WRAICOGS 2025

Writing Aids at the Crossroads of AI, Cognitive Science and NLP

Proceedings of the First Workshop

January 20, 2025

Introduction

It is a truism to say that writing is important. Indeed, it is, but it is also a highly demanding task. To achieve their goal, authors must: (a) analyze the problem, the context, and the audience; (b) consider constraints such as space and time (e.g., deadlines); (c) determine the content (messages) and its organization (coherence, outline planning); (d) choose appropriate linguistic means (lexicalization, syntax, morphology); (e) control emphasis and discourse flow (cohesion); (f) decide on the title, subtitles, and layout (e.g., chunking into paragraphs); (g) check spelling; and finally, (h) evaluate and revise, considering potential rewrites at various levels (content, form, spelling, and tone).

Writing is a special form of communication. Yet, communication relies on knowledge: knowledge of language, knowledge about the world around us, and, of course, knowledge concerning people: what do they know, typically do, feel, and believe in? In summary, writing is complex not only because of the breadth and diversity of knowledge required (e.g., domain-specific, social, cultural, and meta-knowledge), but also because of the critical need for clear, logical, and strategic thinking. While writing is not the same as thinking, it inherently requires this skill. Furthermore, it takes a toll by taxing the brain's information-processing abilities, particularly attention and memory, as intermediate results must be stored and refined. As one can see, writing is not easy, and the reasons given here above explain to some extent why so many students struggle or fail and why developing authoring aids is a logical solution. Hence, our motivation to organize this workshop.

WRAICOGS, an acronym for *Writing Aids at the Crossroads of AI, Cognitive Science, and NLP*, is a workshop dedicated to the development of writing aids aligned with human cognition. It aims to address factors such as attention and memory limitations, as well as information needs.

WRAICOGS is arguably the first event to: (a) consider the entire spectrum of writing – ideation, formulation, and revision – rather than limiting its scope to lower-level aspects such as grammar and spelling; (b) integrate humans into the development cycle of writing aids from the outset; and (c) provide support and feedback at all stages of the writing process – before, during, and after writing – rather than exclusively at the very end. Additionally, it is one of the first workshops to explore the potential applications of large language models (LLMs) across the various stages of the writing process (ideation, formulation, revision).

Finally, the workshop recognizes that writing is rarely a linear process. It is typically cyclic, involving false starts, dead ends, and varying degrees of revision. Importantly, the most critical aspect of writing lies not in the act of writing itself but in the thinking that precedes or follows the creation of the text.

These considerations guided the creation of the call for papers. We received 15 papers, and after careful review, we selected seven for presentation, of which one is non-archival and six are presented in these proceedings.

1. Buhnila et al.'s paper *"Chain of Meta Writing"* explores the potential and limitations of multilingual Small Language Models (SLMs) in assisting with writing tasks, focusing on short story writing for schoolchildren and undergraduate students in French. While SLMs can imitate certain aspects of the human writing process, such as planning and evaluation, their outputs often differ significantly from human-produced texts in terms of coherence, cohesion, and audience-appropriate vocabulary. For example, SLMs struggle with sensitive topics like school violence and they sometimes use words that are too complex for the intended group of readers. Given these facts the authors conclude that SLMs are not yet ripe enough as tools for teaching writing. This work is particularly relevant for this workshop as AI tools like ChatGPT become more integrated into education, underscoring the importance of understanding these tools' capabilities and limitations when applied to complex cognitive tasks like writing.

2. Eugeni et al.'s paper *"Reading Between the Lines"* discusses the importance of readability in writing. We generally write for a specific reader, and effective reading involves going beyond the information given. Experts read between the lines. This paper addresses the challenge of making texts more accessible to people with intellectual disabilities, particularly those with cognitive limitations, low IQs, and difficulties in reading and comprehension. It introduces a novel annotation scheme for identifying textual challenges, grounded in empirical research from psychology and translation studies. The annotated dataset consists of parallel texts (standard English and Easy Read English) available online.

3. Having stressed the importance of revision in scientific writing, the authors of *ParaRev* (Jourdan et al.) redefine the task by focusing on paragraph-level revisions. This latter is superior to sentence-level edits, which often fail to consider the broader context and discourse. Among the key contributions, we can cite: (a) *Task Redefinition*: Shifting the scope of revision from sentences to paragraphs allows for more meaningful and context-aware modifications; (b) *Improved Dataset*: Combining the original and revised scientific paragraphs with annotations improves the quality of automated revisions, regardless of the model or evaluation metric used.

4. Maggi and Vitaletti strive "towards an operative definition of creative writing." Exploring the concept of creativity in AI-generated texts, they express concerns about AI's increasing presence and its potential to replace human efforts. They suggest shifting the perception of AI from a threat to an opportunity by focusing on its creative potential, which is often misunderstood or overlooked. By changing the perspective on evaluating creative writing in AI systems, they provide a foundation for future research and help bridge the gap between AI capabilities and human creativity. Among the key findings, we can cite: (a) *Framework for Creativity*: The authors propose a measurable definition of creativity and operationalize it for evaluating texts; (b) *Comparison of Creativity in LLMs and Human-Produced Texts*: The results demonstrate that human-written texts are more creative than AI-generated ones, supporting the viability of their approach.

5. Tracing the genesis or evolution thoughts (ideation, conceptualizing) is relevant for many tasks including speaking or writing. Brain decoding technology revolutionizes the interpretation of neural activity underlying thoughts, emotions, and movements. Sato and Kobayashi's paper extends current brain decoding technology, which uses functional magnetic resonance imaging (fMRI) data to reconstruct sentences based on neural activity, by employing large language models (LLMs) as generative decoders. While the results demonstrate impressive sentence reconstruction capabilities and potential for advancing brain decoding technology, the paper's true contributions lie primarily in comparative assessments of LLMs and metrics. The lack of transparency in the training data for LLMs, apart from the fine-tuned GPT model, limits deeper analysis of performance differences. Nonetheless, the study underscores the role of text type and semantic similarity in achieving accurate brain decoding.

6. Shi and Penn deal with *Semantic Masking* (SM), a notion referring to a phenomenon where semantically coherent and contextually rich surrounding text (the "haystack") interferes with the retrieval or comprehension of specific information (the "needle") embedded within it. For example, if a piece of information is hidden within a paragraph of text that is thematically or conceptually related, the surrounding information may distract or mislead a reader trying to locate or interpret important details. Hence, SM is not merely about the length of the text but about the semantic similarity or coherence of the surrounding material. Put differently, SM is relevant both for reading and for writing. In the case of reading, it is crucial for understanding how well LLMs handle long-text scenarios, where distinguishing relevant information from semantically similar or dense contexts is a key requirement. In the case of writing, it ensures that key information is easily identifiable and comprehensible within a larger, contextually rich text.

In addition to these papers, there will be an invited talk by Cerstin Mahlow, Professor of Digital

Linguistics and Writing Research at ZHAW Zurich University of Applied Sciences, title: *'Generative AI in Writing: Redefining Collaboration, Cognition, and Creativity'* (for a summary, see here).

As always, selecting the best and most relevant submissions for the workshop was a challenging task. We would like to take this opportunity to thank all the reviewers who contributed to this effort.

Their reviews were helpful not only for us to make the decisions, but also for the authors, helping them to strengthen their work.

While the topics listed on our website are numerous, only some of them have been addressed, highlighting the need for more workshops of this kind. We hope that the work presented here will inspire you, generate fruitful discussions, and possibly lead to new ideas, insights, and collaborations.

Michael Zock, Kentaro Inui & Zheng Yuan
(organizers of the WRAICOGS workshop)

Organizing Committee

Zock, Michael (CNRS & LIS)
Inui, Kentaro (Mohamed bin Zayed University of Artificial Intelligence & Tohoku University)
Yuan, Zheng (King's College London)

Program Committee

Biemann, Chris (Universität Hamburg)
Bryant, Christopher (Writer Inc. & University of Cambridge)
Church, Ken (Northeastern University)
Cristea, Dan (University of Iasi)
Coyne, Steven (Tohoku University)
Dale, Robert (Language Technology Group)
Delmonte, Rodolfo (Università Ca' Foscari)
Ferret, Olivier (CEA LIST)
Fontenelle, Thierry (European Investment Bank)
François, Thomas (Université catholique de Louvain)
Gadeau, Gabriella (University of Cambridge)
Galván, Diana (University of Cambridge)
Gemechu, Debela (University of Dundee)
Guerraoui, Camélia (Tohoku University)
Hernandez, Nicolas (University of Nantes)
Hovy, Edward (University of Melbourne & Carnegie Mellon University)
Iacobacci, Ignacio (Elm Europe)
Ishii, Yutaka (Chiba University)
Ito, Takumi (Langsmith & Tohoku University)
Lafourcade, Mathieu (Université de Montpellier)
Langlais, Felipe (University of Montreal)
Mahlow, Cerstin (ZHAW School of Applied Linguistics)
Matsubayashi, Yuichiro (Tohoku University)
Pease, Adam (Parallax Research)
Pirrelli, Vito (University of Pisa)
Raganato, Alessandro (University of Milano-Bicocca)
Reiter, Ehud (University of Aberdeen)
Schwab, Didier (GETALP-LIG)
Strapparava, Carlo (Fondazione Bruno Kessler)
Varzandeh, Mohsen (Shiraz University of Medical Sciences)
Wanner, Leo (Universitat Pompeu Fabra)
Winniwarter, Werner (CSLEARN - Educational Technologies)

Table of Contents

Program

Monday, January 20, 2025

09:00–09:15   *Opening Remarks*

09:15–10:30   *Keynote Speech by Cerstin Mahlow*

10:30–11:00   *Morning Coffee Break*

11:00–12:30   **Oral Presentations 1**

11:00–11:45   *Chain-of-MetaWriting: Linguistic and Textual Analysis of How Small Language Models Write Young Students Texts*
Ioana Buhnila, Georgeta Cislaru and Amalia Todirascu

11:45–12:30   *Semantic Masking in a Needle-in-a-haystack Test for Evaluating Large Language Model Long-Text Capabilities*
Ken Shi and Gerald Penn

12:30–14:00   *Lunch Break*

14:00–15:30   **Oral Presentations 2**

14:00–14:45   *Reading Between the Lines: A dataset and a study on why some texts are tougher than others*
Nouran Khallaf, Carlo Eugeni and Serge Sharoff

14:45–15:30   *ParaRev : Building a dataset for Scientific Paragraph Revision annotated with revision instruction*
Léane Jourdan, Florian Boudin, Richard Dufour, Nicolas Hernandez and Akiko Aizawa

15:30–16:00   *Afternoon Coffee Break*

16:00–18:15   Oral Presentations 3

16:00–16:45   *Towards an operative definition of creative writing: a preliminary assessment of creativeness in AI and human texts*
Chiara Maggi and Andrea Vitaletti

16:45–17:30   *Decoding Semantic Representations in the Brain Under Language Stimuli with Large Language Models*
Anna Sato and Ichiro Kobayashi

17:30–18:15   *Game Plot Design with an LLM-powered Assistant: An Empirical Study with Game Designers*
Seyed Hossein Alavi, Weijia Xu, Nebojsa Jojic, Daniel Kennett, Raymond T. Ng, Sudha Rao, Haiyan Zhang, Bill Dolan and Vered Shwartz

18:15–18:30   *Closing Remarks*