

Content Repurposing in Knowledge Work: Implications for Generative AI

SIÂN LINDLEY

Microsoft Research, sianl@microsoft.com

CHRIS ELSDEN

University of Edinburgh, chris.elsden@ed.ac.uk

HANNA MOSER

Personio, hanna.schneider@ifi.lmu.de

AMID AYOBİ

University of Bristol, amid.ayobi@bristol.ac.uk.com

We propose that the use of GenAI tools in knowledge work builds on the well-established practice of content repurposing. Drawing on data from three studies into how knowledge workers reuse content, we suggest that GenAI tools have the potential to support this practice in new ways. However, there is a need for these technologies to support knowledge workers in learning through content generation, and in understanding the wider organisational context of the source materials they work with, if they are to support the production of organisational knowledge and the development of worker skill.

Additional Keywords and Phrases: remix, reuse, authorship, organisational knowledge, GenAI

1 INTRODUCTION

Generative AI (GenAI) has the potential to change knowledge work in new and unprecedented ways. Knowledge work can be defined as work undertaken to generate knowledge, and, in contrast to first-line work and manual labour, has been largely believed to be somewhat immune to advances in AI, as creativity and expertise are integral to its successful accomplishment. The discourse around recent GenAI applications, such as ChatGPT, on the one hand highlights their potential to enable this creativity and productivity, and on the other suggests that they may challenge the nature of knowledge work itself, and the job security of knowledge workers [8]. However, scholars have also argued that ML is simply the repackaging of existing human labour and expertise [4]. From this perspective, GenAI draws on a practice that is already extremely prevalent in knowledge work; that of repurposing existing content. The repurposing of personal, shared, and public content has been highlighted as a necessary strategy for knowledge workers, who deal with complex and abundant amounts of information. For instance, in [12], over half of documents, spreadsheets, and presentations were found to be related to at least one other file. In [16], it is argued that “the most urgent challenge for knowledge workers is making effective use of what information is available” [p. 2], and that constructing the best document possible out of available materials, by “pasting up”, is a common and logical approach. However, a nuanced understanding of how content repurposing is accomplished in knowledge work, and what motivates it, is currently lacking. In this paper, we draw on qualitative research into why and how people repurpose content as part of work, and what they perceive as associated risks. We consider (i) what this suggests for the design of GenAI applications, and (ii) how GenAI may support these practices in new ways.

2 RELATED WORK

Research on content repurposing encompasses investigations of ‘remix’, which is framed as a creative activity performed within online communities ([1][6][10][11][14][18]), and studies of ‘reuse’, which is often explored through analyses of workplace content ([12][16][17]). Research on the application of existing content in knowledge work has largely focused on making that content findable and understandable. Researchers have focused on the challenges raised by sharing content in organisational settings and maintaining group repositories ([23][24][19][3]), and have studied the work needed to make that content accessible to others ([2][5]). Investigations of content repurposing in the workplace, including the studies cited in the Introduction ([12][16]), tend to focus on content relationships within a single user’s files. An exception to this is [17], which examined the repurposing of presentation content in a large organization. A significant amount of exact reuse of text and images was found, but this was qualified by role; for example, material was more likely to come from a manager than a non-manager, and was strongly associated with certain

types of work. In a further survey, participants reported that information flows up and down hierarchies, with images being reused most frequently. Most of the material that was repurposed came from a person’s social network, with participants reporting uncertainty about repurposing material without knowing its origin.

Research described in [9] explores in more depth how digital content is obtained from and made available to colleagues for remix. Here it is noted that remix requires both conceptual transformation, the realization that content could be used for some other purpose, and concrete transformation, in which the semantic, structural, or stylistic content of a digital resource is changed or incorporated into something else. The conceptual work of remix might be performed by colleagues, for example, when a co-worker raises awareness of the potential for some piece of content to be appropriated. However, this tends to be done through social interaction rather than manipulation of the content in question, with the work to put it ‘in common’ (cf. [2]) being done without interacting with it. The work required to support concrete transformation of content can be considerable. A study reported in [7] highlights the importance of maintaining consistency and identifies two forms of remix: adaptive remixing, in which content is adapted to fit a new context, and creative remixing, which goes beyond translations of the material to incorporate new elements. This research highlights the need to contextualise reused materials, to make them appropriate for specific settings and audiences.

3 STUDIES

The research reported in this paper was conducted to develop a deeper understanding of how and why repurposing is accomplished by knowledge workers. We were interested in not only how easy it is to cut and paste such materials, but also in the wider set of activities that surround repurposing, such as finding and working with content. Furthermore, we expected these practices to be mediated by different motivations for content repurposing and by different types of content. These aims gave rise to the following research questions: (i) What are the different motivations for content repurposing, and how do they relate to different kinds of action? (ii) How do the material qualities of different types of content mediate repurposing, including how content is identified and worked with? (iii) What are barriers and trade-offs associated with content repurposing?

We conducted three qualitative studies with knowledge workers. The first (n=15) used diaries and interviews to explore the different activities associated with content repurposing, and the motivations for undertaking these. The second (n=19) and third (n=15) looked at the repurposing of presentation materials and text, respectively, to understand how content to be repurposed is identified and worked with, and how this varies for different forms of content. In this paper, we draw on the data collected to consider implications for emerging GenAI systems when used in the context of knowledge work.

4 FINDINGS

4.1 Intentions and actions in content repurposing

First, we present a framework of actions and intentions that are associated with content repurposing (see Figure 1). This was developed through an analysis of data from Study 1, and refined as data was collected and analysed in Studies 2 and 3. The high level actions are: (i) Adopting, which refers to the action of inserting some piece of content into a new context ‘as-is’, either by pasting it into a primary document with little or no editing, or by directly borrowing structure or style. Adoption supports two key intentions in work: ensuring *accuracy* and achieving *consistency*. (ii) Adapting, which refers to bringing some piece of content into a new context, including acting on it to ensure appropriateness for its new setting. Modifications include translating content for a new audience and updating it. Key challenges include getting content into a primary application so that it can be worked on, and tailoring it for its new context by, for example, changing dates, adding or removing personal touches, and matching formatting. Participants wished to both show and hide provenance when adapting content; sometimes it was valuable to

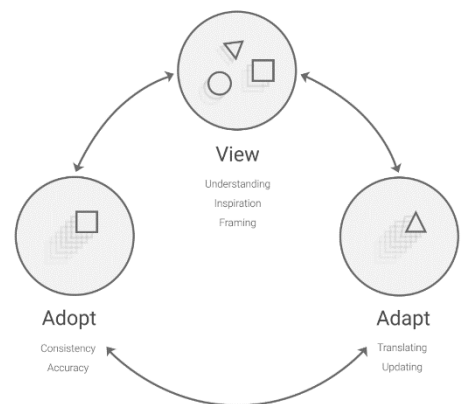


Figure 1. Framework of actions and intentions associated with content repurposing.

acknowledge origin, but in other cases this was unimportant or would undermine the impression of being personal or professional. We found two key intentions associated with adaption: *translating* materials for a new audience and *updating* content. (iii) Producing a View, which typically involves assembling or juxtaposing key pieces of content, which are compiled outside of the working document or put in place as a framework that is then deleted. These actions play an essential part in work, especially during its earlier phases, where ideas and arguments take shape. It is worth emphasizing that, while the term ‘view’ seems passive, this is a dynamic and flexible process; content is actively organized to create a particular view. Three key intentions are associated with this activity: developing *understanding*, seeking *inspiration*, and underpinning organisation by *framing* work.

4.2 Units of content

In Studies 2 and 3, we focused on two types of content respectively: presentation materials and text. Slide decks were found to be more easily compartmentalized than documents, with single slides, as well as sections of slides, being easy to repurpose. However, we also saw differences other than modularity between text and slides, which impacted how the work of repurposing was done. For example, participants found it easier to gain a visual overview of slide decks than documents, and this made it easier not only to see their style and structure, but also to understand their narrative. Structure, style and narrative are all potential components for repurposing but can be difficult to extricate from text-based documents. Interestingly, even where text is being repurposed from one document to another, we found that slide decks could play a supporting role by revealing narrative.

Working with text presents different requirements and challenges to working with presentation material. As documents are typically less visual than presentations, finding content within them was more dependent on recall, as it is harder to recognize it through visual scanning. Consequently, either authorship of, or at least high familiarity with, text was often a precursor to its repurposing, and even then, participants often forgot about relevant content until encountering it whilst looking for something else. Furthermore, ‘chunks’ of content are less obvious with text. In some cases, participants were able to cut and paste exactly what they needed, be it a word or a paragraph. But in other cases, deciding where relevant content begins and ends was less straightforward. Participants copied in extra text around that which they thought they needed, and highlighted all of it, with the expectation that they would need to work it into its new setting. A benefit of text is that it is easier to adapt for a new setting than images or slides, which are typically less pliable; ‘working in’ is necessary but easier for text.

For both presentation materials and text, content could be interpreted as offering: (i) a baseline, to be updated or translated for a new audience; (ii) a structure; (iii) elements of style; or (iv) piecemeal content that could be combined to make something new.

4.3 Risks of content repurposing

The above description of highlighting chunks of pasted-in text hints at one of the risks associated with re-use: failing to tailor content for its new context. Copy and paste errors are a potential source of embarrassment and can also have more serious consequences. In a rare example of templates being used, one of our participants described how they always generated contracts from a template rather than re-using a previously issued contract, because the risk of failing to rewrite certain details was too great.

A more general risk highlighted by our participants was that re-use could increase the potential for writing lazily, introduce bias, or even result in plagiarism. These risks were additional reasons for rewriting content and keeping track of what is pasted into a document during authoring. As one participant explained, “*You write text for a particular context, for a particular audience, for a particular style of paper, for a particular whatever. [...] I mean a lot of people do it [reuse content], but a lot of writing is bad.*” Some participants described how they sometimes deliberately avoided reuse. For them, thinking is bound up with writing and, in some cases, it is beneficial to start with an empty page.

5 DISCUSSION

Content repurposing in the workplace comprises different activities – adoption, adaption, and producing a view – each of which is intended to accomplish a different outcome. Producing a view helps build understanding and foster inspiration (cf. [21]), as well as framing work to lay out points that need to be addressed, and to assess one’s progress through an activity. Adaptation of existing materials can bring them up to date or make them suitable for a new audience (cf. [7]). Adoption supports accuracy and ensures consistency of presentation, either within a team or to a broader community. In contrast to the implication that ‘pasting up’ [16] is

a hurried practice, we found it to be effortful and careful. In this Discussion, we draw on our findings in relation to potential applications of GenAI experiences. We organise our thoughts around three emergent themes.

Working with exemplars. Knowledge workers used exemplars for multiple reasons, including to build understanding, communicate in ways that are consistent with community and organisational norms, align with specific narratives, and be certain of the accuracy of content. They were directed by colleagues to documents that illustrate how to write a proposal; they repurposed emails that illustrate how best to write a request; they reproduced code to learn something new, and they referred to and reused specific regulatory or institutional examples. There is scope for GenAI tools to offer alternatives to such exemplars, raising the question of to what extent prompts (or libraries of prompts) may replace templates, and whether, as is already the case amongst software engineers, communities of practice may craft and share these. It is worth noting here that this may be riskier for knowledge workers who are seeking to learn (and who, ironically, may benefit most from exemplars). As [21]’s analysis of programming with GitHub Copilot shows, the content that GenAI tools produce can appear superficially correct, thus errors are most difficult for novices to identify. The potential for errors also suggests that adoption of content to ascertain accuracy may be one example where GenAI tools could introduce too much risk, especially where reuse is undertaken to follow or include regulatory or institutional language.

Provenance and organisational knowledge. Repurposed content has an important quality that makes it distinct from content that is generated by AI. It is associated with particular organization members, who are seen as knowledgeable about something specific, and who may help add context or deepen understanding. GenAI tools, even when able to point to source content, may weaken this connection. On the one hand, this may help knowledge workers build on content that is hard to reach due to organisational silos. Much research has explored the challenges of disseminating organisational knowledge, and GenAI tools may offer some unique benefits here. Casting the net wide may be especially beneficial in cases of building understanding or finding inspiration. Further, the capabilities of GenAI systems to produce unique and creative outputs could be directed at new ways of supporting organisational memory. Organisational memory is bound up with ‘knowing how to ask’ and ‘knowing how to tell’; the use of memorable ‘war stories’ is key to information sharing [20]. While it is understood that prompting GenAI systems entails ‘knowing how to ask’, an additional question relates to how GenAI systems can be designed to ‘know how to tell’, by making content memorable and suited to the context of the asker.

On the other hand, this points to the risk that, by making content readily accessible across organisational boundaries, it becomes devoid of context and open to misinterpretation. Having AI perform the ‘conceptual shift’ [9] that content is ‘available’ for reuse may be especially risky in organisational settings where ‘live’ documents are the norm and where organisation members struggle to understand what they are implicitly contributing to ML systems [15]. This stands in contrast to the web, where content is made available via some ‘publish’ action. Additionally, having GenAI mediate connections between people and content may weaken the understanding of who produces what in an organisation. This is a significant aspect of organisational knowledge, which enables its understanding, socialisation, and can support future collaboration. Knowledge work entails working with, building on, and respecting the contributions of others (in contrast to the discourse around current GenAI systems, in relation to copyright and theft [13]).

Knowledge workers as experts. A final theme relates to the defining feature of knowledge work: the production of knowledge. Knowledge work is bound up with expertise and deep thinking, which is in part developed through and associated with content production. GenAI tools are sometimes interpreted as offering a ‘first draft’ or framework for further work. In this context, additional work might entail refining a written document, learning and rehearsing the speaking points for a slide deck, or auditing the content included in a synthesis. However, AI-generated frameworks might do more to enable the work that knowledge workers engage in when creating their own ‘first drafts’. For instance, they could scaffold navigation and decision-making across resources and pieces of content, helping a user ‘produce a view’ of relevant resources in support of deeper understanding of the generation of the document. The presentation of AI-generated content summaries, opposing viewpoints, calculations, and visualisations could all be produced in service of this.

6 CONCLUSION

We suggest that the use of GenAI tools in knowledge work builds on a well-established practice of content repurposing, which is performed in different ways and for different purposes. GenAI tools have the potential to support this practice in new ways, but this needs to be done in the context of knowledge work as a building of knowledge and expertise in organisational settings, where knowledge workers learn through content production, and where knowledge socialisation is accomplished through content sharing.

ACKNOWLEDGMENTS

Thanks to Richard Banks for comments on an earlier version of this paper, and to our collaborators Charlene Jennett, Gavin Smyth, Victor Poznanski, and Bernhard Kohlmeier.

REFERENCES

- [1] Yannick Assogba and Judith Donath. 2010. Share: a programming environment for loosely bound cooperation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*, 961-970. <https://doi.org/10.1145/1753326.1753469>
- [2] Liam Bannon and Susanne Bødker. 1997. Constructing common information spaces. In *Proceedings of the Fifth Conference on European Conference on Computer-Supported Cooperative Work (ECSCW'97)*, 81-96. <http://dl.acm.org/citation.cfm?id=1241980.1241986>.
- [3] Ofer Bergman, Steve Whittaker, and Noa Falk. 2014. Shared files: The retrieval perspective. *Journal of the Association for Information Science and Technology* 65(10), 1949-63. <https://doi.org/10.1002/asi.23147>
- [4] Alan F. Blackwell. 2019. Objective functions: (In)humanity and inequity in artificial intelligence. *Journal of Ethnographic Theory*, 9, 1, <https://doi.org/10.1086/703871>
- [5] John Bowers. 1994. The work to make a network work: studying CSCW in action. In *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work (CSCW '94)*, 287-298. <https://doi.org/10.1145/192844.193030>.
- [6] Giorgos Cheliotis and Jude Yew. 2009. An analysis of the social structure of remix culture. In *Proceedings of the fourth international conference on Communities and technologies (C&T '09)*, 165-174. <https://doi.org/10.1145/1556460.1556485>
- [7] Tim Coughlan, Rebecca Pitt, and Patrick McAndrew. 2013. Building open bridges: collaborative remixing and reuse of open educational resources across organisations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*, 991-1000. <https://doi.org/10.1145/2470654.2466127>
- [8] Larry Elliott. 2023. The AI industrial revolution puts middle-class workers under threat this time. *The Guardian* (18 Feb 2023). Retrieved 23 Feb 2023, from <https://www.theguardian.com/technology/2023/feb/18/the-ai-industrial-revolution-puts-middle-class-workers-under-threat-this-time>
- [9] Jessica L. Feuston and Siân E. Lindley. 2018. How social dynamics and the context of digital content impact workplace remix. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*, Paper 597, 13 pages. <https://doi.org/10.1145/3173574.3174171>
- [10] Casey Fiesler and Amy S. Bruckman. 2014. Remixers' understandings of fair use online. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*, 1023-1032. <https://doi.org/10.1145/2531602.2531695>
- [11] Casey Fiesler, Jessica L. Feuston, and Amy S. Bruckman. 2015. Understanding Copyright Law in Online Creative Communities. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*, 116-129. <https://doi.org/10.1145/2675133.2675234>
- [12] Carlos Jensen, Heather Lonsdale, Eleanor Wynn, Jill Cao, Michael Slater, and Thomas G. Dietterich. 2010. The life and times of files and information: a study of desktop provenance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*, 767-776. <https://doi.org/10.1145/1753326.1753439>
- [13] Cait Kelly. 2022. Australian artists accuse popular AI imaging app of stealing content, call for stricter copyright laws. *The Guardian* (11 December, 2022). Retrieved 1 February 2023, from <https://www.theguardian.com/australia-news/2022/dec/12/australian-artists-accuse-popular-ai-imaging-app-of-stealing-content-call-for-stricter-copyright-laws>
- [14] Lawrence Lessig. 2008. *Remix: Making art and commerce thrive in the hybrid economy*. Penguin.
- [15] Siân E. Lindley and Denise J. Wilkins. 2023. Building Knowledge through Action: Considerations for Machine Learning in the Workplace. *Transactions on Computer-Human Interaction*. <https://dl.acm.org/doi/10.1145/3584947>
- [16] Heather Lonsdale, Carlos Jensen, Eleanor Wynn, and Nicolas J. Dedual. 2010. Cutting and pasting up: "documents" and provenance in a complex work environment. In *Proceedings of the 43rd Hawaii International Conference on System Sciences*, 1-10. 10.1109/HICSS.2010.127
- [17] Yelena Mejova, Klaar De Schepper, Lawrence Bergman, and Jie Lu. 2011. Reuse in the wild: an empirical and ethnographic study of organizational content reuse. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*, 2877-2886. <https://doi.org/10.1145/1978942.1979370>
- [18] Andrés Monroy-Hernández and Mitchel Resnick. 2008. Empowering kids to create and share programmable media. *interactions* 15, 2 (March 2008), 50-53. <https://doi.org/10.1145/1340961.1340974>
- [19] Emilee Rader. 2009. Yours, mine and (not) ours: social influences on group information repositories. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*, 2095-2098. <https://doi.org/10.1145/1518701.1519019>
- [20] David Randall, John Hughes, Jon O'Brien, Mark Rouncefield, and Peter Tolmie. 2001. 'Memories are made of this': Explicating organisational knowledge and memory. *European Journal of Information Systems* 10, 2, 113-121.
- [21] Advait Sarkar, Andrew D. Gordon, Carina Negreanu, Christian Poelitz, Sruti Srinivasa Ragavan, and Ben Zorn. 2022. What is it like to program with artificial intelligence? In *Proceedings of the 33rd Annual Conference of the Psychology of Programming Interest Group (PPIG 2022)*. <https://arxiv.org/abs/2208.06213>
- [22] Moushumi Sharmin, Brian P. Bailey, Cole Coats, and Kevin Hamilton. 2009. Understanding knowledge management practices for early design activity and its implications for reuse. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*, 2367-2376. <https://doi.org/10.1145/1518701.1519064>
- [23] Amy Volda, Judith S. Olson, and Gary M. Olson. 2013. Turbulence in the clouds: challenges of cloud-based information work. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*, 2273-2282. <https://doi.org/10.1145/2470654.2481313>
- [24] Stephen Volda, W. Keith Edwards, Mark W. Newman, Rebecca E. Grinter, and Nicolas Ducheneaut. 2006. Share and share alike: exploring the user interface affordances of file sharing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*, 221-230. <http://dx.doi.org/10.1145/1124772.1124806>