How Knowledge Workers Use and Want to Use LLMs in an Enterprise Context

Michelle Brachman
michelle.brachman@ibm.com
IBM Research
Cambridge, MA, USA

Amina El-Ashry
aminaelashry@ibm.com
IBM
Bethesda, Maryland, USA

Casey Dugan
cadugan@us.ibm.com
IBM Research
Cambridge, MA, USA

Werner Geyer
werner.geyer@us.ibm.com
IBM Research
Cambridge, MA, USA

ABSTRACT

Large Language Models (LLMs) have introduced a paradigm shift in interaction with AI technology, enabling knowledge workers to complete tasks by specifying their desired outcome in natural language. LLMs have the potential to increase productivity and reduce tedious tasks in an unprecedented way. A systematic study of LLM adoption for work can provide insight into how LLMs can best support these workers. To explore knowledge workers' current and desired usage of LLMs, we ran a survey (n=216). Workers described tasks they already used LLMs for, like generating code or improving text, but imagined a future with LLMs integrated into their workflows and data. We discuss implications for adoption and design of generative AI technologies for knowledge work.

CCS CONCEPTS

• **Human-centered computing** → **Natural language interfaces**; **Empirical studies in HCI**.

KEYWORDS

survey, large language models, knowledge workers, adoption

ACM Reference Format:
Michelle Brachman, Amina El-Ashry, Casey Dugan, and Werner Geyer. 2024. How Knowledge Workers Use and Want to Use LLMs in an Enterprise Context. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '24), May 11–16, 2024, Honolulu, HI, USA*. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3613905.3650841

1 INTRODUCTION

The public release of Large Language Models (LLMs) has enabled widespread use of generative AI by end-users. The ability to specify a goal in natural language has the potential to transform knowledge work through automation and augmentation [4, 5, 38, 41]. In fact, one CEO stated about knowledge work in 2023: "I could easily see 30% of that getting replaced by AI and automation over a five-year period" [5]. Yet, accounts of knowledge workers using LLMs in the workplace are limited [43] and the evolution of generative AI based tools for knowledge work is just beginning with new tools like CoPilot for MSOffice 365[1], and Zoom AI Companion[2]. A better understanding of how knowledge workers use and want to use LLMs will provide insight into the design of LLM-based systems that support workers' needs.

Knowledge workers, who we will refer to as *workers*, are those "with high degrees of expertise, education or experience and the primary purpose of their jobs involves the creation, distribution, or application of knowledge" [12]. We focus specifically on knowledge workers in an enterprise context, specifically, a large international technology company. Despite technological advancements, there is still pressure for workers to improve productivity [24]. Increased automation can improve productivity, reducing tedious and manual tasks, but also has challenges. Knowledge work is often complex and dynamic, with workflows involving multiple people and various information and data sources [23, 30]. Our aim was to explore both the details and context of current knowledge worker adoption of LLMs and future desired LLM-based tool use, thus providing insights into workers' needs for LLM-based tools [9]. Our survey (n=216) captured workers' descriptions of current and desired use of LLMs within a large international technology company in Summer 2023. Our research questions are: 1) how do knowledge workers currently use LLMs for their work tasks and personal tasks, and how do they try out LLMs, and 2) how do knowledge workers want to use LLMs in the future for their work?

We describe four categories of LLM usage for work: for creation, to find or work with information, to get advice, and for automation. Current LLM usage for work focused on creation of artifacts and ideas, finding or learning new information, and improving existing artifacts. Workers' personal use of LLMs involved requests for ideas and guidance, even on high-risk topics. Workers hope that LLMs can support insight generation, guidance, validation, and automation of tasks. Our contributions are: 1) how knowledge workers currently use LLMs for work tasks, personal tasks, and exploration, 2) how workers want to use LLMs in the future, and 3) implications and challenges for design.

CHI EA '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0331-7/24/05
https://doi.org/10.1145/3613905.3650841

[1]https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your-copilot-for-work/
[2]https://www.zoom.com/en/blog/zoom-ai-companion/

2 RELATED WORK

Our work contributes to research on adoption of LLMs, intelligent agents at work, and knowledge workers.

2.1 Adoption of LLMs

The recent public availability of LLM technologies, such as ChatGPT [8], Bard [3], and GitHub CoPilot [19], has enabled broad swaths of the public to experiment and use LLMs in a variety of contexts. Research has identified the way people use LLMs and generative AI in a variety of contexts, such as education [54], healthcare [55], writing [22], and personal use [47, 50]. Closest to our work, an interview study of 22 information workers in the spring of 2023 found that knowledge workers were using ChatGPT for: answering questions, serving as a search engine, generating content, improving content, generating and improving code, supporting learning, and handling emails and reminders [43]. A study of an LLM programming assistant showed that users were using the tool for tasks like explaining code and recalling syntax, while they wanted to use the tool for tasks like refactoring and getting suggestions for code improvements [44]. We support and expand upon knowledge worker LLM usage, exploring how different adoption groups use LLMs differently, and exploring how knowledge workers want to use LLMs to augment and automate their work in the future.

2.2 Conversational Agents, Chatbots and Intelligent Agents at Work

Though LLMs have many applications and potential usages, ChatGPT [8] and the nature of text generation makes conversational agents a natural application of LLMs. Conversational agents, chatbots, and intelligent agents enable users to interact through natural language with a system to accomplish a variety of goals, such as information search [53], data analysis [17], and coding [44]. In the workplace, research on conversational agents has centered around intention to use and perceptions [20]. The perceptions and use of chatbots and intelligent assistants in the workplace is highly personal [14] and contextual [25]. For example, use of a chatbot for IT workers was impacted by users' understanding of the tool and use of the chatbot in rational vs. emotional ways [21]. Our results are not specific to conversational systems, but may apply to the use and design of conversational agents for knowledge workers.

2.3 Enterprise Knowledge and Information Workers

Researchers have found a variety of challenges in knowledge work. For example, work tasks are often part of larger workflows, which are often implicit, requiring mining to capture them [13, 51, 52]. Within a workflow, work is often interrupted, requiring people to leave and return to contexts [2, 36]. These contexts require many digital artifacts and resources, which workers utilize for individual tasks and across workflows [2, 27]. These challenging aspects of modern knowledge and information work can lead to stress and well-being concerns [37]. Researchers have aimed to support workers in finding and handling resources [45], coordination of work [26], email management [35], skill acquisition [31], well-being [49], and productivity [11]. However, challenges and annoyances remain in knowledge work, as well as a need to better understand how to support productivity [40]. Our work contributes to the understanding of knowledge workers by documenting the ways workers are currently attempting to use LLMs for work and how they would like to use LLMs.

3 METHODS

3.1 Survey

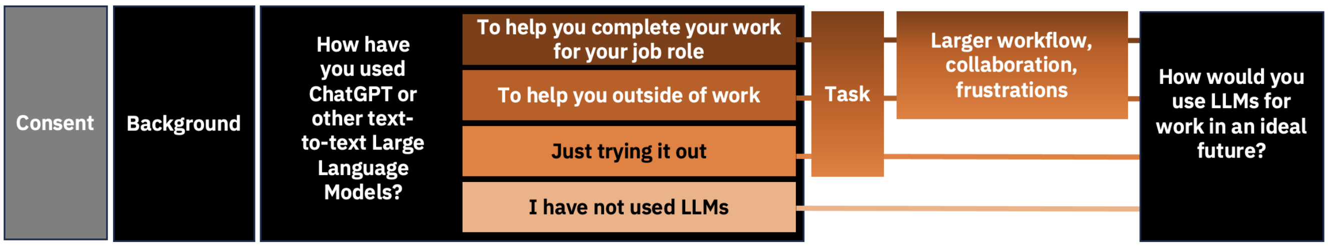
3.1.1 Design Process. We chose to use a survey for this research because we wanted to capture a large number of responses across a broad population of workers. Since the availability of LLMs for the general public is still relatively new, we also wanted to explore how workers described their current and future uses of LLMs in their own words. Thus our survey contains primarily open-ended questions and some multiple choice and Likert-scale questions. We focused our survey on LLMs with text inputs and outputs (such as ChatGPT, Bard, etc) to limit the scope. To design our questions, we were inspired by the activity checklist [29]. The activity checklist has four main concerns: means and ends, environment, learning, and development. We designed questions about knowledge workers' background and their goals in using LLMs (means and ends). Our survey also asks about the larger context and workflow around LLM usage (environment), users' frustrations in using LLMs (learning), and how workers would like to use LLMs in the future (development). This paper will focus on users' backgrounds, current use of LLMs, and desired future use of LLMs. Two authors iterated on the survey questions, including discussion with two other authors and five pilot participants before finalizing the questions. Pilot tests showed that the survey took about 15-20 minutes to complete.

3.1.2 Survey Questions. Our survey began with informed consent and collection of background data: job role, work location, experience with programming and AI, and trust in AI [28]. It then asked participants questions about their adoption of LLMs (use for work, use for personal, only trying it out, or no use, as shown in Figure 1). If they selected that they used it for work, they went down the work-use path, regardless of whatever else they selected. If they didn't select work but did select non-work, the survey directed them down the non-work path. If they did not select work or non-work LLM use, but did report that they had tried out an LLM, they were sent down the tried-it path. Participants who had not used LLMs for other purposes followed the no-use path. Participants who had used LLMs for work or personal tasks received a similar set of questions, probing at details of the tasks, while those who had only tried out LLMs or had not used them at all had similar sets of questions. Finally, the survey asked all participants about their idealized future uses of an LLM at work. The survey questions are available in the supplementary material.

3.2 Participants

We recruited participants broadly across a large international technology company through internal Slack channels from June 23, 2023 to August 1, 2023. We aimed to recruit across a wide range of technical experience, AI experience, job roles, and geographic regions. We analyzed the data from 216 participants who completed the

Figure 1: High-level survey flow, with topics discussed based on question of how they used LLMs.



survey. Our participants were primarily from North America (63%), Europe (22%) or Asia (10%). Many participants had at least partially technical roles (45%), but more than half were non-technical, with roles like design/UX (19%), sales (16%), and marketing (11%) (see Table 4). Our participants were roughly evenly split between having used LLMs for work (24.5%), for personal use but not work (27.8%), having tried out LLMs (26.9%), and not having used LLMs at all (20.8%) (see Table 1).

Table 1: Adoption groups (n=216)

Count	% Users	LLM use	Programming Exp. (1-10)	% Sig. AI work exp.
53	24.5%	*work-use*	M=6.1, SD =3.2	26%
60	27.8%	*personal-use*	M=4.6, SD=3.3	12%
58	26.9%	*tried-it*	M=5.2, SD=3	3%
45	20.8%	*no-use*	M=4.5, SD=3.2	7%

3.3 Data and Analysis

We performed a qualitative analysis of our open-ended responses to categorize the ways participants use and want to use LLMs using an inductive reflexive thematic analysis [6, 7]. We chose an inductive approach, as our study was exploratory in nature. We followed the six phases of thematic analysis: 1) familiarization with the data, 2) generate initial codes, 3) search for themes, 4) review themes, 5) define and name the themes, and 6) produce the report [6]. Two authors familiarized themselves with the data for both current and future uses of LLMs by reading through it multiple times and taking notes. They then generated initial codes, sorted the codes into themes, and collected the data for those themes. They then reviewed the themes together, checking that the data fit into the themes and revising themes when needed. To remain faithful to the reflexive thematic analysis method, we did not aim to establish inter-rater reliability but instead acknowledge the impact of the authors on the analysis, as Research Scientists and Designers at a large international technology company that has a focus on AI [6, 18, 39, 42].

4 RESULTS

We answer our research questions: 1) How do knowledge workers currently use LLMs for their work tasks and personal tasks, and how do they try out LLMs and 2) How do knowledge workers want to use LLMs in the future for their work? We describe participants by their highest level of adoption (in order from work to no-use, where we believe work use is more adoption than personal use).

4.1 RQ1: How do knowledge workers currently use LLMs for their work tasks and personal tasks, and how do they try out LLMs?

We wanted to understand the kinds of tasks workers described currently using LLMs for (171 participants: work-use, personal-use, and tried-it). We describe current LLM usages across three categories: creation, information, and advice.

4.1.1 Creation. There were two main ways users wanted LLMs to help them in creation: text they could use for their goal (creation-artifact) or ideas that they could use to further their thinking on a topic (creation-idea). Workers described generating technical artifacts, like code or commands. P135-work said that they used an LLM *"to generate unit test with lots of mock data for my code."* Current creation requests are often discrete, like SQL commands (P231-work) or an Excel script (P223-work). For non-technical work, workers used LLMs for *"'starter' draft responses [for a chatbot] that I mod in brand voice and edit for accuracy"* (P198-work) and *"to get me started on writing a job req [job requisition]"* (P255-work). Importantly, workers discussed these generations as starting points or drafts, such as P274-work who wrote: *"The code snippet didn't work as-is, but it was helpful as a starting point for my code."* Workers also sought novel ideas or brainstorming support from LLMs. P69-work said: *"We asked ChatGPT to brainstorm a list of design/research skills that would be helpful to help grow our team."* However, workers faced challenges, like in using LLMs for creative text generation: *"My main challenge is that, in doing creative work, the platforms I use try to make everything so neat and tidy, like a boring human trying to sound "correct" but not really that good or interesting"* (P143-work).

Those who used LLMs for personal use or those who had only tried out LLMs completed similar tasks as work use, but focused on using LLMs to generate ideas and creative text. Personal-use users sought ideas throughout their daily life, such as *"travel itineraries"* (P60-personal), *"cooking recipe ideas"* (P61-personal), and *"work-out schedules"* (P220-personal). Those who only tried out LLMs

Table 2: Themes and sub-themes for LLM usage

Theme	Sub-Theme	Description	Example
Creation	Artifact	Generate a new artifact to be used directly or with some modification	*To write a Python code snippet for me*
	Idea	Generate an idea, to be used indirectly	*Use for persona ideas*
Information	Search	Seek a fact or piece of information	*To find better description of the Open Source code which I study*
	Learn	Learn about a new topic more broadly	*To explain to me how different technologies work*
	Summarize	Generate a shorter version of a piece of content that describes the important elements.	*Summarize text from external websites*
	Analyze	Discover a new insight about information or data.	*Log analysis*
Advice	Improve	Generate a better version	*To re-write text that was otherwise too complex*
	Guidance	Get guidance about how to make a decision.	*Try to figure out the ideal amount of time a project should take*
	Validation	Check whether an artifact satisfies a set of rules or constraints.	*Document checking to ensure all required elements are included*
Automation	Automation	Complete a task in a piece of software with less or no human effort.	*Schedule meetings for multiple participants and rearrange their other meetings that conflict*

requested creative generations, such as *"Writing children's stories for amusement purposes"* (P141-tried). While there may be some overlap in needs between personal and work creation uses, such as needing creative capabilities, work use tasks often require knowledge that is specific and detailed, while non-work tasks use more common knowledge, such as cooking or travel.

4.1.2 Information. Workers often described using LLMs to obtain or understand information by: finding particular pieces of information (search), gaining new information about a topic (learn), obtaining a concise description of longer content (summarize), or gathering new insights about information or data (analyze). Despite the known issue of LLMs providing incorrect responses that sound correct [46], workers still used LLMs in place of search even for specialized work topics. For example, P257-work wrote: *"Getting answers on security topics instead of googling them."* One participant described their use of an LLM for learning in detail: *"Once I heard we are about to start working with Kubernetes, I immediately started a chat with ChatGPT about everything pertaining to it. It summarizes really well, and it matches itself to me in terms of complexity and professionality."* (P174-work). While this is useful in some cases, the same participant noted that they found a lack of knowledge about REST APIs, with the LLM even providing examples from a wrong API. Finally, work-use participants used LLMs to summarize lengthy public content (P258-work). Both summarization and analysis have limitations, such as token limits, that can made it challenging for an end-user to use LLMs for these use cases (P322-work). Outside of work, tried-it users asked questions as a way to evaluate the quality of LLMs. For example, P108-tried described how they experimented with LLMs as: *"basics, like answering questions."* These users knew the answers and were testing whether the LLMs could provide the correct answers.

4.1.3 Advice. Workers currently use LLMs to obtain three kinds of advice: 1) improve, in which the LLM takes a user's artifact and makes it better, 2) guide, in which the LLM makes a recommendation about how to proceed, and 3) validate, in which the LLM checks whether an artifact fulfills a set of requirements. Current LLM usage focuses on improvement of artifacts, often generally improving writing. P33-work wrote that they used an LLM to: *"help me improve my writing (e.g. make a sentence more concise or coherent, come up with a more suitable word to describe something)."* However, even in seemingly simple cases like grammar fixes, workers did not always get back what they expected: *"get frustrated if I ask only fix grammar and got back the paragraph with changed meaning"* (P123-work). Several work-use participants also used LLMs for validation or guidance with respect to bugs or errors in code or software.

Outside of work, users requested specialized guidance from LLMs on topics that typically require expertise and personalized context. Users had used LLMs for *"diagnosing a respiratory infection and researching treatment"* (P37-personal), parenting advice (P47-personal), and advice for purchasing a house (P87-personal). Considering known risks with outputs as well as privacy concerns, we were surprised that workers were using LLMs for these types of advice.

4.2 RQ2: How do knowledge workers want to use LLMs in the future for their work?

We describe four categories of desired LLM usages (creation, information, advice, and automation) for the 181 workers who envisioned future LLM uses at work (96% of work-use, 83% of personal-use, 79% of tried-it and 76% of no-use).

4.2.1 Creation. Imagined future uses of LLMs are complex, integrated into work, and include specialized skills beyond code. Workers described wanting to generate documentation and tests for code (P309-personal) and spreadsheet building (P38-nouse), rather than the code snippets users currently generate. Workers also want even further help with creating artifacts that are central to their job roles, like papers (P34-work), *"epics and user stories"* (P282-work), *"sales or training presentations"* (P263-tried), reports (P260-work), and performance reviews (P340-tried). Yet, workers still talk about using LLMs for *"a first draft"* (P353-nouse) and for writing the *"fluffy part of documents"* (P243-nouse).

4.2.2 Information. In the future, workers want to be able to perform information tasks using their own data, especially searching, summarizing, and analyzing. In an ideal future, workers want to use LLMs to search within their own data, such as *"I would love to train* *a bot with all of the guild's content so I can offer a bot to answer FAQs on Slack"* (P140-personal), or *"I would like to be able to upload all of our source code, and the companies policies and best practices into a LLM. Then I could ask it questions when other developers are gone for the day or have left the company"* (P39-work). These scenarios involve more than merely a basic input and output interaction with an LLM due to the large amounts of data. Workers also described focused summarization needs, such as support with meeting minutes (P329-personal), emails (P180-personal) and *"create an on-the-glass view based on multiple projects across multiple dashboards to show me one view with status updates, milestone dates"* (P98-personal). Workers seemed more interested in using LLMs for summarizing and analysis in the future and described using LLMs integrated into their data and workflows.

4.2.3 Advice. Workers also described future scenarios including validation and guidance based on their data. P135-work hoped that LLMs could: *"do code reviews that analyze the pull request description to check if it's properly worded, and to see if the description matches the actual changes in the files."* P33-work wants suggestions on demand: *"It might be helpful if an LLM could be embedded into those tools [for visual artifacts] and suggest ideas or improvements when I ask for them."* Workers also wanted to be able to request feedback on other specialized work, like user research tasks (P110-personal).

4.2.4 Automation. Workers described future scenarios where LLMs would perform tasks for them, requiring access to APIs or control over user interfaces. For example, workers described wanting LLMs to *"handle my calendar"* (P186-work) or provide *"'status updates' from project management activities"* (P332-personal).

5 DISCUSSION

We discuss: 1) our work in the context of prior work, 2) design implications and challenges, and 3) limitations.

5.1 Current and Future LLM Use Cases

We present an analysis of the ways knowledge workers currently use LLMs and want to use LLMs in the future. Our findings both support and expand upon previous findings. Our usage themes and sub-themes capture LLM usages described in multiple prior studies in various domains (see Table 3), providing a detailed and cohesive view of LLM use for knowledge work. The closest work to ours outlined a set of worker LLM usages that align with 6 of the 10 types of LLM usage that we describe [43]. Our results indicate that workers of a variety of roles, including non-technical ones, already use LLMs for generation of work documents, ideas, learning and finding information, and improving their writing. In the future, workers hope to be able to use LLMs to capture insights about their own data as well and automate their tasks. Both current and future tasks require that the LLM has certain abilities in order to generate valuable content or provide correct information. Finally, both current and future use of LLMs at work requires oversight.

5.2 Design Implications and Challenges for LLM-based Tools for Knowledge Workers

5.2.1 Support specialized skills. Across the tasks users described both for current and future LLM usages, specialized skills are critical,

Table 3: Support and expansion of prior work with our findings

Theme	Sub-Theme	Prior work that includes this type of LLM task
Creation	Artifact	work [43], writing [22], education [54], general [47, 48, 50]
	Idea	work [43], programming [44], general [47, 48, 50]
Information	Search	work [43], programming [44], general [47, 48, 50]
	Learn	work [43], programming [44], general [47, 48, 50]
	Summarize	writing [22]
	Analyze	education [54], healthcare [55]
Advice	Improve	work [43], programming [44], writing [22]
	Guidance	education [54], healthcare [55]
	Validation	education [54]
Automation	Automation	work [43]

meaning that 1) workers need to be able to determine if an LLM-based system has been trained on the necessary information for their task, and 2) workers may benefit from being able to customize their LLM-based tools. Transparency, such as through documentation, may help workers determine whether an LLM-based tool has critical knowledge, but is often missing or lacking necessary detail in proprietary models [34]. Reporting frameworks have been developed for AI systems [1], but users likely need more than static documentation [10, 33]. If a user is an expert in their LLM-based task, they may be able to oversee the work, but workers also use LLMs for tasks they are not experts in. Future work could explore how workers are currently evaluating LLM capabilities for tools they are not experts in. One direction may be to leverage the collaborative nature of work through social transparency [16]. Even if an LLM-based system has the capabilities necessary to support a particular task, there may be specific details necessary for a worker or company. For example, one worker found that ChatGPT provided information about Kubernetes at the correct level for them, but that same information would not necessarily be at the appropriate level for another worker. Research has begun to investigate algorithmic personalization [15, 32]. Future work could explore how to enable workers to customize LLM-based tools for their specialized contexts in cases where a prompt is not enough.

5.2.2 Support Integrated Workflows and Data. Workers described desired future LLM usage in which LLMs are integrated into their workflows and leverage their own data. During the time period we collected responses, participants in our population did not have access to integrated LLM tools that have begun to be released, such as Microsoft's CoPilot[3] for tools like Word, PowerPoint, Excel, and email or the AI companion within Zoom[4]. These tools may begin to address some of the needs and desires of workers, such as presentation and spreadsheet building, as these AI tools will have access to the user's data. One challenge in this space is that

[3]https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your-copilot-for-work/
[4]https://www.zoom.com/en/blog/zoom-ai-companion/

workers have complex, individual and contextual workflows that span multiple applications and data sources [30]. For example, a UX researcher might have interview transcripts and survey data that they want to combine, analyze, and present. Similarly, a marketing professional might need to brainstorm and create content, disperse the content, and generate and summarize reports. In these kinds of workflows, future systems will need to support users in connecting the data sources and generated outputs throughout the workflow. End-user programming may be one way to support these needs, such as 'if this then that' [5], which could allow users to import the data they need and set up their own flows. Yet, we know little about how to integrate LLM-based tools into these contexts.

5.2.3 Support LLM oversight. Across both current and future descriptions of LLM usage for work, participants described using LLMs for ideas, drafting, and as an assistant. However, due to the importance of work outputs and the unpredictable nature of generative AI, workers need to oversee LLM outputs. Overseeing the outputs could range from a quick check to extensive editing, leading to extra work regardless of whether it is tedious or a significant effort. We suggest that future systems for knowledge workers provide support for workers to oversee and modify LLM outputs in ways that minimize effort while ensuring the necessary quality of work.

5.3 Limitations

We have several limitations in our methods: 1) our population and 2) not specifying or capturing the LLMs used. We completed this survey internally at one large international technology company. While we recruited broadly across the company, including non-technical roles, many of our participants (45%) did have a technical role. Our population is also heavily North American, Western European, and Asian. Though many had knowledge or experience with LLMs, some did not, which may have limited their ability to imagine how LLMs could help them in the future. Further, the particular company we recruited from had policies in place limiting use of some LLMs for work purposes due to privacy concerns. We do not capture the particular LLMs used, in order to provide anonymity to our participants and the systems used.

6 CONCLUSION

We present a survey of knowledge workers' current uses and future desired uses of LLMs for four categories of LLM adoption. About a quarter of participants used LLMs at work (work-use), primarily requesting *creation* of templates or starting points of code or text that they then planned to modify. Outside of work (personal-use), participants often asked LLMs for *ideas* on a variety of specialized domains, while those who had only experimented with LLMs (tried-it) wanted to see how an LLM would answer questions. Participants' future visions of using LLMs involved *automation* and having the LLMs integrated into their data and workflows. We contribute to discussion of adoption of LLMs and designing LLM-based tools for workers.

[5]https://ifttt.com/

REFERENCES

[1] Matthew Arnold, Rachel KE Bellamy, Michael Hind, Stephanie Houde, Sameep Mehta, Aleksandra Mojsilović, Ravi Nair, K Natesan Ramamurthy, Alexandra Olteanu, David Piorkowski, et al. 2019. FactSheets: Increasing trust in AI services through supplier's declarations of conformity. *IBM Journal of Research and Development* 63, 4/5 (2019), 6–1.

[2] Liam Bannon, Allen Cypher, Steven Greenspan, and Melissa L Monty. 1983. Evaluation and analysis of users' activity organization. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 54–57.

[3] Bard 2023. Bard- Chat Based AI Tool from Google, Powered by PaLM 2. https://bard.google.com/. Accessed: 2023-09-02.

[4] Janine Berg and Pawel Gmyrek. 2023. Automation Hits the Knowledge Worker: ChatGPT and the Future of Work. In *UN Multi-Stakeholder Forum on Science, Technology and Innovation for the SDGs (STI Forum)*.

[5] Bloomberg 2023. IBM to Pause Hiring for Jobs that AI Could Do. https://www.bloomberg.com/news/articles/2023-05-01/ibm-to-pause-hiring-for-back-office-jobs-that-ai-could-kill. Accessed: 2023-07-28.

[6] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[7] Virginia Braun and Victoria Clarke. 2021. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative research in psychology* 18, 3 (2021), 328–352.

[8] ChatGPT 2023. ChatGPT. https://chat.openai.com/. Accessed: 2023-07-28.

[9] Parmit K Chilana, Amy J Ko, and Jacob Wobbrock. 2015. From user-centered to adoption-centered design: a case study of an HCI research innovation becoming a product. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 1749–1758.

[10] Anamaria Crisan, Margaret Drouhard, Jesse Vig, and Nazneen Rajani. 2022. Interactive model cards: A human-centered approach to model documentation. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. 427–439.

[11] Vedant Das Swain, Javier Hernandez, Brian Houck, Koustuv Saha, Jina Suh, Ahad Chaudhry, Tenny Cho, Wendy Guo, Shamsi Iqbal, and Mary P Czerwinski. 2023. Focused Time Saves Nine: Evaluating Computer–Assisted Protected Time for Hybrid Information Work. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 857, 18 pages. https://doi.org/10.1145/3544548.3581326

[12] Thomas H Davenport. 2005. *Thinking for a living: how to get better performances and results from knowledge workers*. Harvard Business Press.

[13] Claudio Di Ciccio and Massimo Mecella. 2013. Mining artful processes from knowledge workers' emails. *IEEE Internet Computing* 17, 5 (2013), 10–20.

[14] Judith Drebert, Sid Suntrayuth, and Stephan Böhm. 2023. The influence of perceived transparency on the acceptance of work task automation: an example of recruiting chatbots in Germany.

[15] Joel Eapen and VS Adhithyan. [n. d.]. Personalization and Customization of LLM Responses. ([n. d.]).

[16] Upol Ehsan, Q Vera Liao, Michael Muller, Mark O Riedl, and Justin D Weisz. 2021. Expanding explainability: Towards social transparency in ai systems. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–19.

[17] Ethan Fast, Binbin Chen, Julia Mendelsohn, Jonathan Bassen, and Michael S Bernstein. 2018. Iris: A conversational agent for complex tasks. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–12.

[18] Robert P Gauthier, Mary Jean Costello, and James R Wallace. 2022. "I Will Not Drink With You Today": A Topic-Guided Thematic Analysis of Addiction Recovery on Reddit. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–17.

[19] GitHub Copilot 2023. GitHub Copilot - Your AI pair programmer. https://github.com/features/copilot. Accessed: 2023-07-28.

[20] L Gkinko and A Elbanna. 2020. The creation of chatbots at work: an organizational perspective. *AI@ Work. ai. reshapingwork. net, Amsterdam* (2020), 5–6.

[21] Lorentsa Gkinko and Amany Elbanna. 2023. The appropriation of conversational AI in the workplace: A taxonomy of AI chatbot users. (2023).

[22] Frederic Gmeiner and Nur Yildirim. 2023. Dimensions for Designing LLM-based Writing Support. In *In2Writing Workshop at CHI*.

[23] Catherine Greene and Jeremy Myerson. 2011. Space for thought: designing for knowledge workers. *Facilities* 29, 1/2 (2011), 19–30.

[24] Hayley Guillou, Kevin Chow, Thomas Fritz, and Joanna McGrenere. 2020. Is your time well spent? reflecting on knowledge work more holistically. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–9.

[25] Jessica He, David Piorkowski, Michael Muller, Kristina Brimijoin, Stephanie Houde, and Justin D Weisz. 2023. Understanding How Task Dimensions Impact Automation Preferences with a Conversational Task Assistant.

[26] Steven Houben, Jakob E Bardram, Jo Vermeulen, Kris Luyten, and Karin Coninx. 2013. Activity-centric support for ad hoc knowledge work: A case study of co-activity manager. In *Proceedings of the SIGCHI Conference on Human Factors*

in Computing Systems. 2263–2272.

[27] Donghan Hu and Sang Won Lee. 2022. Scrapbook: Screenshot-Based Bookmarks for Effective Digital Resource Curation across Applications. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–13.

[28] Sarah A Jessup, Tamera R Schneider, Gene M Alarcon, Tyler J Ryan, and August Capiola. 2019. The measurement of the propensity to trust automation. In *Virtual, Augmented and Mixed Reality. Applications and Case Studies: 11th International Conference, VAMR 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019, Proceedings, Part II 21*. Springer, 476–489.

[29] Victor Kaptelinin, Bonnie A Nardi, and Catriona Macaulay. 1999. Methods & tools: The activity checklist: a tool for representing the "space" of context. *interactions* 6, 4 (1999), 27–39.

[30] Sandra L Kogan and Michael J Muller. 2006. Ethnographic study of collaborative knowledge work. *IBM Systems Journal* 45, 4 (2006), 759–771.

[31] Jennifer Lai, Jie Lu, Shimei Pan, Danny Soroker, Mercan Topkara, Justin Weisz, Jeff Boston, and Jason Crawford. 2014. Expediting Expertise: Supporting Informal Social Learning in the Enterprise. In *Proceedings of the 19th International Conference on Intelligent User Interfaces* (Haifa, Israel) *(IUI '14)*. Association for Computing Machinery, New York, NY, USA, 133–142. https://doi.org/10.1145/2557500.2557539

[32] Cheng Li, Mingyang Zhang, Qiaozhu Mei, Yaqing Wang, Spurthi Amba Hombaiah, Yi Liang, and Michael Bendersky. 2023. Teach LLMs to Personalize–An Approach inspired by Writing Education. *arXiv preprint arXiv:2308.07968* (2023).

[33] Q Vera Liao, Hariharan Subramonyam, Jennifer Wang, and Jennifer Wortman Vaughan. 2023. Designerly understanding: Information needs for model transparency to support design ideation for AI-powered user experience. In *Proceedings of the 2023 CHI conference on human factors in computing systems*. 1–21.

[34] Q Vera Liao and Jennifer Wortman Vaughan. 2023. AI Transparency in the Age of LLMs: A Human-Centered Research Roadmap. *arXiv preprint arXiv:2306.01941* (2023).

[35] Jie Lu, Shimei Pan, Jennifer C. Lai, and Zhen Wen. 2011. Information at Your Fingertips: Contextual IR in Enterprise Email. In *Proceedings of the 16th International Conference on Intelligent User Interfaces* (Palo Alto, CA, USA) *(IUI '11)*. Association for Computing Machinery, New York, NY, USA, 205–214. https://doi.org/10.1145/1943403.1943434

[36] Gloria Mark, Victor M Gonzalez, and Justin Harris. 2005. No task left behind? Examining the nature of fragmented work. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 321–330.

[37] Gloria Mark, Daniela Gudith, and Ulrich Klocke. 2008. The cost of interrupted work: more speed and stress. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 107–110.

[38] McKinsey 2023. Jobs lost, jobs gained: What the future of work will mean for jobs, skills, and wages. https://www.mckinsey.com/featured-insights/future-of-work/jobs-lost-jobs-gained-what-the-future-of-work-will-mean-for-jobs-skills-and-wages. Accessed: 2023-07-28.

[39] Josh Aaron Miller and Seth Cooper. 2022. Barriers to expertise in citizen science games. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–25.

[40] Miikka Palvalin. 2019. What matters for knowledge work productivity? *Employee Relations* 41, 1 (2019), 209–227.

[41] Sharon K Parker and Gudela Grote. 2022. Automation, algorithms, and beyond: Why work design matters more than ever in a digital world. *Applied Psychology* 71, 4 (2022), 1171–1204.

[42] Nathan Renney, Benedict Gaster, Tom Mitchell, and Harri Renney. 2022. Studying How Digital Luthiers Choose Their Tools. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–18.

[43] Paavo Ritala, Mika Ruokonen, and Laavanya Ramaul. 2023. Transforming boundaries: how does ChatGPT change knowledge work? *Journal of Business Strategy* (2023).

[44] Steven I Ross, Fernando Martinez, Stephanie Houde, Michael Muller, and Justin D Weisz. 2023. The programmer's assistant: Conversational interaction with a large language model for software development. In *Proceedings of the 28th International Conference on Intelligent User Interfaces*. 491–514.

[45] Jianqiang Shen, Werner Geyer, Michael Muller, Casey Dugan, Beth Brownholtz, and David R Millen. 2008. Automatically finding and recommending resources to support knowledge workers' activities. In *Proceedings of the 13th international conference on Intelligent user interfaces*. 207–216.

[46] Xinyue Shen, Zeyuan Chen, Michael Backes, and Yang Zhang. 2023. In ChatGPT We Trust? Measuring and Characterizing the Reliability of ChatGPT.

[47] Marita Skjuve. 2023. Why people use chatgpt. *Available at SSRN 4376834* (2023).

[48] Marita Skjuve, Asbjørn Følstad, and Petter Bae Brandtzaeg. 2023. The User Experience of ChatGPT: Findings from a Questionnaire Study of Early Users. In *Proceedings of the 5th International Conference on Conversational User Interfaces*. 1–10.

[49] Aaron Tabor, Scott Bateman, Erik Scheme, Book Sadprasid, and m.c. schraefel. 2021. Understanding the Design and Effectiveness of Peripheral Breathing Guide Use During Information Work. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 516, 13 pages. https://doi.org/10.1145/3411764.3445388

[50] Viriya Taecharungroj. 2023. "What Can ChatGPT Do?" Analyzing Early Reactions to the Innovative AI Chatbot on Twitter. *Big Data and Cognitive Computing* 7, 1 (2023), 35.

[51] Wil MP Van der Aalst, Boudewijn F Van Dongen, Joachim Herbst, Laura Maruster, Guido Schimm, and Anton JMM Weijters. 2003. Workflow mining: A survey of issues and approaches. *Data & knowledge engineering* 47, 2 (2003), 237–267.

[52] Ludger Van Elst, F-R Aschoff, Ansgar Bernardi, and S Schwarz. 2003. Weakly-structured workflows for knowledge-intensive tasks: An experimental evaluation. In *WET ICE 2003. Proceedings. Twelfth IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, 2003*. IEEE, 340–345.

[53] Alexandra Vtyurina, Denis Savenkov, Eugene Agichtein, and Charles LA Clarke. 2017. Exploring conversational search with humans, assistants, and wizards. In *Proceedings of the 2017 chi conference extended abstracts on human factors in computing systems*. 2187–2193.

[54] Lixiang Yan, Lele Sha, Linxuan Zhao, Yuheng Li, Roberto Martinez-Maldonado, Guanliang Chen, Xinyu Li, Yueqiao Jin, and Dragan Gašević. 2023. Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology* (2023).

[55] Yue Zheng, Laduona Wang, Baijie Feng, Ailin Zhao, and Yijun Wu. 2023. Innovating Healthcare: the role of ChatGPT in streamlining hospital workflow in the future. *Annals of Biomedical Engineering* (2023), 1–4.

APPENDIX

Table 4: Participant demographics of all participants (n=216). We list all job roles in the table with more than one participant listing them. Participants also listed several other roles, like content creation, education, audit, communications, operations, strategy, content writing, contract preparation, procurement, and proposal management.

Work Location (select one)	#	%	Job Responsibility (select all that apply)	#	%
North America	136	63%	Technical	97	45%
Europe	47	22%	Design	41	19%
Asia	22	10%	Sales	35	16%
South America	5	2%	Research	28	13%
Oceania	3	1%	Analyst	27	12%
Africa	3	1%	User Research	24	11%
AI experience (select all that apply)	#	%	Marketing	24	11%
Tried consumer AI tools	151	70%	Management	18	8%
Use consumer AI systems regularly	92	43%	Customer Service	15	7%
Closely follow AI news	86	40%	Product/project Management	12	6%
Some work experience/education	69	32%	Administrative	10	5%
Work related to AI	53	25%	Human Resources	6	3%
Significant AI work experience	26	12%	Executive	4	2%
Don't know much about it	17	8%	Finance	3	1%
			Copy editing	3	1%
			Consulting	3	1%
			Legal	2	1%
			Accounting	2	1%