A Survey of Personalized Large Language Models and Text Summarization Techniques

www.surveyx.cn

Abstract

This survey explores the transformative potential of personalized large language models (LLMs) and text summarization techniques in enhancing user experience across diverse domains. Personalized LLMs, by tailoring outputs based on user-specific data, significantly improve content relevance and effectiveness. Frameworks like APIDocBooster demonstrate advancements in informativeness, relevance, and faithfulness, highlighting the critical role of personalization in long-text generation. In healthcare, personalized LLMs enhance patient care documentation and deliver contextually relevant responses, improving patient outcomes. Educational tools benefit from personalized summarization techniques, such as HumSum, which enhance user engagement, while mixed reality technologies like RealitySummary further demonstrate potential in enhancing reading experiences. Abstractive summarization techniques, exemplified by StructSum, improve abstractiveness and content coverage, showcasing their potential in generating meaningful summaries. The survey identifies challenges in managing risks associated with personalized LLMs, emphasizing the need for future research to address issues of factual accuracy, scalability, and privacy. By overcoming these challenges, personalized LLMs and text summarization techniques can deliver tailored and meaningful experiences, ultimately enhancing user satisfaction and engagement across various applications.

1 Introduction

1.1 Importance of Personalization in NLP

Personalization in natural language processing (NLP) is crucial for enhancing user experience by tailoring interactions to individual preferences and needs. This is particularly relevant in recommendation systems and creative applications, where personalized content significantly boosts user satisfaction and interaction quality. For example, personalized dialog summarization in customer support leads to more effective and satisfactory interactions [1].

The emergence of end-user applications utilizing large language models (LLMs) has further expanded the scope of personalization, albeit with accompanying privacy challenges that users may overlook [2]. Despite these concerns, adapting content to meet individual needs remains essential for improving engagement, as evidenced by applications like PersonaMark in personalized watermarking and various interior design tools. By addressing evolving user expectations, personalization in NLP not only enhances interaction quality but also meets diverse user requirements across multiple domains [3].

1.2 Structure of the Survey

This survey offers a thorough examination of the challenges and opportunities presented by large language models (LLMs) in personalization, aiming to connect personalized text generation with

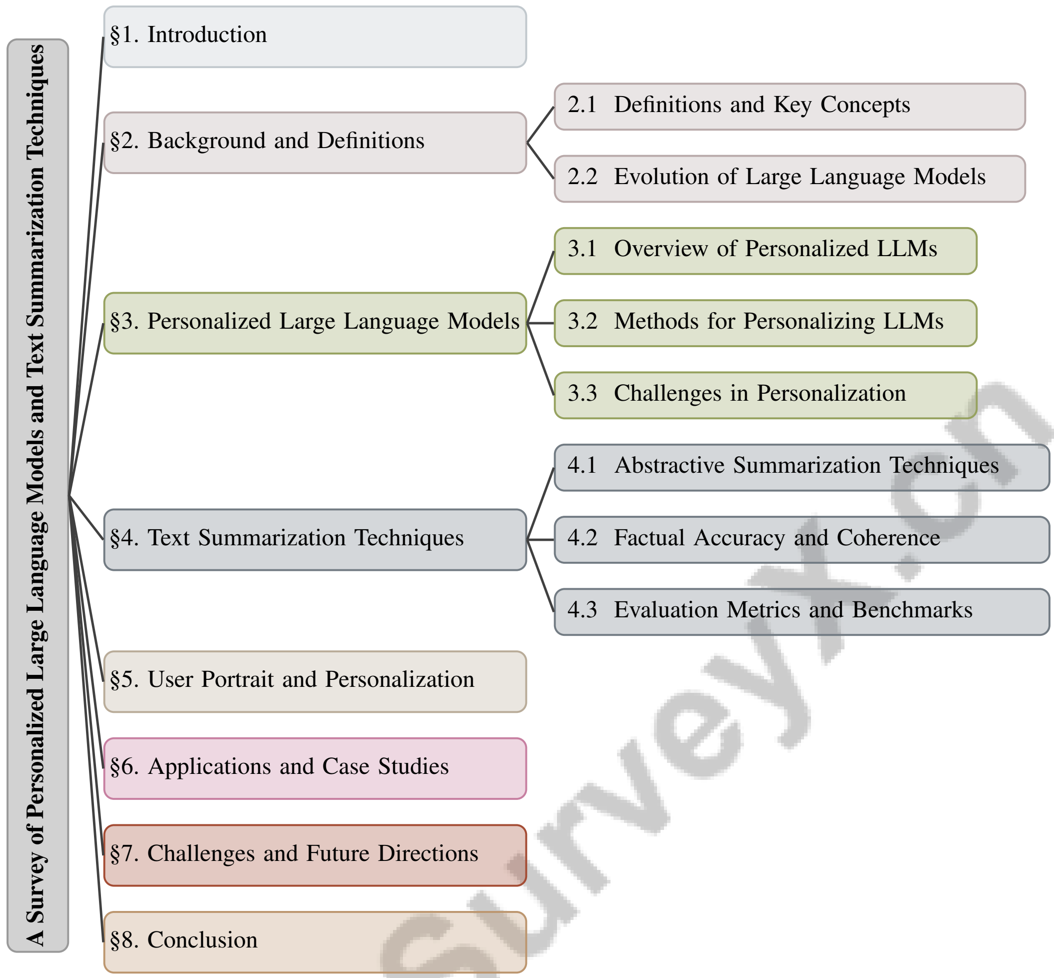


Figure 1: chapter structure

LLM applications for downstream personalization. The paper commences with an introduction to personalized LLMs and their relevance in NLP, followed by a discussion on the importance of personalization for enhancing user experience. The background section defines key concepts and explores the evolution of LLMs, laying the groundwork for understanding personalized LLMs.

Following this, the survey investigates the development and implementation of personalized LLMs, detailing various personalization methods and their associated challenges. The text summarization techniques section emphasizes abstractive summarization, focusing on critical aspects such as factual accuracy and coherence, along with the evaluation metrics and benchmarks for assessing these techniques.

The survey further explores user portraits, underscoring their importance in personalized content generation. It outlines techniques for developing user portraits, including leveraging social media data to profile personality traits and utilizing historical user interactions to inform content preferences. The discussion highlights how user portraits enhance personalization strategies and evaluate personalized text generation and summarization models, demonstrating their role in improving the relevance and quality of content tailored to individual users [4, 5, 6, 7, 8]. The survey also presents applications and case studies across domains such as healthcare, e-commerce, and education, showcasing the practical use of personalized LLMs.

In examining the current challenges and future directions in user portrait development, the paper highlights critical issues like privacy, bias, resource constraints, and scalability. Addressing these challenges is vital for enhancing user profiling techniques, especially concerning social media data for understanding personality traits. The paper identifies significant research opportunities, particularly in innovative methodologies such as personalized text watermarking, which can improve accountability and traceability while ensuring data privacy in large-scale applications [7, 8, 9, 10]. The conclusion encapsulates key insights, reinforcing the significance of personalized LLMs and text summarization in elevating user experience.The following sections are organized as shown in Figure 1.

2 Background and Definitions

2.1 Definitions and Key Concepts

Personalized large language models (LLMs) are advanced AI systems that customize outputs based on user-specific data, preferences, and interactions, thereby enhancing user engagement through tailored content. These models excel in dialogue summarization by adapting responses to individual needs and condensing multi-participant dialogues into coherent summaries while maintaining key information [11]. The integration of multimodal information fusion further enriches text summarization by incorporating diverse sources, such as speech, for a comprehensive understanding of content [12].

Text summarization aims to distill essential information from extensive texts into concise summaries tailored to specific tasks and user needs [1]. Abstractive summarization, which synthesizes novel summaries rather than merely extracting content, is crucial for factual accuracy and coherence, especially in contexts like Vietnamese news, where concise outputs are essential [13]. Personalized LLMs enhance the summarization of online discussions, often characterized by informal language and topic drift, challenges that traditional methods struggle to address [14].

User portraits, encapsulating preferences, behaviors, and demographic information, form the foundation for personalized content generation [15]. These profiles enable systems to tailor interactions to align with user expectations, enhancing engagement and satisfaction. Leveraging user portraits alongside personalized LLMs aligns content generation with individual preferences, improving interaction quality across applications [3]. For example, in interior design, tools like I-Design convert unstructured user input into structured scene graphs representing personalized layouts [16].

Interactive summarization allows users to receive tailored overviews and detailed summaries based on their selections, contributing to personalized content generation [1]. Personalized LLMs utilize multimodal information to address existing methods' inadequacies in providing personalized experiences, such as in museum tours [12]. Unsupervised abstractive opinion summarization generates coherent summaries reflecting consensus opinions from multiple reviews without relying on labeled data [11].

2.2 Evolution of Large Language Models

The evolution of large language models (LLMs) has been marked by significant advancements in model architecture and data availability, enhancing their capabilities in natural language processing tasks [17]. Initially, the field relied on extractive summarization due to limited corpora size in the 1990s, constraining task complexity [18]. However, the advent of larger datasets and sophisticated models shifted the focus to abstractive summarization methods, generating more coherent and contextually relevant summaries.

The development of sequence-to-sequence models has been foundational for tasks like machine translation and abstractive summarization [19]. Recent advancements allow models to operate directly on raw input text, reducing reliance on engineered features and demonstrating impressive performance across various tasks [20]. Innovations such as integrating Role-Playing Language Agents (RPLAs) with LLMs enhance their ability to simulate human-like interactions [21].

Research categorization on integrating LLMs into personalization systems highlights their unique capabilities, including in-context learning and instruction following, essential for personalized content generation [22]. Advancements in user-level, persona-level, and global preference personalization broaden LLM applicability across diverse domains [23].

Surveys of LLMs in user modeling categorize approaches and applications, showcasing the advancements these models bring [17]. Datasets comprising 500,000 unique examples across multiple natural language task categories provide a robust foundation for comprehensive evaluation, facilitating nuanced and effective model development [15]. For instance, a benchmark for Vietnamese abstractive multi-document summarization offers a human-annotated dataset aiding in model evaluation, emphasizing context-specific data's importance in advancing LLM capabilities [13].

The trajectory of LLMs reflects continuous enhancement of capabilities, driven by advances in model architecture and data availability. These developments have significantly improved LLM performance in applications like zero-shot abstractive summarization, legal case judgment summarization, personalized information retrieval, and incident reporting, establishing them as essential tools in natural language processing. LLMs demonstrate state-of-the-art performance in generating coherent summaries, even in complex domains like law, while addressing challenges such as model hallucination and context understanding. Frameworks like LaMSUM leverage LLMs for effective extractive summarization, showcasing their versatility and importance in processing large datasets and improving user interactions across multiple contexts [24, 25, 26, 27].

The exploration of personalized large language models (LLMs) necessitates a comprehensive understanding of their hierarchical structure and the multifaceted applications they present. As illustrated in Figure 2, this figure provides a detailed overview of the categorization of various applications, techniques, and benefits associated with personalized LLMs. The methods for personalizing these models are systematically divided into three primary approaches: fine-tuning, plug-and-play, and hybrid frameworks. Furthermore, the figure delineates the challenges inherent in the personalization process, which encompass technical difficulties, privacy concerns, and ongoing developments aimed at addressing these issues. This visual representation not only enhances our understanding of the complexities involved in personalized LLMs but also serves as a pivotal reference point for the subsequent discussions in this review.

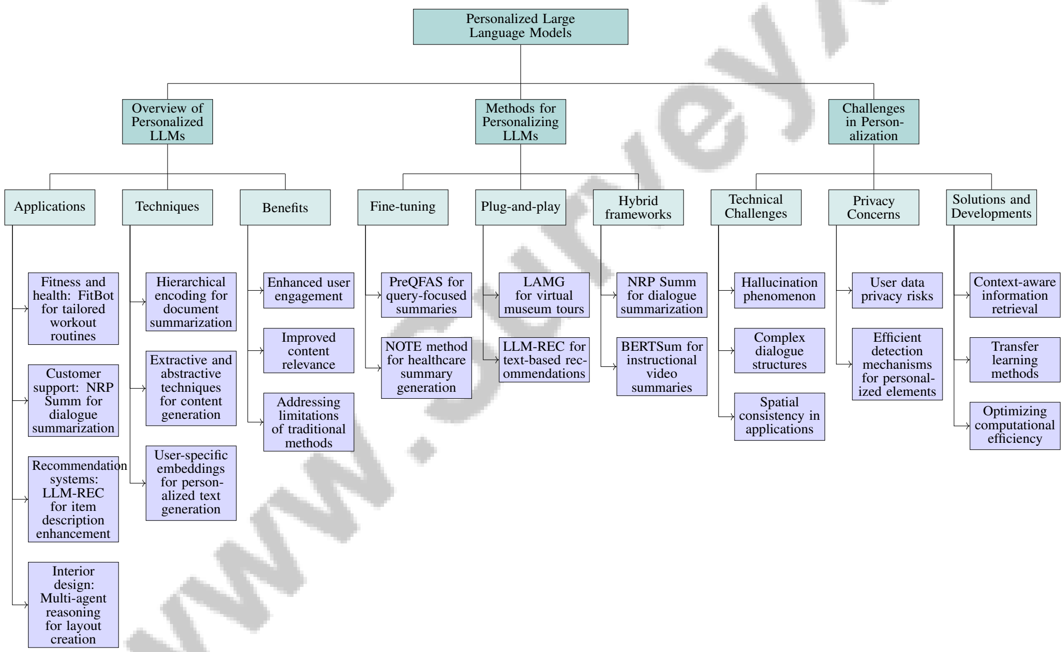


Figure 2: This figure illustrates the hierarchical structure of personalized large language models (LLMs), highlighting their applications, methods, and challenges. The overview section categorizes various applications, techniques, and benefits of personalized LLMs. Methods for personalizing LLMs are divided into fine-tuning, plug-and-play, and hybrid frameworks. Challenges in personalization include technical challenges, privacy concerns, and proposed solutions and developments.

3 Personalized Large Language Models

3.1 Overview of Personalized LLMs

Personalized large language models (LLMs) represent a significant advancement in natural language processing, enabling interactions and content generation tailored to individual user preferences. By leveraging user-specific data, these models enhance engagement and deliver relevant outputs, transforming user experiences across various applications. Tan et al. categorize LLM-user modeling

AI-generated, for reference only.

techniques based on approaches and applications, highlighting LLMs' critical role in user modeling [17].

In fitness and health, FitBot exemplifies personalized LLM applications by generating tailored workout routines and advice based on user interactions, providing customized fitness solutions [28]. The CSDS benchmark further underscores the importance of personalized LLMs in producing role-oriented and topic-based structural summaries, capturing diverse speaker perspectives [29].

To provide a comprehensive overview of the landscape of personalized LLMs, Figure 3 illustrates the key applications, techniques, and challenges associated with these models. This figure highlights specific applications such as FitBot, the CSDS benchmark, and the Hierarchical Transformer, while also exploring techniques like LLM-REC, multi-agent reasoning, and extractive-abstractive methods. Furthermore, it addresses the challenges in user modeling, watermarking, and data processing that are inherent to the deployment of personalized LLMs.

The versatility of personalized LLMs is illustrated by models like the Hierarchical Transformer, which processes multiple documents to create coherent abstractive summaries through hierarchical encoding, capturing inter-document relationships [11]. In customer support, the NRP Summ method enhances user interactions by summarizing dialogues with a focus on next response prediction, identifying salient sentences [1].

LLM-REC showcases personalized LLMs' potential in recommendation systems by utilizing prompting strategies to augment item descriptions [3]. In interior design, combining multi-agent reasoning with scene graph representations demonstrates personalized LLMs' adaptability in practical applications, facilitating complex layout creation [16].

Integrating extractive and abstractive techniques, as proposed by Subramanian et al., improves conditioning on relevant information, enhancing personalized content generation effectiveness [30]. PersonaMark illustrates an innovative approach by embedding user-specific watermarks into generated text, contributing to personalized text watermarking [7].

Personalized LLMs are poised to enhance content generation and consumption by delivering experiences finely tuned to individual user profiles and contextual nuances. By employing user-specific embeddings from historical interactions, these models capture user preferences and styles, leading to more engaging outputs. This transformation addresses traditional methods' limitations, which often lack continuity in user history and fail to reflect diverse interests. Bridging personalized text generation and recommendation systems, these LLMs are positioned as a cornerstone technology in natural language processing evolution [24, 23, 3, 5].

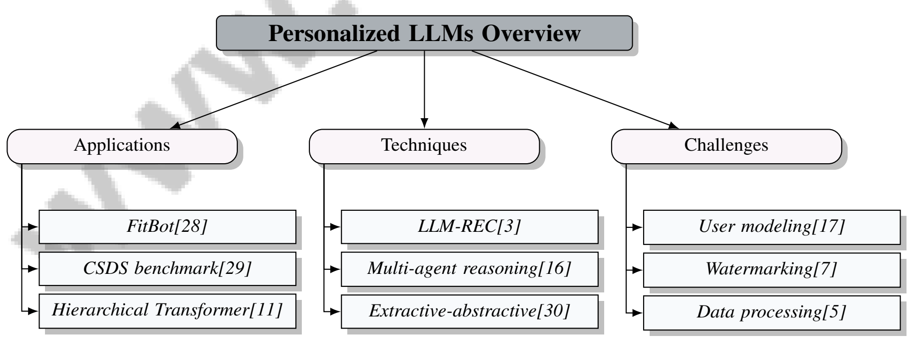


Figure 3: This figure illustrates the key applications, techniques, and challenges associated with personalized large language models (LLMs). It highlights specific applications such as FitBot, CSDS benchmark, and Hierarchical Transformer, explores techniques like LLM-REC, multi-agent reasoning, and extractive-abstractive methods, and addresses challenges in user modeling, watermarking, and data processing.

3.2 Methods for Personalizing LLMs

Personalizing large language models (LLMs) involves sophisticated techniques to tailor outputs to user-specific requirements, enhancing relevance and engagement. Fine-tuning is a common method where pre-trained transformer models are adapted using domain-specific datasets to improve

AI-generated, for reference only.

performance in specialized contexts. For instance, PreQFAS generates query-focused abstractive summaries by fine-tuning pre-trained models on small datasets [31]. Similarly, the NOTE method optimizes summary generation in healthcare settings through structured processes based on clinical ontologies [32].

Plug-and-play approaches offer dynamic customization without altering the model's core architecture. The LAMG method personalizes virtual museum tours by aligning exhibits with user preferences, showcasing these strategies' potential in delivering personalized experiences [12]. Furthermore, LLM-REC enhances item descriptions through various prompting strategies, improving text-based recommendations [3].

Hybrid frameworks are crucial in personalizing LLMs. For example, the NRP Summ method employs an unsupervised extractive summarization approach that assesses sentence importance based on their influence on subsequent responses, enhancing dialogue summarization personalization [1]. BERTSum illustrates the integration of extractive and abstractive techniques, using a BERT-based encoder-decoder architecture to generate coherent summaries from instructional videos [33].

Cross-modal approaches enrich personalization capabilities, such as segmenting long speech inputs and extracting features with Q-Former to generate summaries through autoregressive LLMs, enhancing speech summarization personalization [12]. PersonaMark employs a sentence-level generation strategy with a unique hashing function to create personalized watermarks, demonstrating innovative methods for embedding personalized elements into generated text [7].

The methodologies for personalizing LLMs—including fine-tuning, plug-and-play techniques utilizing historical texts, and integrating hybrid frameworks with cross-modal capabilities—demonstrate a comprehensive range of strategies aimed at enhancing user engagement and satisfaction. These approaches address the challenges of aligning model outputs with individual preferences, optimizing relevant information retrieval, and improving recommendation quality through innovative prompting techniques, ultimately leading to more tailored interactions with LLMs [24, 3, 5]. By leveraging advances in machine learning and data processing, these techniques deliver personalized experiences aligned with individual user profiles and contextual factors.

3.3 Challenges in Personalization

The development of personalized large language models (LLMs) faces several challenges that complicate their implementation and efficacy. A primary issue is the hallucination phenomenon, where models generate outputs that do not accurately reflect input data, posing significant risks in applications demanding factual accuracy [17]. Privacy concerns are also critical, as users may prioritize immediate utility over long-term privacy, inadvertently sharing personal information without fully understanding potential risks [2]. This emphasizes the need for robust privacy-preserving mechanisms in personalized systems.

In recommendation systems, a core challenge is the insufficiency of information within item descriptions, which hampers effective personalized recommendations [3]. Additionally, dialogue summarization challenges arise from complex dialogue structures, multiple participants, and topic drifts, complicating the collection of labeled data necessary for training effective summarization models [34]. Accurately modeling user interactions and preferences is crucial for delivering personalized experiences [17].

The complexity of implementing personalized hashing functions for large user bases, as highlighted in the PersonaMark method, presents significant challenges, necessitating efficient detection mechanisms to manage these personalized elements [7]. In spatial applications, ensuring spatial consistency and managing complex scenes with numerous objects can lead to placement conflicts, posing additional challenges to maintaining coherence and relevance in generated content [16].

Addressing these challenges requires innovative solutions to enhance the efficiency and accuracy of personalized LLMs. Ongoing development focuses on sophisticated methods for capturing user preferences, improving information retrieval systems through context-aware techniques and effective summarization. Additionally, enhancing data integration and transfer learning methods facilitates seamless interactions between LLMs and information retrieval systems. A critical aspect of this development is ensuring user privacy and optimizing computational efficiency during model alignment, addressing issues like model hallucination—where inaccurate or misinterpreted data is generated—thus fostering a user-centric approach in the digital landscape [24, 10]. By overcoming these obstacles, personalized LLMs can deliver more tailored user experiences, capturing fine-grained user information and adapting to the complexities of user historical content.

4 Text Summarization Techniques

4.1 Abstractive Summarization Techniques

Abstractive summarization techniques significantly advance text summarization by generating concise and semantically enriched summaries. Unlike extractive methods, which select verbatim sentences, abstractive techniques synthesize new sentences encapsulating core ideas. These methods employ advanced frameworks, such as attention-based neural networks and dual attentional seq2seq models, to enhance semantic relevance and coherence. Integrating extractive elements has been shown to improve summary quality, addressing challenges like out-of-vocabulary words and summary diversity, achieving competitive performance in evaluations and human assessments [35, 36, 37, 38]. This approach is particularly beneficial for complex documents and dialogues requiring nuanced understanding.

As illustrated in Figure 4, the figure highlights the key techniques in abstractive summarization, showcasing advanced frameworks, integration approaches, and challenges in handling specific applications. The Concept Pointer Network by Wenbo et al. exemplifies progress by selecting expressive concepts from a knowledge base, enhancing summary relevance [39]. BERTSum effectively generates abstractive summaries from both spoken and written instructions, showcasing its versatility across modalities [33]. Methods like Abstractive Meeting Summarization (AMS) utilize segmentation and fusion to produce coherent summaries from meeting transcripts, addressing coherence and readability challenges [40]. Additionally, a two-step process integrating extractive and abstractive techniques, as proposed by Subramanian et al., improves coherence by conditioning on significant extracted sentences [30].

Despite advancements, challenges persist in ensuring that abstractive summaries do not omit critical information or introduce factual inaccuracies, which can hinder practical applications [41]. However, approaches like PersonaMark illustrate the potential of abstractive summarization to enhance copyright protection and attribution in generated text, highlighting broader applicability [7].

Abstractive summarization techniques offer significant advantages by generating summaries that are informative, contextually aligned, and engaging. Recent innovations include a two-staged strategy producing multiple candidate summaries tailored to user preferences, ensuring grammatical fluency and fidelity to source material. The integration of extractive techniques, such as WordNet-based sentence ranking, further enhances semantic relevance. Advances in dual attentional seq2seq frameworks address challenges like out-of-vocabulary words and redundancy, achieving state-of-the-art performance in automated metrics and human evaluations across various datasets [35, 36, 37]. These evolving techniques aim to improve informativeness and faithfulness, enhancing the overall efficiency and impact of text summarization.

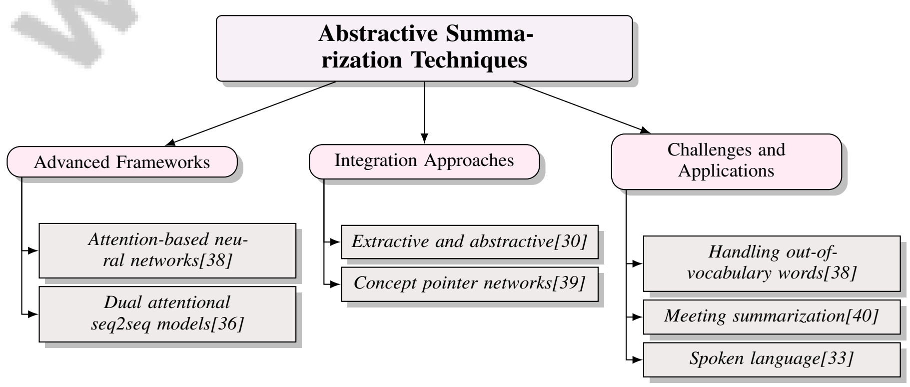


Figure 4: This figure illustrates the key techniques in abstractive summarization, highlighting advanced frameworks, integration approaches, and challenges in handling specific applications.

AI-generated, for reference only.

4.2 Factual Accuracy and Coherence

Ensuring factual accuracy and coherence in generated summaries is a critical challenge, particularly with large language models (LLMs) that often produce hallucinated content, deviating from source material and undermining reliability [42]. The scattered nature of essential information across utterances complicates the generation of coherent and concise summaries [43].

To address these challenges, several methodologies have emerged. The NRP Summ method employs an unsupervised extractive approach to identify key sentences in dialogue that influence subsequent responses, enhancing coherence [44]. The Concept Pointer Network enhances semantic depth by incorporating higher-level abstract concepts into summaries, reflecting both source content and abstract ideas [39].

Inconsistency and stylistic variations in spoken language further complicate summarization, impeding the effectiveness of existing methods [33]. Techniques like AMS address these issues through segmentation and fusion, producing coherent summaries from meeting transcripts and overcoming disfluencies common in noisy extractive summaries [40].

Traditional evaluation metrics such as ROUGE often fail to accurately reflect human assessments of summary quality, particularly in the presence of hallucinations [41]. While metrics like ROUGE-1 and ROUGE-2 measure relevance and informativeness by assessing overlap with reference summaries, they may not capture human judgment nuances, necessitating more sophisticated evaluation methods [45]. Current methods also struggle with managing long-range dependencies, leading to summaries overly reliant on original text and less abstract [30]. Addressing these challenges requires innovative training objectives and methodologies that bolster both factual accuracy and coherence, ensuring generated summaries reliably capture the essence of source material.

4.3 Evaluation Metrics and Benchmarks

Benchmark	**Size**	**Domain**	**Task Format**	**Metric**
FEQA[42]	288,000	Text Summarization	Faithfulness Evaluation	F1 score
CNewSum[46]	304,307	News Summarization	Summarization	ROUGE-1, ROUGE-2
USE-SUM[47]	10,000	Text Summarization	Question Answering	EM, F1
wiki$_{b}io_{g}pt3_{h}allucination$[48]	238	Text Summarization	Hallucination Detection	LLM-based methods, Ensemble methods
BLANC[49]	300	Text Summarization	Quality Assessment	Correlation Coefficient, BLANC Score
ChatGPT-Sum[50]	50	Text Summarization	Abstractive Summarization	ROUGE, METEOR
AQUAMUSE[51]	5,519	Natural Language Processing	Multi-Document Summarization	ROUGE-1, ROUGE-2
PerSEval[4]	200	News Summarization	Summarization	EGISES, PerSEval

Table 1: This table provides a comprehensive overview of various benchmarks used in the evaluation of text summarization techniques. It details the size, domain, task format, and metrics employed for each benchmark, highlighting the diversity and specificity of evaluation methods across different summarization tasks.

Evaluating summarization techniques is crucial for advancing the field, ensuring generated summaries are accurate and useful. Traditional metrics like ROUGE, including ROUGE-1, ROUGE-2, and ROUGE-L, assess overlap between generated summaries and human-written references, providing a quantitative measure of content similarity. However, these metrics often inadequately capture human judgment nuances, particularly in high-abstraction contexts [41].

To address these limitations, additional evaluation methods like BERTScore have been introduced, measuring semantic similarity between generated texts and reference summaries for a more nuanced quality assessment [43]. Human evaluations are also essential, offering insights into the informativeness and readability of summaries, particularly in highly abstractive contexts where automatic metrics may not fully reflect quality improvements [41].

Innovative approaches like the energy-based re-ranking model aim to enhance summary quality, although these improvements may not always align with human evaluations, underscoring the complexity of measuring summary quality in abstractive contexts [41]. Methodologies like the FEQA benchmark evaluate the faithfulness of generated summaries by comparing them to source documents, addressing existing automatic metrics' inadequacies [42].

The combination of automatic metrics, such as ROUGE and BERTScore, alongside qualitative human evaluations, creates a comprehensive framework for assessing summarization techniques. Methods like RISE and extrinsic human assessments ensure generated summaries meet rigorous standards of accuracy, coherence, and relevance. These approaches correlate highly with human evaluations and demonstrate adaptability across various languages and tasks. By focusing on the usefulness of summaries in downstream applications like question answering and text classification, these methods enhance the applicability of summarization across diverse domains, ensuring summaries are informative and tailored to specific user needs [36, 52, 49, 47]. Table 1 presents a detailed comparison of benchmarks utilized in the assessment of summarization methodologies, illustrating the varied approaches to measuring summary quality and effectiveness.

5 User Portrait and Personalization

5.1 Defining User Portrait

User portraits are detailed profiles encapsulating individual attributes such as personality traits, preferences, behaviors, and demographics. They are pivotal for personalized content generation, enabling systems to tailor interactions in line with user expectations. Social media data significantly enhances user portraits by providing insights into personality traits, which inform personalized marketing strategies [8]. In education, particularly within the humanities, user portraits accommodate unique learning preferences, enhancing educational tools' effectiveness [53]. In healthcare, they integrate patient data, like hospitalization records, to tailor medical content and improve care [54].

As illustrated in Figure 5, user portraits find diverse applications across various domains, underscoring their role in enhancing insights within social media, improving educational tools, and facilitating healthcare integration. The construction of user portraits requires comparative analyses of human-generated summaries, often revealing the absence of attributes like scope and quality statements in automated metadata. This underscores the need for structured annotations to improve dialogue summarization and user understanding [9]. Surveys highlight the scarcity of specific summary items in existing resources, emphasizing structured annotations' importance in capturing comprehensive user interactions and preferences [55].

By providing nuanced insights into individual characteristics, user portraits significantly enhance content personalization, informing the customization of interactions across various applications, including personalized language models and tailored content generation [6, 7, 8, 5]. Integrating detailed user data improves the relevance and effectiveness of personalized systems, fostering user engagement and satisfaction.

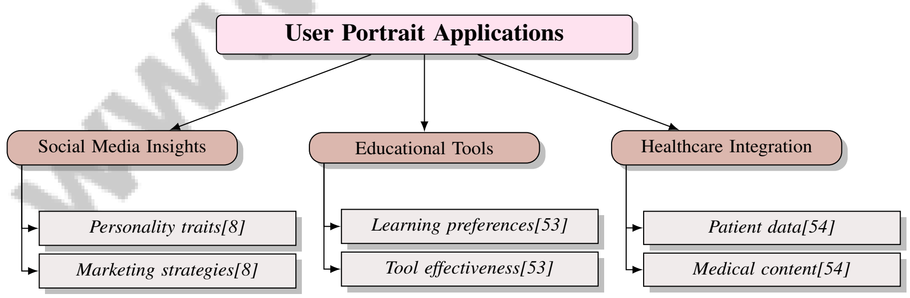


Figure 5: This figure illustrates the diverse applications of user portraits across different domains, highlighting their role in enhancing social media insights, educational tools, and healthcare integration.

5.2 Techniques for Developing User Portraits

Developing user portraits involves integrating diverse data sources and analytical techniques to create detailed profiles for personalized content generation. Analyzing social media data reveals personality traits and preferences, utilizing user-generated content on platforms like Twitter and Facebook for behavioral insights [8]. In educational contexts, techniques focus on understanding individual learning styles, particularly in the humanities, using user interaction data to tailor educational experiences

[53]. This personalization is achieved by integrating user feedback and performance data, informing educational material customization.

In healthcare, user portraits incorporate detailed patient data, such as medical histories, to generate personalized medical content and improve care [54]. Electronic health records and patient surveys enrich these portraits, offering a holistic view of patient profiles. Structured annotations and metadata further enhance understanding of user interactions and preferences. Comparative analyses of human-generated summaries highlight the significance of attributes like scope and quality statements, often missing in automated systems [9]. Incorporating these elements allows systems to deliver more nuanced, personalized content.

The techniques for developing user portraits combine data analysis, user feedback, and structured annotations, creating detailed profiles crucial for tailoring content across domains. These portraits enhance user engagement and satisfaction by aligning outputs with individual preferences and contextual needs, as shown by advancements in personalized text generation evaluation methods using large language models [6, 4].

5.3 Applications of User Portraits

User portraits, as comprehensive profiles capturing individual preferences, behaviors, and demographics, are essential for personalizing experiences across various domains. In marketing, they enable businesses to align strategies with consumer preferences, improving targeted advertising and personalized recommendations [8]. By leveraging social media and digital interactions, marketers gain insights into consumer behavior, facilitating precise audience segmentation and message customization.

In education, user portraits personalize learning experiences by accommodating diverse student learning styles and preferences, particularly in the humanities, where personalized content enhances engagement and outcomes [53]. Understanding individual needs allows educational tools to provide tailored resources, fostering effective learning environments.

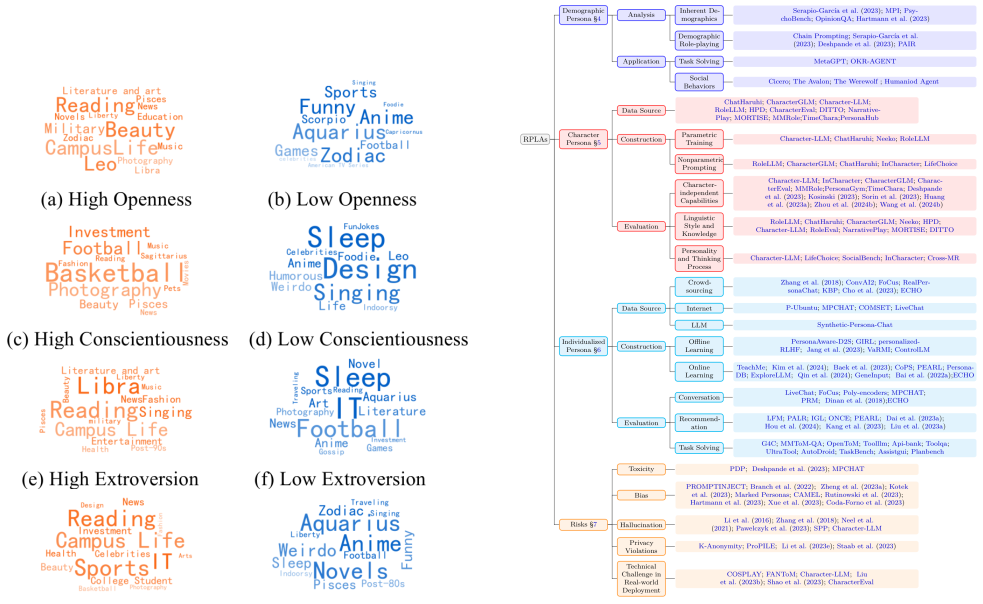
In healthcare, user portraits support patient care customization by integrating detailed patient data, such as medical histories, enabling targeted treatments and improved outcomes [54]. They also aid in developing patient-centered communication strategies, ensuring information resonates with individual patients.

In content generation, user portraits tailor media and entertainment experiences to individual preferences. Analyzing user interactions and feedback helps content creators curate personalized recommendations aligned with audience tastes. LLMOps integration into recommendation systems enhances user engagement by delivering content that aligns with preferences, increasing consumption and loyalty. Advanced machine learning techniques streamline decision-making, reduce search time, and foster deeper connections between users and platforms, as seen in services like Netflix and Amazon [56, 57].

User portraits play a crucial role in personalizing applications by enabling customization of interactions and content to align with individual preferences and needs. Constructed using historical behaviors and preferences, systems like large language models (LLMs) generate outputs reflecting users' styles and interests. This personalization enhances user experience by providing relevant recommendations, improving engagement, and facilitating content discovery across platforms like social media and e-commerce. Advanced techniques, such as embedding user-specific information into LLMs, enhance personalization while maintaining data privacy and security, driving personalized systems' effectiveness and leading to greater satisfaction and loyalty [56, 5, 6, 7, 8]. Detailed user profiles enable systems to deliver more relevant and engaging experiences, enhancing satisfaction and loyalty.

As shown in Figure 6, user portraits are pivotal for tailoring experiences and services to individual preferences and traits. The first example, depicted through word clouds, categorizes personality traits into levels of openness, conscientiousness, and extroversion, derived from words associated with activities like reading, music, and sports, serving as tools for understanding and predicting behavior and preferences. The second example addresses complexities in creating and utilizing personas within AI systems, visualized through a tree diagram exploring demographic, character, and individualized personas, while acknowledging inherent risks and challenges. Together, these

AI-generated, for reference only.



(a) Word Clouds Representing Different Personality Traits[8]

(b) Risks and Challenges in the Development and Deployment of Personas for AI Systems[21]

Figure 6: Examples of Applications of User Portraits

examples underscore user portraits' dual role in enhancing personalization through insightful analysis and careful navigation of challenges in AI applications [8, 21].

6 Applications and Case Studies

6.1 Healthcare Applications

Personalized large language models (LLMs) are revolutionizing healthcare by tailoring medical content to enhance patient care and information delivery. Utilizing user-specific data, these models refine the accuracy and pertinence of healthcare information for patients and practitioners. For example, CLIP-SynTEL benchmarks facilitate the comparison of models by generating multimodal summaries that combine textual and visual data, improving medical response accuracy [58]. This advancement allows healthcare providers to offer precise and contextually relevant information, thereby improving patient outcomes.

The PEGASUS model demonstrates the effectiveness of personalized LLMs in producing clinically acceptable medical impressions, with 89

These models also enhance patient education and engagement by providing personalized medical summaries, which improve comprehension and recall, particularly in educational contexts like mobile reading platforms [59]. This personalized approach ensures patients receive accessible and pertinent information, fostering better understanding and adherence to medical guidance.

The integration of personalized LLMs in healthcare marks a significant advancement in patient care and information dissemination. As illustrated in Figure 7, these models focus on enhancing patient care, improving patient education, and streamlining healthcare processes through multimodal summarization and personalized information delivery. By aligning medical content with individual patient needs, these models enhance healthcare service delivery. For instance, combining visual aids with textual data in the Multimodal Medical Question Summarization framework deepens the understanding of patient inquiries, while the ConSum model synthesizes psychotherapy dialogues into coherent summaries to facilitate clinical decision-making. Moreover, the NOTE system streamlines discharge summary generation, reducing clinicians' administrative burdens and improving care continuity [54, 58, 60].

AI-generated, for reference only.

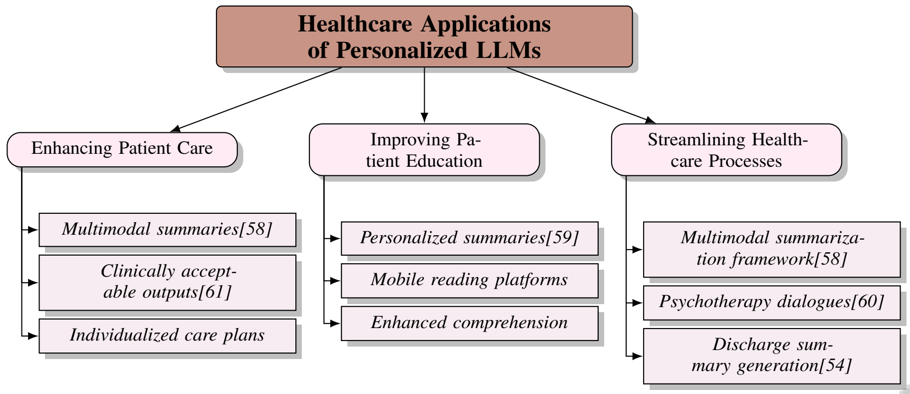


Figure 7: This figure illustrates the applications of personalized large language models (LLMs) in healthcare, focusing on enhancing patient care, improving patient education, and streamlining healthcare processes through multimodal summarization and personalized information delivery.

6.2 E-commerce and Customer Service

In e-commerce, personalized large language models (LLMs) have transformed business-customer interactions, enhancing shopping experiences and customer service. Integrating LLMs with advanced techniques, such as tensor factorization, has led to context-aware product discovery systems that capture user-object-content interactions, significantly improving the relevance of e-commerce search results [62]. These systems provide tailored product recommendations by understanding user context and preferences, enhancing customer satisfaction and engagement.

LLMs also improve customer service by developing topic-oriented dialogue summarization datasets, crucial for call center efficiency. For instance, datasets from an e-commerce company's Mandarin Chinese call center have helped train models to generate concise customer interaction summaries, enabling representatives to quickly grasp conversation essentials and resolve issues effectively [63].

The implementation of multimodal autonomous multi-agent recommendation systems (MAMARS) further illustrates LLMs' potential in delivering personalized product recommendations. These systems analyze user preferences and contextual data to create customized shopping experiences, driving sales and customer loyalty [64].

The integration of personalized LLMs in e-commerce and customer service signifies a major leap in creating tailored experiences that meet user needs. By employing advanced AI techniques, businesses can facilitate highly personalized and contextually relevant customer interactions, significantly boosting engagement, satisfaction, and retention rates [47, 57, 65].

6.3 Educational and Writing Support

Personalized large language models (LLMs) significantly impact educational tools and writing assistance by offering tailored experiences that enhance learning and writing efficiency. In educational settings, these models adapt content to individual student needs, improving engagement and comprehension. A notable application is a tutoring system designed to teach English writing concepts, showcasing how personalized LLMs can deliver customized educational content aligned with student learning styles and preferences [66].

In writing support, personalized LLMs provide advanced capabilities for generating and refining text, aiding users in producing high-quality written content. These models offer context-aware suggestions and corrections, enhancing the writing process and helping users develop their skills. For instance, the integration of LLMs in fact-checking tools has significantly improved user efficiency, with users verifying claims six times faster using AggChecker compared to traditional SQL methods, highlighting LLMs' effectiveness in streamlining fact-checking and ensuring content accuracy [67].

The role of personalized LLMs in educational and writing support is transformative, offering tailored solutions that address individual user needs. By leveraging advanced AI techniques, such as continuous automatic text summarization and personalized literature management systems, educational tools and writing assistance platforms provide real-time feedback and insights into writing processes. These innovations facilitate better content planning and structuring while enhancing engagement through tailored recommendations and concise summaries. Consequently, learners and writers are empowered to improve their writing proficiency and achieve more effective learning outcomes, ultimately transforming their educational experiences [57, 65].

7 Challenges and Future Directions

7.1 Privacy and Bias Concerns

The ethical deployment of personalized large language models (LLMs) necessitates addressing privacy and bias issues to ensure responsible use. Data bias poses a significant challenge by potentially skewing model evaluations and conclusions, with the quality of sourced data affecting the reliability of automatic metrics like ROUGE-L and their alignment with human judgment. The complexity of conversational data further complicates summarization, raising ethical concerns [42]. Copyright protection is crucial, with solutions such as PersonaMark embedding user-specific watermarks into generated text to mitigate privacy issues [7]. Additionally, inaccuracies in automatic speech recognition (ASR) can compromise summary quality [33].

Efforts like the CLEAR framework aim to enhance user privacy awareness by providing contextual information and actionable insights, though current benchmarks may still lack topic diversity and dataset representativeness, affecting the generalizability of findings [13]. Innovative approaches are required to prioritize ethical considerations, bias mitigation, and resource optimization. Enhancing summary item organization and adopting query-based methodologies can advance LLM development while safeguarding user privacy and promoting equitable outcomes. This includes creating user-specific embeddings that reflect individual preferences and historical contexts, thereby improving output relevance and accuracy while addressing challenges such as model hallucination and interpretability in information retrieval systems [68, 69, 24, 23, 5].

7.2 Resource Constraints and Scalability

The deployment of personalized LLMs is hindered by resource constraints and scalability issues, which limit their effectiveness and broader adoption. The substantial computational resources required for training and deploying these models are often inaccessible in low-resource environments [70]. The DEEP method, for example, demands more computational power than existing fine-tuned encoder models. Scalability is also challenged by the quality of input documents and user data, with methods like the Hierarchical Transformer relying heavily on input data quality, posing difficulties in processing large document clusters [11]. Extensive datasets are necessary to evaluate model scalability, highlighting the computational demands of training across diverse domains [71].

Innovative approaches such as AggChecker demonstrate automation's potential to enhance efficiency in tasks like claim verification [67]. However, capturing evolving user preferences and the diverse range of application fields remains challenging, as illustrated by the limitations of methods like IntellectSeeker [57]. Future research should explore techniques like quantization and knowledge distillation to improve model efficiency and scalability [19]. Enhancing evaluation metrics' robustness and expanding tools to address complex summarization tasks may also help overcome scalability challenges [72]. Addressing these resource and scalability issues will facilitate broader applicability and improved efficiency of personalized LLMs across various applications and environments.

7.3 Future Directions in User Portrait Research

Future directions in user portrait research, especially its integration with LLMs, offer opportunities for enhanced personalization across domains. A key area involves improving entity extraction methods and refining planning mechanisms to elevate summary quality, facilitating the creation of more accurate user portraits [73]. Developing hybrid approaches that combine user-level, persona-level, and global preference personalization is critical, necessitating enhancements in datasets and evaluation methods for comprehensive personalization [23].

Research should aim to expand benchmarks to include more diverse tasks and integrate user feedback to refine evaluation processes, thereby improving the adaptability and precision of user portraits [15].

AI-generated, for reference only.

Additionally, examining hyperparameters' impact on decoding methods and utilizing supervised information for guided summarization could further enhance user portrait accuracy [74]. Integrating reinforcement learning techniques with existing models could improve summarization quality, resulting in more nuanced and contextually relevant user portraits [35]. Moreover, research should focus on developing models capable of handling multi-modal inputs and addressing low-resource settings, broadening user portrait applicability across diverse applications [34]. This includes enhancing model capabilities in managing discourse relations and coreferences, as well as exploring methodologies for summarizing spoken dialogues [45].

Refining QA-based evaluation methods and investigating dimensions of summarization quality, such as coherence and relevance, are essential for advancing user portrait research [42]. Pursuing these research directions will significantly advance user portrait development, ensuring seamless integration with LLMs and enhancing personalization across a wide array of applications.

8 Conclusion

The exploration of personalized large language models (LLMs) and text summarization techniques underscores their transformative potential across diverse sectors. By customizing outputs based on user-specific data, personalized LLMs significantly enhance content relevance and efficacy. This is exemplified by innovative frameworks like APIDocBooster, which surpass traditional methods in terms of informativeness and relevance. Such advancements highlight the critical role of personalization in generating long-form content, as evidenced by benchmarks like LongLaMP, which pave the way for future research.

In the healthcare domain, personalized LLMs are instrumental in refining patient care documentation, offering contextually pertinent responses that address existing limitations and improve patient outcomes. Similarly, educational platforms benefit from personalized summarization methods such as HumSum, which notably increase user engagement. The fusion of LLMs with mixed reality technologies, as demonstrated by RealitySummary, further exemplifies their capacity to enrich reading experiences.

Abstractive summarization techniques, exemplified by StructSum, enhance the depth and scope of summaries while mitigating layout biases, thereby serving as robust frameworks for generating insightful content. The importance of maintaining factual accuracy and coherence in AI-generated content is accentuated by advancements in multi-source meeting summarization, which enhance contextual understanding and personalization.

Despite these developments, challenges remain, particularly in managing the risks associated with personalized LLMs to maintain societal cohesion. This survey outlines key challenges in personalization and proposes research directions to address these issues, ensuring the continued evolution of LLMs to enhance user experiences. By addressing challenges such as factual accuracy, scalability, and privacy, personalized LLMs and text summarization techniques can offer tailored and meaningful experiences, ultimately elevating user satisfaction and engagement across various applications.

References

[1] Guy Feigenblat, Chulaka Gunasekara, Benjamin Sznajder, Sachindra Joshi, David Konopnicki, and Ranit Aharonov. Tweetsumm–a dialog summarization dataset for customer service. *arXiv preprint arXiv:2111.11894*, 2021.

[2] Chaoran Chen, Daodao Zhou, Yanfang Ye, Toby Jia-jun Li, and Yaxing Yao. Clear: Towards contextual llm-empowered privacy policy analysis and risk generation for large language model applications. *arXiv preprint arXiv:2410.13387*, 2024.

[3] Hanjia Lyu, Song Jiang, Hanqing Zeng, Yinglong Xia, Qifan Wang, Si Zhang, Ren Chen, Christopher Leung, Jiajie Tang, and Jiebo Luo. Llm-rec: Personalized recommendation via prompting large language models. *arXiv preprint arXiv:2307.15780*, 2023.

[4] Sourish Dasgupta, Ankush Chander, Parth Borad, Isha Motiyani, and Tanmoy Chakraborty. Perseval: Assessing personalization in text summarizers, 2024.

[5] Jiongnan Liu, Yutao Zhu, Shuting Wang, Xiaochi Wei, Erxue Min, Yu Lu, Shuaiqiang Wang, Dawei Yin, and Zhicheng Dou. Llms + persona-plug = personalized llms, 2024.

[6] Yaqing Wang, Jiepu Jiang, Mingyang Zhang, Cheng Li, Yi Liang, Qiaozhu Mei, and Michael Bendersky. Automated evaluation of personalized text generation using large language models. *arXiv preprint arXiv:2310.11593*, 2023.

[7] Yuehan Zhang, Peizhuo Lv, Yinpeng Liu, Yongqiang Ma, Wei Lu, Xiaofeng Wang, Xiaozhong Liu, and Jiawei Liu. Personamark: Personalized llm watermarking for model protection and user attribution, 2024.

[8] Haiqian Gu, Jie Wang, Ziwen Wang, Bojin Zhuang, and Fei Su. Modeling of user portrait through social media, 2019.

[9] Laura Koesten, Elena Simperl, Emilia Kacprzak, Tom Blount, and Jeni Tennison. Everything you always wanted to know about a dataset: studies in data summarisation, 2018.

[10] Charlie George and Andreas Stuhlmüller. Factored verification: Detecting and reducing hallucination in summaries of academic papers, 2023.

[11] Yang Liu and Mirella Lapata. Hierarchical transformers for multi-document summarization, 2019.

[12] Hengchao Shang, Zongyao Li, Jiaxin Guo, Shaojun Li, Zhiqiang Rao, Yuanchang Luo, Daimeng Wei, and Hao Yang. An end-to-end speech summarization using large language model, 2024.

[13] Mai-Vu Tran, Hoang-Quynh Le, Duy-Cat Can, and Quoc-An Nguyen. Overview of the vlsp 2022 – abmusu shared task: A data challenge for vietnamese abstractive multi-document summarization, 2023.

[14] Ignacio Tampe Palma, Marcelo Mendoza, and Evangelos Milios. Neural abstractive unsupervised summarization of online news discussions, 2021.

[15] Yuyang Deng, Ni Zhao, and Xin Huang. Early chatgpt user portrait through the lens of data, 2023.

[16] Ata Çelen, Guo Han, Konrad Schindler, Luc Van Gool, Iro Armeni, Anton Obukhov, and Xi Wang. I-design: Personalized llm interior designer. *arXiv preprint arXiv:2404.02838*, 2024.

[17] Zhaoxuan Tan and Meng Jiang. User modeling in the era of large language models: Current research and future directions. *arXiv preprint arXiv:2312.11518*, 2023.

[18] Erion Çano and Ondřej Bojar. Two huge title and keyword generation corpora of research articles, 2020.

[19] Daniel Campos and ChengXiang Zhai. To asymmetry and beyond: Structured pruning of sequence to sequence models for improved inference efficiency, 2023.

[20] Jonathan Pilault, Jaehong Park, and Christopher Pal. On the impressive performance of randomly weighted encoders in summarization tasks, 2020.

[21] Jiangjie Chen, Xintao Wang, Rui Xu, Siyu Yuan, Yikai Zhang, Wei Shi, Jian Xie, Shuang Li, Ruihan Yang, Tinghui Zhu, et al. From persona to personalization: A survey on role-playing language agents. *arXiv preprint arXiv:2404.18231*, 2024.

[22] Jin Chen, Zheng Liu, Xu Huang, Chenwang Wu, Qi Liu, Gangwei Jiang, Yuanhao Pu, Yuxuan Lei, Xiaolong Chen, Xingmei Wang, et al. When large language models meet personalization: Perspectives of challenges and opportunities. *World Wide Web*, 27(4):42, 2024.

[23] Zhehao Zhang, Ryan A Rossi, Branislav Kveton, Yijia Shao, Diyi Yang, Hamed Zamani, Franck Dernoncourt, Joe Barrow, Tong Yu, Sungchul Kim, et al. Personalization of large language models: A survey. *arXiv preprint arXiv:2411.00027*, 2024.

[24] Samira Ghodratnama and Mehrdad Zakershahrak. Adapting llms for efficient, personalized information retrieval: Methods and implications. In *International Conference on Service-Oriented Computing*, pages 17–26. Springer, 2023.

[25] Hadi Askari, Anshuman Chhabra, Muhao Chen, and Prasant Mohapatra. Assessing llms for zero-shot abstractive summarization through the lens of relevance paraphrasing, 2024.

[26] Aniket Deroy, Kripabandhu Ghosh, and Saptarshi Ghosh. How ready are pre-trained abstractive models and llms for legal case judgement summarization?, 2023.

[27] Garima Chhikara, Anurag Sharma, V. Gurucharan, Kripabandhu Ghosh, and Abhijnan Chakraborty. Lamsum: Amplifying voices against harassment through llm guided extractive summarization of user incident reports, 2025.

[28] LLM Personalization Using Summarization and Arie Klaver. Bachelor computer science & datascience and artificial intelligence.

[29] Haitao Lin, Liqun Ma, Junnan Zhu, Lu Xiang, Yu Zhou, Jiajun Zhang, and Chengqing Zong. Csds: A fine-grained chinese dataset for customer service dialogue summarization, 2021.

[30] Sandeep Subramanian, Raymond Li, Jonathan Pilault, and Christopher Pal. On extractive and abstractive neural document summarization with transformer language models, 2020.

[31] Md Tahmid Rahman Laskar, Enamul Hoque, and Jimmy Xiangji Huang. Domain adaptation with pre-trained transformers for query focused abstractive text summarization, 2021.

[32] Zelalem Gero, Chandan Singh, Yiqing Xie, Sheng Zhang, Praveen Subramanian, Paul Vozila, Tristan Naumann, Jianfeng Gao, and Hoifung Poon. Attribute structuring improves llm-based evaluation of clinical text summaries, 2024.

[33] Alexandra Savelieva, Bryan Au-Yeung, and Vasanth Ramani. Abstractive summarization of spoken and written instructions with bert, 2020.

[34] Xiachong Feng, Xiaocheng Feng, and Bing Qin. A survey on dialogue summarization: Recent advances and new frontiers. *arXiv preprint arXiv:2107.03175*, 2021.

[35] Niantao Xie, Sujian Li, Huiling Ren, and Qibin Zhai. Abstractive summarization improved by wordnet-based extractive sentences, 2018.

[36] Kaiqiang Song, Bingqing Wang, Zhe Feng, and Fei Liu. A new approach to overgenerating and scoring abstractive summaries, 2021.

[37] Amanuel Alambo, Cori Lohstroh, Erik Madaus, Swati Padhee, Brandy Foster, Tanvi Banerjee, Krishnaprasad Thirunarayan, and Michael Raymer. Topic-centric unsupervised multi-document summarization of scientific and news articles, 2020.

[38] Jacob Krantz and Jugal Kalita. Abstractive summarization using attentive neural techniques, 2018.

[39] Wang Wenbo, Gao Yang, Huang Heyan, and Zhou Yuxiang. Concept pointer network for abstractive summarization, 2019.

[40] Siddhartha Banerjee, Prasenjit Mitra, and Kazunari Sugiyama. Generating abstractive summaries from meeting transcripts, 2016.

[41] Diogo Pernes, Afonso Mendes, and André F. T. Martins. Improving abstractive summarization with energy-based re-ranking, 2022.

[42] Esin Durmus, He He, and Mona Diab. Feqa: A question answering evaluation framework for faithfulness assessment in abstractive summarization, 2020.

[43] Junpeng Liu, Yanyan Zou, Hainan Zhang, Hongshen Chen, Zhuoye Ding, Caixia Yuan, and Xiaojie Wang. Topic-aware contrastive learning for abstractive dialogue summarization. *arXiv preprint arXiv:2109.04994*, 2021.

[44] Guy Feigenblat, Chulaka Gunasekara, Benjamin Sznajder, Sachindra Joshi, David Konopnicki, and Ranit Aharonov. Tweetsumm – a dialog summarization dataset for customer service, 2021.

[45] Yulong Chen, Yang Liu, Liang Chen, and Yue Zhang. Dialogsum: A real-life scenario dialogue summarization dataset. *arXiv preprint arXiv:2105.06762*, 2021.

[46] Danqing Wang, Jiaze Chen, Xianze Wu, Hao Zhou, and Lei Li. Cnewsum: A large-scale chinese news summarization dataset with human-annotated adequacy and deducibility level, 2021.

[47] Xiao Pu, Mingqi Gao, and Xiaojun Wan. Is summary useful or not? an extrinsic human evaluation of text summaries on downstream tasks, 2023.

[48] Grant C. Forbes, Parth Katlana, and Zeydy Ortiz. Metric ensembles for hallucination detection, 2023.

[49] Oleg Vasilyev, Vedant Dharnidharka, Nicholas Egan, Charlene Chambliss, and John Bohannon. Sensitivity of blanc to human-scored qualities of text summaries, 2020.

[50] Mayank Soni and Vincent Wade. Comparing abstractive summaries generated by chatgpt to real summaries through blinded reviewers and text classification algorithms, 2023.

[51] Sayali Kulkarni, Sheide Chammas, Wan Zhu, Fei Sha, and Eugene Ie. Aquamuse: Automatically generating datasets for query-based multi-document summarization, 2020.

[52] David Uthus and Jianmo Ni. Rise: Leveraging retrieval techniques for summarization evaluation, 2023.

[53] Zahra Kolagar and Alessandra Zarcone. Humsum: A personalized lecture summarization tool for humanities students using llms. In *Proceedings of the 1st Workshop on Personalization of Generative AI Systems (PERSONALIZE 2024)*, pages 36–70, 2024.

[54] Imjin Ahn, Hansle Gwon, Young-Hak Kim, Tae Joon Jun, and Sanghyun Park. Note: Notable generation of patient text summaries through efficient approach based on direct preference optimization, 2024.

[55] Don Tuggener, Margot Mieskes, Jan Milan Deriu, and Mark Cieliebak. Are we summarizing the right way?: a survey of dialogue summarization data sets. In *Conference on Empirical Methods in Natural Language Processing (EMNLP), Punta Cana, Dominican Republic (online), 7-11 November 2021*, pages 107–118. Association for Computational Linguistics, 2021.

[56] Chenxi Shi, Penghao Liang, Yichao Wu, Tong Zhan, and Zhengyu Jin. Maximizing user experience with llmops-driven personalized recommendation systems. *arXiv preprint arXiv:2404.00903*, 2024.

[57] Weizhen Bian, Siyan Liu, Yubo Zhou, Dezhi Chen, Yijie Liao, Zhenzhen Fan, and Aobo Wang. Intellectseeker: A personalized literature management system with the probabilistic model and large language model. In *International Conference on Knowledge Science, Engineering and Management*, pages 270–282. Springer, 2024.

[58] Akash Ghosh, Arkadeep Acharya, Raghav Jain, Sriparna Saha, Aman Chadha, and Setu Sinha. Clipsyntel: clip and llm synergy for multimodal question summarization in healthcare. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 22031–22039, 2024.

[59] Namrata Srivastava, Rajiv Jain, Jennifer Healey, Zoya Bylinskii, and Tilman Dingler. Mitigating the effects of reading interruptions by providing reviews and previews, 2021.

[60] Aseem Srivastava, Tharun Suresh, Sarah P Lord, Md Shad Akhtar, and Tanmoy Chakraborty. Counseling summarization using mental health knowledge guided utterance filtering. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 3920–3930, 2022.

[61] Xin Tie, Muheon Shin, Ali Pirasteh, Nevein Ibrahim, Zachary Huemann, Sharon M Castellino, Kara M Kelly, John Garrett, Junjie Hu, Steve Y Cho, et al. Personalized impression generation for pet reports using large language models. *Journal of Imaging Informatics in Medicine*, 37(2):471–488, 2024.

[62] Gaike Wang, Xin Ni, Qi Shen, and Mingxuan Yang. Leveraging large language models for context-aware product discovery in e-commerce search systems. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, 3(4):300–312, 2024.

[63] Yicheng Zou, Lujun Zhao, Yangyang Kang, Jun Lin, Minlong Peng, Zhuoren Jiang, Changlong Sun, Qi Zhang, Xuanjing Huang, and Xiaozhong Liu. Topic-oriented spoken dialogue summarization for customer service with saliency-aware topic modeling, 2021.

[64] Param Thakkar and Anushka Yadav. Personalized recommendation systems using multimodal, autonomous, multi agent systems. *arXiv preprint arXiv:2410.19855*, 2024.

[65] Hai Dang, Karim Benharrak, Florian Lehmann, and Daniel Buschek. Beyond text generation: Supporting writers with continuous automatic text summaries, 2022.

[66] Minju Park, Sojung Kim, Seunghyun Lee, Soonwoo Kwon, and Kyuseok Kim. Empowering personalized learning through a conversation-based tutoring system with student modeling. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, pages 1–10, 2024.

[67] Saehan Jo, Immanuel Trummer, Weicheng Yu, Daniel Liu, Xuezhi Wang, Cong Yu, and Niyati Mehta. Verifying text summaries of relational data sets, 2018.

[68] Frederic Kirstein, Terry Ruas, Robert Kratel, and Bela Gipp. Tell me what i need to know: Exploring llm-based (personalized) abstractive multi-source meeting summarization. *arXiv preprint arXiv:2410.14545*, 2024.

[69] Cameron Lui, Aizhan Ibraimova, David C White, Rajeshri Prashant Survase, Bachir Chihani, and Tiffany K Pham. Optimizing llm-based queries and responses using secure, portable, and personalized encoded user profile. 2024.

[70] Alex Chandler, Devesh Surve, and Hui Su. Detecting errors through ensembling prompts (deep): An end-to-end llm framework for detecting factual errors, 2024.

[71] Tohida Rehman, Raghubir Bose, Soumik Dey, and Samiran Chattopadhyay. Analysis of multidomain abstractive summarization using salience allocation, 2024.

[72] Jesse Vig, Wojciech Kryściński, Karan Goel, and Nazneen Fatema Rajani. Summvis: Interactive visual analysis of models, data, and evaluation for text summarization, 2021.

[73] Shashi Narayan, Yao Zhao, Joshua Maynez, Gonçalo Simoes, Vitaly Nikolaev, and Ryan McDonald. Planning with learned entity prompts for abstractive summarization, 2021.

[74] Janko Vidaković, Filip Karlo Došilović, and Domagoj Pluščec. Abstractive summarization as augmentation for document-level event detection, 2023.

Disclaimer:

SurveyX is an AI-powered system designed to automate the generation of surveys. While it aims to produce high-quality, coherent, and comprehensive surveys with accurate citations, the final output is derived from the AI's synthesis of pre-processed materials, which may contain limitations or inaccuracies. As such, the generated content should not be used for academic publication or formal submissions and must be independently reviewed and verified. The developers of SurveyX do not assume responsibility for any errors or consequences arising from the use of the generated surveys.

AI-generated, for reference only.