Reader-aware Writing Assistance through Reader Profiles

Ge Li
University of Bologna
Bologna, Italy
ge.li2@studio.unibo.it

Danai Vachtsevanou
University of St. Gallen
St. Gallen, Switzerland
danai.vachtsevanou@unisg.ch

Jérémy Lemée
University of St. Gallen
St. Gallen, Switzerland
jeremy.lemee@unisg.ch

Simon Mayer
University of St. Gallen
St. Gallen, Switzerland
simon.mayer@unisg.ch

Jannis Strecker
University of St. Gallen
St. Gallen, Switzerland
jannisrene.strecker@unisg.ch

ABSTRACT

Establishing rapport between authors and readers of scientific texts is essential for supporting readers in understanding texts as intended, facilitating socio-discursive practices within disciplinary communities, and helping in identifying interdisciplinary links among scientific writings. We propose a Reader-aware Congruence Assistant (RaCA), which supports writers to create texts that are adapted to target readers. Similar to user-centered design which is based on user profiles, RaCA features reader-centered writing through *reader profiles* that are dynamically computed from information discovered through academic search engines. Our assistant then leverages large language models to measure the congruence of a written text with a given reader profile, and provides feedback to the writer. We demonstrate our approach with an implemented prototype that illustrates how RaCA exploits information available on the Web to construct reader profiles, assesses writer-reader congruence and offers writers color-coded visual feedback accordingly. We argue that our approach to reader-oriented scientific writing paves the way towards the more personalized interaction of readers and writers with scientific content, and discuss how integration with Semantic Web technologies and Adaptive User Interface design can help materialize this vision within an ever-growing Web of scientific ideas, proof, and discourse.

CCS CONCEPTS

• **Information systems** → **World Wide Web**; • **Human-centered computing** → **Interactive systems and tools**; • **Computing methodologies** → *Natural language processing*.

KEYWORDS

Reader Profile, Natural Language Processing, Text Congruence, Personalized Text Adaptation

ACM Reference Format:

Ge Li, Danai Vachtsevanou, Jérémy Lemée, Simon Mayer, and Jannis Strecker. 2024. Reader-aware Writing Assistance through Reader Profiles. In *34th ACM Conference on Hypertext and Social Media (HT '24), September 10–13, 2024, Poznań, Poland*. ACM, New York, NY, USA, 7 pages.

1 INTRODUCTION

"[W]hen the economist invokes 'competition' or the ecologist utters 'niche,' they are bringing to the discussion powerful imagery that invokes neoclassical production theory with the former, identifiable components of ecosystems with the latter. These deeper meanings are only clear to the properly initiated practitioner." [41] Indeed, disciplinary keywords carry a lot of metaphorical baggage, and this observation echoes Vannevar Bush's "As We May Think" [7], which emphasizes that the development of *technologies and systems that facilitate access and linkage* of diverse knowledge domains is crucial for overcoming such barriers.

One of the challenges in the creation and consumption of (scientific) texts consequently lies in establishing congruence of a writer's text with a reader's background and experience. Here, we use the term *congruence* to highlight the association with peer-assisted learning where (cognitive or social) congruence describes the situation that *tutors and tutees share the same knowledge framework* [23]. Complicating matters, this encompasses socio-discursive practices unique to disciplinary communities [16], where research traditions, objectives [8], and knowledge hierarchy [42] vary. A related issue is the difficulty in creating texts that *bridge* across disciplines [41]: Cross-disciplinary scientific texts are hard to formulate appropriately and often remain siloed within their own domains, hindering broader comprehension and collaboration across fields.

To this end, current approaches to the creation of systems that provide assistance in the writing process focus on optimizing content creation, argument structuring, and information retrieval. This is true, for instance, for VISAR [44], Semantic Reader [22], and Autogen [30]. However, to the best of our knowledge, there is currently no writing assistant that addresses the socio-discursive relationship between writers and their potential readers.

We observe that the creation of texts that explicitly takes into account the expected, stereotypical, readership bears similarity to the *design of everyday things* [25] that takes into account the expected, stereotypical, users of these things and their needs and abilities. In that field, "good" design implies that the designed items themselves transmit cues that can be *intuitively and reliably* discovered and interpreted by users to provide guidance about *what* are the possible behaviors of the item and *how* these behaviors can be performed [38]. Humans experience this with well-designed and well-placed signs that guide them to their destination with little effort, furniture that invites them to use it, and, famously, door handles [25]—and it is this idea that established the field of User-Centered Design (UCD) [17].

In this paper, we propose taking inspiration from UCD to develop an approach to the more *reader-centered design* of texts. To this end, we propose the concept of *Reader Profiles* (RP) that contain relevant information about target readers (see Section 3). Based on such profiles, we propose the automatic evaluation of the congruence of given texts with RPs using large language models (LLM), and present a prototype of a Web-based writing interface that provides feedback about this congruence to its users (see Section 4). We furthermore discuss the potential of RPs beyond the mere evaluation of texts (see Section 5): Similar to how UCD may be coupled with run-time interface adaptation to cater not merely to user *stereotypes* but to *actual users* (cf. [38] about transferring this idea to artificial agents), RPs may be used to adapt texts for individual target users and in real time, creating a conceptual foundation for *text personalization*.

2 RELATED WORK

We briefly survey related work on (academic) writing assistants in the narrow sense introduced above, and link this field to UCD in the following.

2.1 Assisted Text Design

In the literature, already before the advent of current LLMs, writing assistants, like *Creative Help* [32], have been presented. These assist users with creative writing by tracking their edits and using a recurrent neural network (RNN)—or another fitting prediction model—to suggest new sentences. Suggestions can be generated while varying the degree of randomness in the probability distribution to stimulate creativity (cf. [4]). Within academic writing and dissemination, existing writing assistant solutions like VISAR [44], Autogen [30], Semantic Reader [22], and Magic Paper[1] provide advanced content creation, argument structuring, and information retrieval.

However, existing approaches and methodologies exhibit a notable gap in embracing the socio-discursive dynamics between writers and their intended audiences. We believe that this gap could be addressed through writing assistants that leverage UCD to enhance mutual understanding through feed-forward writer-reader communication (cf. [25]). Concretely, we foresee that *reader-oriented* text adaptation could transcend conventional content-focused enhancements by acknowledging and actively engaging with the diverse backgrounds of target readers, such as their knowledge, experience, and interests, at the time of writing or even at the time of reading (i.e., run-time text adaptation). Although this approach is particularly relevant for academic writing, reader-orientation could have the potential for much broader application, in the same way as UCD applies across domains.

2.2 UCD and Adaptive Interface Design

UCD is an approach to design that "puts human needs, capabilities, and behavior first, then designs to accommodate those needs, capabilities, and ways of behaving" [25, 26]. UCD methods form a general approach to design that is applicable to any product, ranging from manufactured physical objects to computer interfaces, and stretching across various types of content [36]. UCD hence offers a framework for creating products that enable efficient and intuitive use. Its methods not only facilitate the tailoring of digital and physical interfaces and content to suit target users during the design phase, but also at run time: To this end, *adaptive user interfaces*—i.e., any "software artifact that improves its ability to interact with a user by constructing a user model based on partial experience with that user" [19, p.358]—dynamically adjust to the needs and preferences of individual users by capturing and gauging their cognitive and sensorimotor abilities, and provide tailored user experiences that cater to a wide range of such abilities.

For the input of adaptive interfaces, *user modeling* is a widely applied method, where typically one or multiple user models are created based on stereotypical user characteristics [18]. When personal data (i.e., "any information relating to an identified or identifiable natural person" [13, p.33]) is included in the user model or the adaptation of an interface, we can speak of a *personalized interface*. In the literature, the term *personalization* is used to refer to processing of personal data as input by a personalized system, the adaptation of its functioning (i.e., the personalization algorithm), as well as the output (i.e., interfaces that convey personalized content) to the individual user (cf. [5, 14, 33, 37, 39]). Additionally to making interfaces more inclusive and accessible [21], personalization can, for instance, reduce information and option overload [5], improve experiences and communication [37], and provide better preference matching and services to users [39]. Possible *harms* of personalization, on the other hand, include a possible lack of transparency [1, 5], privacy risks [39], the creation of filter bubbles [28] and manipulation possibilities [43]—see [15, 35] for an overview.

3 TEXT ADAPTATION THROUGH READER PROFILES

Based on the discussed related work and specifically inspired by *user modeling* approaches in UCD, we conceptualize *Reader Profiles* (RP) as a means for managing information about readers, and thus, as a driver of the more reader-oriented text adaptation. Although using RPs alone does not amount to the adoption of a UCD approach—due, for example, to the potential lack of reader involvement in the creation of RPs—we argue that examining what type of information is relevant for modelling readers of scientific text is a valuable step towards taking the reader more into account during the design of scientific texts. For this, we consider that an RP comprises information representing a potential reader's domain knowledge, focuses, interests, and learning path. This profile can be created from various data sources; in a scholarly context, it could be generated to encompass not only a scholar's academic outputs but also their interests, methodologies, and the intellectual networks they engage with; it could hence be assembled from a scholar's own publications, citations, conferences they attended and/or contributed to, and their reading history; and finally, RPs may be hypermedia themselves: in this way, an RP could be linked to up-to-date resources that represent the reader's evolving research interests, and to connect the reader's intellectual map more broadly so that the RP can integrate online resources from several research communities in order to enable interdisciplinary discourse.

[1] https://magicpaper.ai/

A scholar's research focus can be discerned through their publications, which encapsulate both the primary contributions delineated within the text and the referenced citations. The content of these publications further provides insights into the scholar's chosen research topics, expertise, and preferred terminology. Simultaneously, citations establish connections to other works, potentially elucidating how the scholar's core contribution and research focus align with other works and domains. This similarly applies to citations of the scholar's publications by other authors. Citations thereby could serve to map the landscape of various research domains, providing insights into how different areas of study are interlinked, and containing information about the frontiers of individual domains [9]. Moreover, the reading history and note-taking record of a scholar could be used to map their knowledge across domains. For example, scholars often rely on bibliographic support systems like Zotero[2] to track their reading history, or use tools such as *Roam Research*[3], *Obsidian*[4], or *Logseq*[5] to keep personal knowledge records [31]. Notably, *Roam Research* allows users to construct personalized knowledge graphs with graphical data structures[6], facilitating the manual curation and interconnection of concepts for synthesizing knowledge across thoughts, projects, and domains. Similar to other types of user profiles, RPs may be represented with different formats (e.g., plain text or hypermedia formats), and may be updated at run time (we explore these aspects further in Sections 5.1 and 5.2, respectively). Finally, they are not limited to containing information about individual readers; rather, *community RPs* may be derived (e.g., in a scholarly context) from tracked conference proceedings or journal articles across authors, capturing research trends, topics, and activities on a macroscopic level [20].

We propose that RPs that capture relevant information about a reader may be used as input to a writing assistance system that evaluates the congruence between an RP and a given piece of text, proposes modifications to the text, comments the text from the perspective of the RP, or even generates RP-adapted text. To elicit requirements that users might hold on such profile-based writing assistants, we conducted an exploratory survey among 15 scholars from diverse sub-fields of computer science (ranging from Human-Computer Interaction to Multi-Agent Systems). Participants were asked to envision interacting with a system that can adapt written content to given RPs. They completed a structured survey with seven linear 5-point scale questions, rated from 1 (least useful) to 5 (most useful), designed to gauge the perceived utility of the system's features. The survey results highlighted three key features as particularly valuable (i.e., having an average rating of 4 to 5): Writing assistance systems are desired to support the *identification of frequently cited sources* and the *understanding of essential historical perspectives* relevant to a given topic, as well as to provide *terminology mapping* to enhance interdisciplinary communication by suggesting community-specific terms that align with the central concepts of the text. In addition, an open question provided space for our survey participants to express their specific needs and ideas. In this feedback, participants emphasized the need for relevant and up-to-date sources, and expressed a desire for a system that does not only adapt terminologies to audience-specific jargon but also provides direct feedback similar to the visual correction of the *Grammarly* tool[7]. These responses directly influenced our priorities in developing and implementing RPs, considering dynamic data sources, and developing features that support real-time feedback on text congruence.

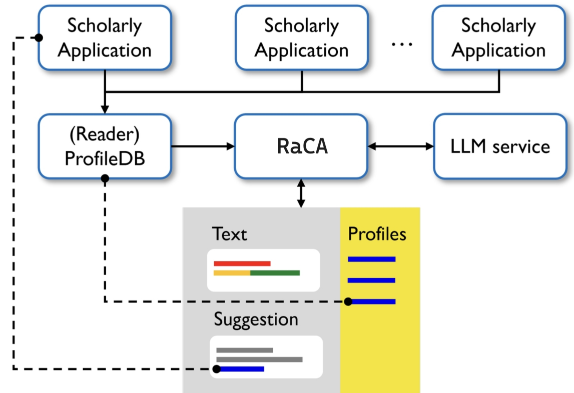

Figure 1: Diagram of the system using the Reader-aware Congruence Assistant (RaCA) to support congruence between readers' profiles (RPs) and writers' text.

4 READER-AWARE CONGRUENCE ASSISTANT

To validate our approach of evaluating and adapting texts through RPs, we implemented[8] a prototype system that we refer to as a *Reader-aware Congruence Assistant (RaCA)*. Our RaCA system leverages RPs of target readers and then provides visual feedback about the estimated text-reader congruence.

4.1 Conceptual Design of RaCA

The RaCA system (see Fig. 1) is designed to enhance textual congruence by utilizing RPs together with generative AI models (specifically, LLMs). Central to this approach is the *ProfileDB*, a curated database of reader profiles sourced from popular scholarly Web platforms like *Semantic Scholar*[9], to represent readers' academic background and research interests.

A user of RaCA enters a *text*, which may range from an abstract to a full manuscript. The text reflects the user's current discourse and is expected to contain ideas that need to be communicated clearly to a target reader. The user furthermore selects an RP from the ProfileDB, for initiating the process of analyzing the congruence between their written text and the profiled reader. The congruence analysis process employs *customized prompts* tailored to bridge the user input to the RP, and thereby uses the LLMs' trained statistical models to identify areas of semantic alignment or divergence. These prompts are tailored based on the content of the user input and the specific aspects of the selected RP (see Section 3). For example, if

[2] https://www.zotero.org/
[3] https://roamresearch.com/
[4] https://obsidian.md/
[5] https://logseq.com/
[6] https://www.zsolt.blog/2021/01/Roam-Data-Structure-Query.html
[7] https://www.grammarly.com/
[8] Our implementation is available online: https://github.com/Interactions-HSG/Reader-Aware-Coherence-Assistant
[9] https://www.semanticscholar.org/

the submitted text is a research pitch from a scholar whose domain is in computer science and who desires to efficiently communicate their ideas to a scholar with a psychology background, the prompt would focus on bridging the terminology and concepts of the two domains.

Following the congruence analysis, results are visualized by highlighting the alignment between the user input and the selected reader profile. Additionally, the results include elements of the RP that influenced any generated suggestions, such as hyperlinks to academic papers or intellectual works associated with the reader (e.g., through authorship, interaction history, etc.). Consequently, users can interpret results with respect to the factors that shaped them, receive directed guidance upon further exploring the reader's academic background, and iteratively refine their writings for improved clarity and engagement.

4.2 Proof-of-Concept Implementation

To demonstrate our approach, we implemented a system that supports academic writing with reader-oriented text evaluation according to the presented conceptual design. Our implementation uses the *Semantic Scholar* API[10] to source academic content given a scholar's ID, and feeds information obtained about the individual from this source into an RP. Resulting RPs contain data on publications, citations, and areas of expertise. For congruence analysis, we employ the OpenAI GPT-4o mini[11], which allows for real-time text analysis and adaptation based on the created RPs. RaCA includes two features to automatically analyze and measure congruence.

First, it provides an analysis of the alignment of different texts, through prompts such as: "You are an assistant for analyzing congruence between user's text: `user_input` and the following abstracts: `abstract_list`. Find the section(s) from the abstracts that match the user's text." This prompt engages the LLM to identify sections within the retrieved abstracts that closely align with the user's input, and then returns these sections as suggestions to the user.

Second, we attempt to quantify the congruence between an RP and an input text. Specifically, we use the LLM to evaluate each input text clause against the selected RP and compute a congruence score in the range $[-1, 1]$. A score of -1 indicates a significant alignment gap, suggesting that the text might likely be misunderstood or classified as irrelevant by readers who match the RP. A score of 0 suggests that the text is somewhat relevant but requires further clarification or adjustment to fully align with the reader's expectations or knowledge. A score of 1 indicates optimal congruence, indicating a high expectation that the text will be understood easily by readers who match the RP. In our prototype, the computed results are visualized for users by color-coding the respective clauses in the text: sections with particularly low congruence values relative to the RP are highlighted in red, while green highlights indicate optimal congruence. Figure 2 depicts the RaCA user interface, where a writer receives congruence analysis results for their text. For this, a prompt is submitted to the OpenAI API after being customized against the user input and RP to express the following: "Evaluate the congruence of each text clause in `{user_input}` against `{reader_profile}`. Assign a congruence score for each clause within a scale of -1, 0, 1, where -1 indicates a significant alignment gap, 0 suggests relevance with required adjustments, and 1 indicates optimal congruence."

[10] https://www.semanticscholar.org/product/api

[11] Prior to the release of GPT-4o mini (https://platform.openai.com/docs/models/gpt-4o-mini), our implementation used GPT-3.5 Turbo (https://platform.openai.com/docs/models/gpt-3-5-turbo).

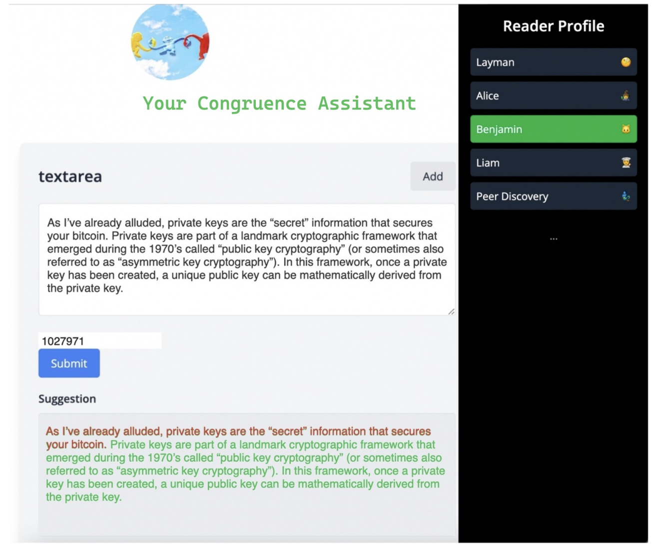

Figure 2: User interface of our proof-of-concept RaCA implementation showing color-coded feedback about writer-reader congruence based on input text and a target reader's RP. Red highlights show significant alignment gaps and green highlights indicate optimal congruence; (orange highlights would indicate moderate congruence that requires adjustments).

5 DISCUSSION AND LIMITATIONS

In our prototypical implementation of the RaCA system, RPs are constructed from static text-based data. This design choice inherently limits the system's ability to respond to the evolving academic landscape. For instance, text-based RPs can not be automatically updated to reflect recent research contributions or changes in a researcher's focus areas, emerging research trends, or personal interests. Furthermore, the lack of rich, contextual information—such as linked data that includes dynamic references to current papers, active research projects, and integrated academic social media content—limits the contextual exploration available to users.

Additionally, our implementation relies on the GPT-4o mini model for performing congruence analysis. This choice presents transparency and reliability challenges, as the model's "black box" nature obscures the logic behind its suggestions, making it difficult for users to trust or critically evaluate the AI-generated advice without additional qualitative insights. This opacity could be particularly problematic in academic settings, where understanding the rationale behind recommendations is crucial.

Our preliminary investigation of reader-oriented text adaptation focuses on providing authors with recommendations for enhancing the congruence of their scientific writings with scholars' background knowledge and research interests. However, there is room for significant enhancements towards enabling a congruence assistant to cater to a wider range of potential readers with heterogeneous characteristics, and to more flexibly support the process of scientific content writing and reading. In the remaining sections, we examine two directions in order to pursue this objective: the *use of ontologies* in order to create RPs that can capture a wide range of reader metadata, and *real-time text personalization* to facilitate the generation and consumption of scientific content.

5.1 Ontologies for Reader Profiles

We envision the process of academic knowledge dissemination not as a one-size-fits-all generation and discovery of static text documents, confined by the writer's background and preferred terminology. Instead, we conceive it as a dynamic, reader-oriented interaction tailored to readers' needs and preferences, within a Web-scale knowledge network. In this context, a software assistant capable of analyzing the alignment between texts should be able to access, reason upon, and suggest links among scientific contents and reader metadata, contributing to an ever-growing Web that is populated by scientific content and its consumers. This process aims to guide writers in navigating interconnected pieces of knowledge with respect to target readers towards increasing writer-reader alignment. Moreover, the same analysis could *additionally assist readers* in interacting with scientific content based on their individual needs and preferences.

We consider Semantic Web technologies as key to the representation and management of scientific content, and scientific content consumers. For example, Web ontologies [3] could enable the representation of RPs in machine-readable formats, e.g., in RDF formats [10]), towards automating congruence analysis. RDF-formatted RPs would allow software components, that differ from and complement LLM-based services, to conduct similar alignment analyses, while remaining adequately understandable also by people and NLP-based components. Such analyses could be, for instance, conducted by components capable of automated reasoning, that may lack natural language understanding, but could enhance transparency, explainability, and consistency during the analysis process. Additionally, employing Web ontologies would enable representing reader metadata as Web resources, which afford to be queried, navigated, and updated by software components following Linked Data principles [24]. This could facilitate the continuous discovery and management of interconnected information in a hypermedia environment for consumers of scientific works, bridging information that could otherwise remain isolated in a purely text-based RP, or the database of a single scholarly application.

We argue that reader-oriented and Web-scale text adaptation motivates the definition of a set of Web ontologies that will serve as a narrow waist for representing interlinked metadata of scientific content and readers. Minimally, an ontology is required for capturing information about readers' publications, citations, reading history, and related conferences, given the direct relation such metadata bare to a reader's scientific background and interests (as discussed in Section 3). For this, ontology suites that support representing publication metadata, like Dublin Core [12] and SPAR [29], could offer valuable input in addressing the core representation requirements of RPs. Additional ontologies could further enhance the tailoring of scientific content to readers drawing input from adaptive user interface methods (see Section 2.2): For instance, ontologies could be used to capture accessibility requirements, such as the ones considered in [11, 12, 40], or even metadata that relate to the real-time state of readers such as their current intentions (e.g., regarding their intention to gather information upon reviewing a manuscript, writing a survey, investigating research gaps etc.). Finally, such ontologies may be used to describe reader profiles but can be extended by other ontologies (existing or to be created) in order to create different types of reader profiles based on the nature of the reader (actual reader or ideal reader of a journal) and the topics related to the reader (e.g., computer science, psychology, etc.).

5.2 Towards Real-Time Text Personalization

While AI-enabled support tools for academic writing, such as Writefull[13], wisio[14], or Jenni[15] are focused on improving grammar, writing style and generic reader alignment, our RP-based approach can be extended to provide personalized text suggestions that are specifically targeted to certain (individual) readers or communities. That is, the current passive approach of providing congruence feedback through color-coding could be complemented with an assistant that actively rewrites terms and phrases to better fit the selected RP.

RPs could also be valuable in collaborative writing settings, and specifically when multiple researchers from different disciplines coedit a cross-disciplinary research paper. In this case, an extension of RaCA could in real-time highlight terms or phrases that might be understood differently by the individual researchers based on their *own* RPs. This could prevent misunderstandings and facilitate cross-disciplinary collaboration and paper writing.

Beyond providing writing support, our approach could also be implemented as a *reader support system*. Based on the *reader's* configured own RP, it could explain or paraphrase sections of a given text. While there are already similar approaches for specific reader groups such as healthcare consumers in the literature (cf. [2]), our RP-based approach could offer *personalized* explanations that are tailored to individual readers. Moving to real-time adaptation of texts, an RP-based system could furthermore be used to dynamically extend (or shrink) texts *while the reader is consuming them*. This could be implemented for instance through a system that integrates eye tracking with optical character recognition: The system could track a user's reading behavior—specifically, the dwell time of the user's eye gaze on specific words and their reading speed across different phrases (cf. [34] for a discussion of the predictive value of eye gaze for subjective text understanding); then, in combination with the user's RP, it could predict which aspects of the text are hard to understand for the user, and dynamically extend the text with further in-line explanations. This approach is conceptually similar to the explorations into adaptive hypertext in the late 1990s (cf. [6]), but enriches the discourse with adaptation based on individual RPs.

To enhance reader involvement and reader-orientation, and hence strengthen the rapport of our approach with UCD, RaCA

[12] https://www.dublincore.org/specifications/dublin-core/

[13] https://www.writefull.com/

[14] https://wisio.app/

[15] https://jenni.ai/

could allow readers to take part in the creation and management their own profiles when joining the system, and it could update them dynamically to reflect their evolving abilities and reading preferences. This could be achieved directly by granting readers greater control over their RPs or indirectly by using questionnaires to collect readers' academic focus, interests, and reading preferences, which can be updated periodically to keep profiles current. Readers could additionally choose from default settings if they prefer not to personalize their profiles, such as "Layman", "Scientist", or "Long memory 10-year-old", etc. By considering readers as active stakeholders rather than passive consumers, we aim to foster a more participatory and truly reader-centered design.

We note that the realization of such fine-grained personalization targeting individuals or small groups of people should always balance the application's utility with privacy risks to the targeted users, especially when sensitive personal data is handled. Such applications hence need to at the minimum adhere to the relevant regulations governing data handling such as the European Union's GDPR [13] or the California CPA [27].

6 CONCLUSIONS

In this paper, we present an approach for more reader-oriented scientific writing towards enhancing reader comprehension, supporting socio-discursive practices within disciplinary communities, and exploring interdisciplinary links among scientific content. Our approach offers guidance in a) the design of *reader profiles* that capture information about potential consumers of scientific work, and b) the design of a congruence assistant capable of providing directions for real-time and more reader-oriented text adaptation to enhance the alignment between (scientific) texts and target readers. Our proof-of-concept implementation demonstrates how our reader-aware congruence assistant can enable writers to convey complex ideas to their target readership in an academically coherent manner. This preliminary implementation focuses on supporting scholars' reading comprehension, by employing reader profiles that capture readers' domain-specific knowledge in the form of their publication metadata and abstracts, and an LLM-based service that evaluates the congruence between users' text input and readers' abstracts.

While the current implementation of RaCA was not utilized in the preparation of this manuscript, its potential applications in similar contexts are noteworthy. Future iterations of this system could enable authors to model RPs, including those of reviewers, to establish shared context and knowledge, which could facilitate the more effective adaptation of texts to potential audiences.

Even though our preliminary system design does not fully embrace UCD methods, it aims to initiate a discussion towards incorporating UCD principles in the creation and consumption of written texts. We argue that further integration of reader-aware text adaptation with Semantic Web technologies and adaptive interfaces could further support the dissemination of scientific content by catering to the objectives, abilities, and preferences of heterogeneous readers within a Web of scientific content. In this sense, we envision a set of Web ontologies that shall enable the design of reader profiles that can be discovered and reasoned upon with methods that complement the limitations of LLM-based congruence analysis. We argue that such research directions could further support real-time text personalization with respect to readers and writers' background, accessibility requirements, run-time behavior and objectives.

Acknowledgements. This research has received funding from the European Union's Horizon 2020 research and innovation program under grant No. 957218 (*IntellIoT*) and the Swiss National Science Foundation under grant No. 189474 (*HyperAgents*). The authors thank Thierry Sorg for his support with the implementation.

REFERENCES

[1] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for Human-AI Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland, UK) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300233

[2] Tal August, Lucy Lu Wang, Jonathan Bragg, Marti A. Hearst, Andrew Head, and Kyle Lo. 2023. Paper Plain: Making Medical Research Papers Approachable to Healthcare Consumers with Natural Language Processing. *ACM Trans. Comput.-Hum. Interact.* 30, 5, Article 74 (sep 2023), 38 pages. https://doi.org/10.1145/3589955

[3] Tim Berners-Lee, James Hendler, and Ora Lassila. 2001. The Semantic Web. *Scientific American* 284, 5 (2001), 34–43. https://www.scientificamerican.com/article/the-semantic-web/

[4] Margaret A. Boden. 2004. *The Creative Mind: Myths and Mechanisms* (2nd ed.). Routledge, London; New York.

[5] Maria Brincker. 2021. Disoriented and Alone in the "Experience Machine" – On Netflix, Shared World Deceptions and the Consequences of Deepening Algorithmic Personalization. *SATS* 22, 1 (July 2021), 75–96. https://doi.org/10.1515/sats-2021-0005

[6] Peter Brusilovsky, Alfred Kobsa, and Julita Vassileva. 1998. *Adaptive HyperText and Hypermedia.* Kluwer Academic Publishers, USA.

[7] Vannevar Bush. 1945. As We May Think. *The Atlantic* (July 1945), 8.

[8] Christopher N Candlin and Ken Hyland. 2014. *Writing: Texts, processes and practices.* Routledge.

[9] Arman Cohan, Waleed Ammar, Madeleine van Zuylen, and Field Cady. 2019. Structural Scaffolds for Citation Intent Classification in Scientific Publications. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Association for Computational Linguistics, Minneapolis, Minnesota, 3586–3596. https://doi.org/10.18653/v1/N19-1361

[10] Richard Cyganiak, David Wood, and Markus Lanthaler. 2014. *RDF 1.1 Concepts and Abstract Syntax.* W3C Recommendation. W3C. https://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/.

[11] Alexandre A. C. de Freitas, Murilo B. Scalser, Simone D. Costa, and Monalessa P. Barcellos. 2022. Towards an ontology-based approach to develop software systems with adaptive user interface. In *Proceedings of the 21st Brazilian Symposium on Human Factors in Computing Systems* (Diamantina, Brazil) *(IHC '22)*. Association for Computing Machinery, New York, NY, USA, Article 43, 7 pages. https://doi.org/10.1145/3554364.3559139

[12] Mirette Elias, Steffen Lohmann, and Sören Auer. 2016. Towards an Ontology-based Representation of Accessibility Profiles for Learners.. In *EKM@ EKAW*. CEUR Workshop Proceedings, Bologna, Italy, 51–59.

[13] European Parliament and Council of the European Union. 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union* 59, L119 (April 2016), 1–88. http://data.europa.eu/eli/reg/2016/679/oj/eng

[14] Haiyan Fan and Marshall Scott Poole. 2006. What Is Personalization? Perspectives on the Design and Implementation of Personalization in Information Systems. *Journal of Organizational Computing and Electronic Commerce* 16, 3-4 (2006), 179–202. https://doi.org/10.1080/10919392.2006.9681199 arXiv:https://doi.org/10.1080/10919392.2006.9681199

[15] Eelco Herder, Laura Stojko, Jannis Strecker, Thomas Neumayr, Enes Yigitbas, and Mirjam Augstein. 2024. Towards new realities: implications of personalized online layers in our daily lives. *i-com* (2024). https://doi.org/doi:10.1515/icom-2024-0017

[16] Ken Hyland. 2016. *Academic publishing: Issues and challenges in the construction of knowledge.* Oxford University Press.

[17] ISO 9241-210:2019 2019. *Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems.* Standard. International Organization for Standardization, Geneva, CH.

[18] Alfred Kobsa. 2001. Generic User Modeling Systems. *User Modeling and User-Adapted Interaction* 11, 1 (2001), 49–63. https://doi.org/10.1023/A:1011187500863

[19] Pat Langley. 1999. User Modeling in Adaptive Interface. In *UM99 User Modeling*, Judy Kay (Ed.). Springer Vienna, Vienna, 357–370. https://doi.org/10.1007/978-3-7091-2490-1_48

[20] Bongshin Lee, Mary Czerwinski, George Robertson, and Benjamin B. Bederson. 2005. Understanding research trends in conferences using paperLens. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (Portland, OR, USA) *(CHI EA '05)*. Association for Computing Machinery, New York, NY, USA, 1969–1972. https://doi.org/10.1145/1056808.1057069

[21] Xiangdong Li, Yunzhan Zhou, Wenqian Chen, Preben Hansen, Weidong Geng, and Lingyun Sun. 2019. 8. Towards personalized virtual reality touring through cross-object user interfaces. In *Personalized Human-Computer Interaction*. De Gruyter, 201–222. https://doi.org/10.1515/9783110552485-008

[22] Kyle Lo, Joseph Chee Chang, Andrew Head, Jonathan Bragg, Amy X. Zhang, Cassidy Trier, Chloe Anastasiades, Tal August, Russell Authur, Danielle Bragg, et al. 2023. The Semantic Reader Project: Augmenting Scholarly Documents through AI-Powered Interactive Reading Interfaces. arXiv:2303.14334 [cs.HC]

[23] Teresa Loda, Rebecca Erschens, Hannah Loenneker, Katharina E. Keifenheim, Christoph Nikendei, Florian Junne, Stephan Zipfel, and Anne Herrmann-Werner. 2019. Cognitive and social congruence in peer-assisted learning – A scoping review. *PLOS ONE* 14, 9 (09 2019), 1–15. https://doi.org/10.1371/journal.pone.0222224

[24] Ashok Malhotra, Steve Speicher, and John Arwe. 2015. *Linked Data Platform 1.0*. W3C Recommendation. W3C. https://www.w3.org/TR/2015/REC-ldp-20150226/.

[25] Don Norman. 2013. *The design of everyday things: Revised and expanded edition*. Basic books.

[26] Donald A. Norman and Stephen W. Draper. 1986. *User Centered System Design; New Perspectives on Human-Computer Interaction*. L. Erlbaum Associates Inc., USA.

[27] Office of the California Attorney General. 2018. California Consumer Privacy Act of 2018. https://www.oag.ca.gov/privacy/ccpa

[28] Eli Pariser. 2011. *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*. Penguin Books, New York, NY, USA.

[29] Silvio Peroni and David Shotton. 2018. The SPAR Ontologies. In *The Semantic Web – ISWC 2018*. Springer International Publishing, Cham, 119–136. https://doi.org/10.1007/978-3-030-00668-6_8

[30] Sebastian Porsdam Mann, Brian D. Earp, Nikolaj Møller, Suren Vynn, and Julian Savulescu. 2023. AUTOGEN: A Personalized Large Language Model for Academic Enhancement-Ethics and Proof of Principle. *The American Journal of Bioethics* 23, 10 (Oct. 2023), 28–41. https://doi.org/10.1080/15265161.2023.2233356

[31] Yvette Pyne and Stuart Stewart. 2022. Meta-Work: How We Research Is as Important as What We Research. *The British Journal of General Practice: The Journal of the Royal College of General Practitioners* 72, 716 (March 2022), 130–131. https://doi.org/10.3399/bjgp22X718757

[32] Melissa Roemmele and Andrew S. Gordon. 2018. Automated Assistance for Creative Writing with an RNN Language Model. In *Companion Proceedings of the 23rd International Conference on Intelligent User Interfaces* (Tokyo, Japan) *(IUI '18 Companion)*. Association for Computing Machinery, New York, NY, USA, Article 21, 2 pages. https://doi.org/10.1145/3180308.3180329

[33] Ville Salonen and Heikki Karjaluoto. 2016. Web personalization: The state of the art and future avenues for research and practice. *Telematics and Informatics* 33, 4 (Nov. 2016), 1088–1104. https://doi.org/10.1016/j.tele.2016.03.004

[34] Charles Lima Sanches, Olivier Augereau, and Koichi Kise. 2017. Using the Eye Gaze to Predict Document Reading Subjective Understanding. In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, Vol. 08. IEEE, Kyoto, Japan, 28–31. https://doi.org/10.1109/ICDAR.2017.377

[35] Jannis Strecker, Simon Mayer, and Kenan Bektas. 2023. Sharing Personalized Mixed Reality Experiences. Mensch und Computer 2023 - Workshopband. https://doi.org/10.18420/muc2023-mci-ws12-263

[36] Jason Chew Kit Tham, Tharon W. Howard, and Gustav Verhulsdonck. 2023. *UX Writing: Designing User-Centered Content*. Routledge, Taylor & Francis Group, New York, NY.

[37] Horst Treiblmaier and Irene Pollach. 2007. Users' Perceptions of Benefits and Costs of Personalization. In *ICIS 2007 Proceedings*. Montreal, QC, Canada, 16. http://aisel.aisnet.org/icis2007/141

[38] Danai Vachtsevanou, Andrei Ciortea, Simon Mayer, and Jérémy Lemée. 2023. Signifiers as a First-class Abstraction in Hypermedia Multi-Agent Systems. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems* (London, United Kingdom) *(AAMAS '23)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1200–1208.

[39] Jari Vesanen. 2007. What is personalization? A conceptual framework. *European Journal of Marketing* 41, 5/6 (June 2007), 409–418. https://doi.org/10.1108/03090560710737534

[40] Konstantinos Votis, Rui Lopes, Dimitrios Tzovaras, Luis Carriço, and Spiridon Likothanassis. 2009. A Semantic Accessibility Assessment Environment for Design and Development for the Web. In *Universal Access in Human-Computer Interaction. Applications and Services*, Constantine Stephanidis (Ed.). Springer, Berlin, Heidelberg, 803–813. https://doi.org/10.1007/978-3-642-02713-0_86

[41] David N. Wear. 1999. Challenges to Interdisciplinary Discourse. *Ecosystems* 2, 4 (1999), 299–301. jstor:3659022 http://www.jstor.org/stable/3659022

[42] J. M. Wolfe. 1998. Visual Memory: What Do You Know about What You Saw? *Current Biology* 8, 9 (April 1998), R303–304. https://doi.org/10.1016/s0960-9822(98)70192-7

[43] Karen Yeung. 2017. 'Hypernudge': Big Data as a Mode of Regulation by Design'. *Information, Communication & Society* 20, 1 (2017), 118–136. https://doi.org/10.1080/1369118X.2016.1186713

[44] Zheng Zhang, Jie Gao, Ranjodh Singh Dhaliwal, and Toby Jia-Jun Li. 2023. VISAR: A Human-AI Argumentative Writing Assistant with Visual Programming and Rapid Draft Prototyping. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (San Francisco, CA, USA) *(UIST '23)*. Association for Computing Machinery, New York, NY, USA, Article 5, 30 pages. https://doi.org/10.1145/3586183.3606800