, "Trochoidal milling: Investigation of dynamic stability and time domain simulation in an alternative path planning strategy," *The International Journal of Advanced Manufacturing Technology*, vol. 102, no. 5, pp. 1405–1419, Jan. 7, 2019. DOI: 10.1007/s00170-018-03280-y.

[106] H. Khatter and A. Ahlawat, "An intelligent personalized web blog searching technique using fuzzy-based feedback recurrent neural network," *Soft Computing*, vol. 24, no. 12, pp. 9321–9333, Apr. 12, 2020. DOI: 10.1007/s00500-020-04891-y.

[107] S. Laato, M. Mäntymäki, A. K. N. Islam, S. Hyrynsalmi, and T. Birkstedt, "Trends and trajectories in the software industry: Implications for the future of work," *Information Systems Frontiers*, Apr. 20, 2022. DOI: 10.1007/s10796-022-10267-4.

[108] R. R. Sola-Guirado, G. Guerrero-Vacas, and Ó. Rodríguez-Alabanda, "Teaching cad/cam/cae tools with project-based learning in virtual distance education," *Education and information technologies*, vol. 27, no. 4, pp. 1–23, Nov. 25, 2021. DOI: 10.1007/s10639-021-10826-3.

[109] K. Mahmud and M. Usman, "Trust establishment and estimation in cloud services: A systematic literature review," *Journal of Network and Systems Management*, vol. 27, no. 2, pp. 489–540, Oct. 1, 2018. DOI: 10.1007/s10922-018-9475-y.

[110] G. Gadelha, F. Ramalho, and T. Massoni, "Traceability recovery between bug reports and test cases-a mozilla firefox case study," *Automated Software Engineering*, vol. 28, no. 2, pp. 1–46, Jul. 7, 2021. DOI: 10.1007/s10515-021-00287-w.

[111] A. C. Junior, E. Casanova, A. Soares, *et al.*, "Coraa asr: A large corpus of spontaneous and prepared speech manually validated for speech recognition in brazilian portuguese," *Language Resources and Evaluation*, vol. 57, no. 3, pp. 1139–1171, Nov. 21, 2022. DOI: 10.1007/s10579-022-09621-4.

[112] S. Pal, P. K. D. Pramanik, and P. Choudhury, "Enhanced metadata modelling and extraction methods to acquire contextual pedagogical information from e-learning contents for personalised learning systems," *Multimedia Tools and Applications*, vol. 80, no. 16, pp. 25 309–25 366, Apr. 17, 2021. DOI: 10.1007/s11042-020-10380-z.

[113] A. Basu *et al.*, "Iconic interfaces for assistive communication," in *Encyclopedia of Human Computer Interaction*, IGI Global, 2006, pp. 295–302.

[114] G. Woolcott, D. Chamberlain, Z. Hawes, *et al.*, “The central position of education in knowledge mobilization: Insights from network analyses of spatial reasoning research across disciplines,” *Scientometrics*, vol. 125, no. 3, pp. 2323–2347, Sep. 14, 2020. DOI: 10.1007/s11192-020-03692-2.

[115] J. Beese, M. K. Haki, S. Aier, and R. Winter, “Simulation-based research in information systems: Epistemic implications and a review of the status quo,” *Business & Information Systems Engineering*, vol. 61, no. 4, pp. 503–521, Feb. 23, 2018. DOI: 10.1007/s12599-018-0529-1.

[116] H. Zhang, A. Ibrahim, B. Parsia, E. Poliakoff, and S. Harper, “Passive social sensing with smartphones: A systematic review,” *Computing*, vol. 105, no. 1, pp. 29–51, Aug. 12, 2022. DOI: 10.1007/s00607-022-01112-2.