IRC2024 Draft Text and Institutional Background  
University of Tartu, Estonia: Djuddah Leijen and Helen Hint

Malmö University, Sweden: Anna Wärnsby and Asko Kaupinnen

Provisional Title: The future is now: How emerging disruptive digital tools are transforming PhD students’ research writing

1. Institutional Description

The project unites two universities who are both founding members of NB!Write (The Nordic and Baltic Network for Writing in Higher Education). This project was initiated by researchers who are responsible for the teaching and training of writing to PhD students. Given the recent emergence of LLMs, such as ChatGPT, both institutions have determined a more longitudinal approach is warranted rather than a kneejerk approach to implementation of generative AI tools.

University of Tartu:

Starting this year, the University of Tartu has formed a small group of staff members across our four faculties to tackle the application and challenges of LLMs and AI tools in our educational context. The development program has set out specific practical and research-related targets regarding our various institutional needs and incentives.

1. Description of the project

The project explores how using emerging disruptive digital research writing tools – knowledge graphs and archives, digital note-taking tools, and the quickly developing generative AI tools – changes PhD students’ research writing and communication practices. As such, this project aligns with the existing global initiatives, such as *TeachAI.org*, who seek to bring together the tech and educational fields on the issue of disruptive digital tools. At the moment we lack empirically grounded evidence of how and what these tools disrupt and do not disrupt. More specifically, where we are now, the disruption seems to be driven primarily by developers of digital tools often with market-driven incentives rather than the needs and requirements of research. This makes education and research respond in a panic mode unable to align their current practices with this development. The biggest problem is that most of these panic mode responses are snapshots and not analyses of the bigger paradigm shift, often aimed at maintaining the status quo rather than accommodating and supporting the new normal. We need longitudinal empirical evidence (data collected across the duration of a PhD degree program) to develop empirically anchored and theoretically grounded responses. We use PhD students as the fulcrum to pivot an understanding of a larger shift which plays out between different stakeholders in research writing (e.g., educational institutions, program administrators, supervisors, tool builders, publishers).

1. The research questions are:
2. What relationships do PhD students develop with disruptive digital tools?
3. How do PhD students adopt these tools for their PhD research writing process?
4. What do these tools change in PhD research writing product?
5. Are there any cultural, institutional, disciplinary, or gender differences related to questions 1-3?
6. How do stakeholders respond to or influence the shift in PhD research writing?
7. A rationale for the project

PhD students and their engagement with writing is central to maintaining a sustainable relationship with research and the growing body of knowledge it produces for society and the betterment of humanity. PhD students are uniquely placed to register the push and pull of changes in academic learning and research practices. On the one hand, they are schooled to established and recognized academic practices. On the other hand, pushed to produce new knowledge, they can be open to new frameworks, methods, and tools. PhD students are open to both old and new ways of working that suit them and are therefore a meaningful group to investigate in the project. PhD research writing occurs in an educational research environment, creates new knowledge, encorporates various research methods, demonstrates critical thinking, communicates, educates, is shared with the research community, and, as any research, can be translated to form the basis for public policy. Research writing is thus an invaluable skill PhD students need to hone.

Research writing is described as recursive rather than linear (Palmeri, 2012); formal, objective, and impersonal (Bailey, 2011); and used to transform knowledge (T. Lillis & Curry, 2015). Research writing is also depicted as challenging (Boice & Jones, 1984; Elton, 2010; T. M. Lillis & Curry, 2010; Murray & Moore, 2006) because it is a) cognitively demanding (Jatin et al., 2009), b) requires a set of learned organizational strategies (Alley, 2018; Singh & Mayer, 2014), c) needs an understanding of the shared social context (Alley, 2018), d) is confusing and abstract in terms of knowledge reconstruction (Millar, 1991), e) requires a shared understanding and application of language use (T. Lillis & Curry, 2015; Soler, 2021), and f) requires rhetorical strategies that clarify without oversimplifying (Gopen & Swan, 1990). In other words, research writers balance these complex processes as they write, regardless of the “digital” intelligent technology they use to achieve these tasks.

These are precisely the challenges addressed by the disruptive digital tools – including knowledge graphs, generative AI for knowledge production and peer review, and bi-directional linking of metadata – to instantaneously connect, shape, and share knowledge while writing. Integrating these types of tools into PhD students’ knowledge production will necessarily disrupt their modes of research writing, thereby, at least at this point, fostering personal experiential relationships to academic text production and consumption (Molinari, 2021; Prechelt et al., 2018). We aim to understand this shift caused by these new tools longitudinally to identify what types of outcomes are produced by the shift.

Digital tools are transforming the status and future of research writing (Aalbersberg et al., 2012; Barbour, 2019; Edwards et al., 2016; Mingo, 2012; Santo, 2009; Yarris et al., 2020). Until very recently, Yarris et al. (2020) provide perhaps the most vivid conceptualization of what this future might look like because of this transformation: research articles become shorter, more like infographics, and supplemented with materials such as audio and video; readers become curators and collaborators of knowledge as articles are shared prepublication. Furthermore, Yarris et al. (2020) maintain that the digital transformation will make research writing more accessible, shareable, and interactive. For example, two recent disruptive academic publishing tools – Octopus and ResearchEquals – promote modular publishing strategies, where individual units of research, “research problem, hypothesis, method, results, analysis, interpretation, real-world application, and peer review,” can be published and accessed separately (Dhar, 2023). Recently, with the arrival of GPT-4 and other generative language models, the debate about the accelerated impact AI digital transformation will have on academic writing has exploded (Bridgeman et al., 2023; Curtis & others, 2023; Marche, 2023; Salvagno et al., 2023; Schäfer, 2023; Thorp, 2023; Yousr, 2023) and has sparked the development of various generative AI-driven applications (e.g., scite.ai, elicit.org, perplexity.ai) which may or may not support PhD research writing and may or may not replace many of the research tasks required. These tools are transforming research writing, knowledge retrieval, data handling, image modeling, research communication, and peer reviewing (Schäfer, 2023). They will also transform PhD education: learning and assessment practices, research processes, sharing and connecting to others’ research, research communication, and archiving. It would not be an overstatement to say that we are witnessing a paradigm shift in progress. The trends outlined above, and the very recent emergence of generative AI applications for writing, denote a fundamental problem that this project aims to address – the current misalignment between the traditional modes of research writing and research publications, and the new technological affordances and societal move to utilizing AI and other digital tools.

Research shows that men and women tend to use digital tools in different ways: they differ in “their motivations for using the software, their style of processing information, their computer self-efficacy, their attitudes toward risk, and their willingness to tinker” (Burnett et al., 2016), which may delimit the perceived usefulness of the tools for female researchers. Thus, women may be less prone to adopt and adapt to new technologies, which in the case of new disruptive tools for writing and publication may exacerbate the already existing gender gap in academia (e. g., Baker, 2012; McGee et al., 2022; Weisshaar, 2017). Moreover, since AI tools learn from historical material which is “laden with stereotypical concepts of gender,” (Leavy, 2018, p. 14) there is a danger that “the resulting application of the technology will perpetuate this bias”. We aim to understand the ongoing shift also in relation to how it may affect different genders, cultures, institutions, and disciplines and their academic practices and output.

Previous longitudinal studies within the area of PhD writing are limited in number and, to a degree in scope, as they focus on collaborative aspects of writing (Lockheart, 2010) and post-PhD writing trajectories (Castelló et al., 2021). This project expands the existing longitudinal research on PhD students by including new disruptive digital tools for connecting, sharing, and shaping knowledge. In the short term, the project seeks to illuminate what kind of disruptive and transformational potential these tools have for research writing, publication, and communication and whether there are differences across disciplines, genders, and educational institutions and cultures. In the long term, the subsequent reconfigured models of research writing will be used to prepare early and established scholars, publishers and gatekeepers, educators and educational institutions, policymakers, and other stakeholders to deal with the current realities of knowledge production and communication. The research front is moved forward by adding new data on PhD writing practices and digital tool use; documenting the paradigm shift in research writing, publishing, and communication; and exploring possible risks for gender bias in this shift. Moreover, based on this new knowledge, we aim to develop empirically grounded educational interventions and offer policy recommendation to stakeholders based on PhD students’ needs, preferences, possible gender biases, cultural and institutional constraints, as well as broad ethical concerns that need to be addressed in relation to disruptive digital tool use.

1. Conceptual Framework

To guide the project, we apply two main conceptual frameworks to investigate the transformation disruptive digital tools are exerting on PhD research writing and the various consequential relationships this has on stakeholders such as educational institutions, program administrators, supervisors, tool builders, and publishers, to name a few. The first conceptual framework, practitioner research paradigm (Cochran-Smith & Lytle, 2009), places and justifies the researchers’ unique position in the context of the project’s research objectives. The second conceptual framework, Activity Theory (Engeström et al., 1999), is applied to contextualise how multiple actors e.g., the state-of-the-art, PhD research writing, digital tools, and stakeholders form complex relationships.

To further elaborate how the project utilizes these frameworks, the practitioner research paradigm (Cochran-Smith & Lytle, 2009) merges theory with practice and researcher with practitioner roles. As writing scholars, PhD supervisors, and PhD educators ourselves, we can access and influence the ongoing shift in writing practices in real-time. Therefore, we will both explore and develop writing practices from the inside to illuminate how PhD research writing shifts with the introduction of new disruptive digital tools. We will collect etic and emic data (Pike, 1967), which will allow us to access the doings, sayings, and relatings (Kemmis et al., 2014) to the ongoing research writing shift in PhD students (who use or do not use new digital tools) in their complex educational contexts.

Activity Theory (Engeström et al., 1999) is used to explore how PhD research writing mediates through complex activities introduced by disruptive digital tools, teaching and learning, research processes, and writing processes and products (Bazerman, 1988; Bruffee, 1984; Russell, 1990). In the project, Activity Theory will be applied to contextualize the survey, interview, observation, and text data to better understand the complex relationships between PhD students and disruptive digital tools, and the changing practices of research writing and society (the broader context in which PhD research writing is placed). Furthermore, Activity Theory will be used to understand strategy in relation to digital tools, stakeholders, education, policy, and other forces which induce or hinder change (Karanasios et al., 2021; Spinuzzi, 2020). Activity Theory will be used to reconceptualize modes of PhD research writing and education based on the answers to research questions (RQs) 1-5.

Finally, in line with the longitudinal study, and the fast moving changes disruptive digital tools have on the entire PhD context, we will apply the Center for Open Science’s strategy to recommend a sustained strategy for change (Nosek, 2019) based on our empirical findings.

1. Research Method

The project’s proposed research method is a longitudinal mixed-methods study where the central objects of observation are PhD students at, at least, two research universities in Europe: Malmö University in Sweden, and the University of Tartu, in Estonia. The research and dissemination practices in the project will follow the ethical principles for research in the humanities and social sciences described in The European Code of Conduct for Research Integrity (2017), the Swedish Research Council (2017), the Estonian Research Council (2017), and follows the EU data protection regulation (GDPR).

Participants: At least 40 PhD students, across various disciplines and equally distributed across our instutions, will be recruited at Malmö University and the University of Tartu from PhD courses on academic writing and research communication, writing retreats, and various PhD programs. PhD students will be surveyed at the initial stages of the project; some of the PhD students surveyed will also participate in individual and focus-group interviews. 12 PhD students will be recruited for the longtitudinal part of the project (6 PhD students from Malmö, and 6 PhD students from Tartu). PhD students will be selected to constitute a representative sample in relation to disciplinary and gender distribution in the two universities, Malmö and Tartu. In addition, calls for PhD student volunteer participation from various other contexts (for example, partner universities connected to our networks: NB!Write, EATAW, ENLIGHT, Erasmus+) will be distributed to reqruit additional diverse participants to add to our primary target numbers from both institutions. Relevant stakeholders connected to the recruited PhD students and Malmö University and the University of Tartu (e.g., supervisors, administrators, institutional government) will also be surveyed and interviewed multiple times during the project.

Data collection instruments: The research questions will be answered with the help of a mix of established quantitative and qualitative methods: surveys, questionnaires, participant observations, focus group interviews, and text analyses. These data collection instruments will generate an array of etic and emic data (Pike, 1967), which will allow us to access the doings, sayings, and relatings (Kemmis et al., 2014) to the ongoing PhD research writing shift of the participants and the relevant stakeholders.

To answer the research questions 1 (*What relationships do PhD students develop with disruptive digital tools?*) and 2 (*How do they adopt these tools for their PhD research writing process?*), we will survey, and conduct focus group and individual interviews with PhD students to understand their developmental trajectories as knowledge creators and research writers in the ongoing shift in writing practices induced by new tools. 12 PhD students – with different disciplinary, gender, and institutional characteristics – will be specifically recruited for the longitudinal part of the study over the four years of the project’s duration. We will observe, collect think-aloud protocols, and interview these students on multiple occasions.

To answer research question 3 (*What do these tools change in PhD research writing product?*), we will collect and compare various texts produced by PhD students over the four years of the project’s duration. We will also interview and observe the PhD students and collect think-aloud protocols related to these texts’ production.

To answer research question 4 (*Are there any cultural, institutional, disciplinary, or gender differences related to questions 1-3?*), we will compare and triangulate the data collected for research questions 1-3 to illuminate emerging patterns related to cultural, institutional, disciplinary, or gender differences.

To answer research question 5 (*How do stakeholders respond to or influence the shift in PhD research writing?*), we will survey and interview PhD supervisors in focus groups or individually to gain insight into their understanding of and attitudes to the ongoing shift in writing practices induced by digital tools, and their strategies to support PhD students’ research writing trajectories in the ongoing shift. We will collect and study policy documents and interview educational administrators and governments at Malmö University and the University of Tartu and other relevant stakeholders.

Modes of analysis: The survey data will be analyzed quantitatively using linear mixed model analysis. The observation data, think-aloud protocols, and interviews will be analyzed qualitatively using thematic analysis (Braun & Clarke, 2006) within a grounded theory tradition (Strauss & Corbin, 1997). Textual data will be analysed using various text analysis methods such as Corpus Analysis (Aull et al., 2017; Dryer, 2013) and Discourse Analysis (Ädel & Reppen, 2008; Baker, 2012), Intertextuality (Groom, 2000), Rhetorical Analysis (Geisler, 2016) and Rhetorical Genre Analysis (Cotos et al., 2017; Swales, 1990).

1. Key Theorists

Activity Theory (AT) (Engeström, 1999) using Survey, interview, observation, and text data. Plus, we intent to use practitioner research paradigm (e.g. Cochran-Smith & Lytle, 2009). The research questions will, consequently, be answered with the help of a mix of established quantitative and qualitative methods: surveys, observations, interviews, and text studies. These data collection methods will generate an array of etic and emic data (Pike, 1967), which will allow us to access the doings, sayings, and relatings (Kemmis et al, 2014) to the ongoing research writing shift of the scholars (who use new digital tools), PhD students, and their supervisors.

1. Glossary

LLM – Large Language Models

Generative Artificial Intelligence

Activity Theory