

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

7,200

Open access books available

192,000

International authors and editors

210M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

14%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



Chapter

# Toward an Optimized Human-AI Reviewing Strategy for Contract Inspection

*Melanie Bancilhon, Alexa Siu, Ryan Rossi and Nedim Lipka*

## Abstract

Contracts are high-value documents that mediate many day-to-day business transactions. Knowledge workers, such as auditors and financial analysts often need to review large collections of contracts containing complex clauses. While prior work across other applications has evaluated the benefits of human-AI collaboration when dealing with large amounts of data, there is a lack of human-centered approaches for contract inspection tools. To address this gap, we present findings from qualitative interviews conducted with six knowledge workers at a large enterprise and discuss their reviewing strategies, usage of tools and perception of AI. We identify that an important but often overlooked aspect of contracts is their cross-functional use as a knowledge base for revenue recognition and forecasting, which can in turn impact business decisions. We propose a framework and preliminary tool that strives to support knowledge workers in adopting a reviewing strategy that creates a more efficient and optimal business pipeline. We believe that this framework may provide a foundation to bridge the gap between knowledge acquisition and decision-making and encourage researchers to diversify their design and evaluation methods.

**Keywords:** contract inspection, human-AI collaboration, decision-making, uncertainty, interactive interfaces

## 1. Introduction

Contracts are central to the business operations of a large number of organizations as they are often their main source of revenue. They not only represent a record of commitment for both parties but also serve as a valuable knowledge base. Knowledge workers typically inspect contracts to collect information for tasks such as revenue recognition and forecasting, which directly impact decision-making and business operations. In cases where there is a large number of contracts, this process can be time-consuming and tedious.

With rapid advancements in machine learning and AI techniques, there has been a growing interest in human-AI collaboration in the HCI community, spanning applications such as end-user auditing [1], healthcare [2] or education [3] with the goal of optimizing performance by leveraging the efficiency of the AI and the domain expertise of the human. Across a number of applications, human-AI collaborative

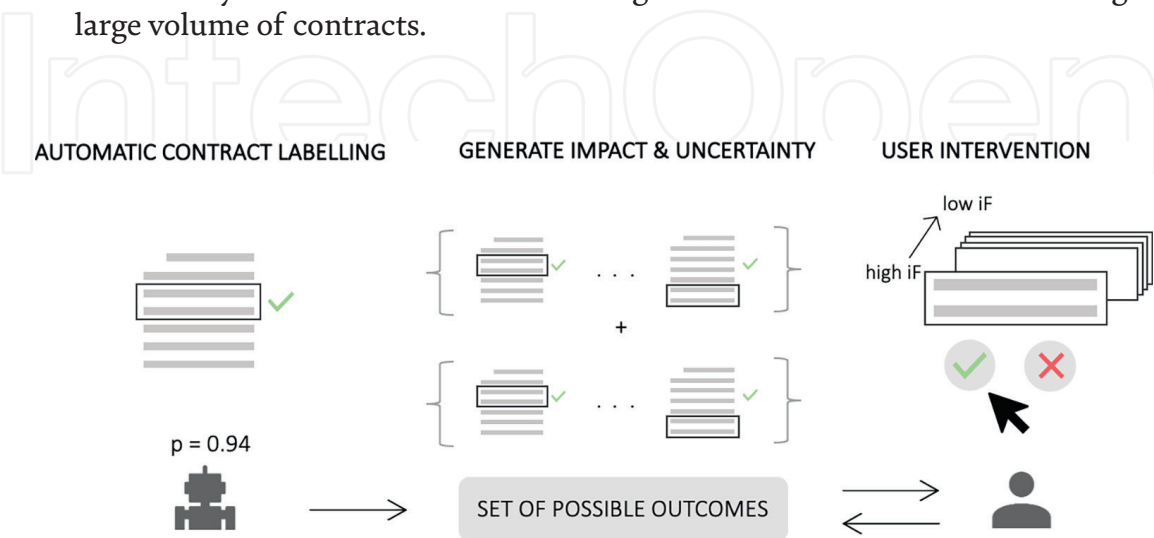
approaches have been used to refine imperfect systems. Prior work has demonstrated their ability to enhance speed, accuracy, and decision-making across various domains. Human-AI collaboration is often enabled through a mixed-initiative interface, whose approach points to a number of design principles to foster more intuitive and adaptive interactions between humans and intelligent systems, a central theme being uncertainty comprehension and the evolution of human goals through action [4].

While prior work has proposed tools for automated contract inspection [5–9], there is a lack of user-centered approaches for contract inspection tasks. Moreover, prior work only focuses on document comprehension and does not account for organizational pipelines and subsequent business operations. Our research focuses on understanding the workflow of knowledge workers and how AI can enhance the contract reviewing process and subsequent business operations. How does an AI tool fit into knowledge workers’ current workflow? How can AI optimize business operations that stem from contracts? What characteristics of a mixed-initiative interface would facilitate human-AI collaboration? Our research addresses a gap in user-centered research for AI-assisted contract inspection. We consider the perspective of knowledge workers and inquire about ways to design effective human-AI contract reviewing tools.

We conducted qualitative interviews with knowledge workers from a large organization and inquired about their contract reviewing strategies, current tools and perception of a potential AI assistant. We found that knowledge workers take a risk-based approach when reviewing contracts, where they prioritize contracts with higher monetary value to minimize errors in revenue recognition and forecasting. Knowledge workers also highlighted the usefulness of AI when flagging certain clauses despite its occasional errors. Informed by these findings, we propose a framework for incorporating business operations into human-AI contract reviewing. Informed by insights from our interviews and established visual analytics principles, we developed a prototype of a mixed-initiative contract reviewing interface that supports a Termination for Convenience (TFC) case study.

Our work makes the following contributions:

- We identify the risk factors that knowledge workers consider when reviewing a large volume of contracts.



**Figure 1.**

*Our proposed framework: (1) clauses that affect revenue recognition are identified (2) use AI uncertainty to generate and communicate forecasts (3) user chooses to review the contracts that have the biggest impact on forecasting accuracy.*

- We propose a framework for adopting human-AI collaboration for contract inspection and forecasting.
- We propose design recommendations and present a preliminary interface supporting Termination for Convenience clauses.
- We motivate the need to conduct further investigations into the integration of business operations in human-AI contract reviewing (**Figure 1**).

## **2. Related work**

Our work is situated within HCI research that strives to understand user needs to inform the design of AI-powered tools. In this section, we review and discuss two areas relevant to our work: the application of human-AI collaboration for contract inspection designing mixed-initiative interfaces.

### **2.1 Human-AI collaboration for contract inspection**

Although AI has shown tremendous potential in the past couple years, it has faced challenges for successful performance in complex real-world environments. Human-AI teaming frameworks have been used across several applications such as healthcare [10, 11], auditing [1], or education [12], with the goal of optimizing performance by leveraging the efficiency of AI tools and the domain expertise of professionals [13]. Several studies have demonstrated that human-AI collaboration can improve speed, accuracy [14] and decision-making [15]. For example, Ashktorab et al. found that AI-assistance in data labeling tasks, in which a human annotator makes decisions for which labels to apply to data, sped up the data labeling process and also increased the accuracy of data labelers [14].

Others works have investigated human's perception and adoption of AI in different settings [16–18]. For example, Weisz et al. has shown that when using AI for code translation, communication of uncertainty and alternate outcomes helped developers better understand the model and its output, as well as detect errors in translation [17]. There has been an extensive body of work on user's perception of AI fairness and trust in AI. Ashktorab et al. has conducted a study to examine machine learning practitioners' perspectives of individual and group fairness of AI models [19]. Bach et al. has conducted a survey of user trust in AI-enabled system [20].

Several AI-assisted tools for document processing have been developed [5–9, 21–25]. For example, Collins et al. [21] proposed an agent-based mixed-initiative decision support system for automated contracting. LegalVis [9] is a recent visual analytics system focusing on the exploration and inference of legal documents that cite or could potentially cite binding precedents. Rind et al. [5] proposed a system called ContractVis Highlighter that automatically highlights keywords in legal text in an online shopping setting. However, there lacks research on understanding the needs of knowledge workers when it comes to contract inspection, as well as their needs for an AI-assisted tool.

### **2.2 Designing mixed-initiative interfaces**

Designing interfaces and tools that support human-AI collaboration is not trivial. Literature across disciplines has demonstrated that both the nature and

format of the information communicated can impact how users perceive and interact with various systems. In the area of explainable AI, several studies have shown that different types AI explanations may naturally show distinctive impact on human decision makers [8]. Researchers in the field of visual analytics have developed various tools that strive to facilitate interactions between human and machines for a number of tasks [26–34]. Monadjemi et al. found that when interacting with a visualization in a guided data discovery task, participants tended to ignore recommendations despite their relevance to the task. The authors argued that the presentation style of the recommendation might have caused this effect and highlight the importance of investigating different ways of presenting suggestions [35]. Whitworth et al. argues that making suggestions obtrusive can cause users to ignore or disable them, a known example being prior assistance agent Clippy built by Microsoft [36]. While investigations into the design of human-AI collaborative tools span various applications, studies in the area of contract inspection have been limited.

### **3. Formative interviews: the challenges of reviewing contracts for large enterprises**

In order to gain more understanding into the contract reviewing process and its underlying mechanisms, we interviewed six professionals from our organization who were either responsible for reviewing contracts, or managed teams of knowledge workers. Among the six experts, two worked in Revenue Assurance, two in Forecasting & Planning and two in Procurement. The interviews, which were conducted via video conferencing, lasted about 30 minutes each. First, we asked participants about their role and daily tasks regarding contract inspection. Then we inquired about the volume of contracts they review. Next, we inquired about their reviewing strategy, and factors they consider when reporting revenue recognition forecasts. Finally, we asked participants for their thoughts on using an AI assistant to assist them in reviewing contracts. Each interview was attended by two of the authors, who took turns questioning participants. The two authors separately annotated and extracted common responses and themes that emerged from the interviews, then collaborated on consolidating these themes. For the sake of confidentiality, some details involving monetary values and stakeholder names have been anonymized.

#### **3.1 Contracts as a knowledge base for business operations**

Several operations in the business pipeline depend on contracts, which contain valuable information about the business. Knowledge workers review contracts to optimize the best deal for the company, or to determine if it is worth reviewing. S5 states that *“We’ll analyze the contract, usage and whether the contracts are optimally designed. And then, when come a renewal time, we’ll build a strategy of what should be our approach. So in a lot of cases, we have gone and changed the structure of the agreement to suit our needs.”* Moreover, a significant portion of the revenue of several enterprises stem from negotiated contracts. Participants mentioned that contracts serve as a reference document to ensure that revenue is being reported correctly and to forecast future revenue. S3 reports that *“We have a [amount] revenue target I believe, and we’re responsible for the reporting of net revenue against that.”*



### 3.2 A risk-based approach for a large number of contracts

This organization has between 12,000 and 15,000 negotiated contracts per year and knowledge workers have to prioritize which contracts to review. While their approach is not systematic, there is a general consensus on the nature and magnitude of risk factors to consider in order to minimize error in revenue recognition. S3 mentioned that on average, his team reviews 20 contracts per week and given the large number of contracts, they attempt to review 60–70% of their contract value in a given quarter. Knowledge workers tend to prioritize high-value contracts which lead to a more accurate revenue forecast. S1 states *“We get a report from [external software] and anything that’s over total contract value of [amount] will show up in this report and be flagged to the analysts associated with the territory so they can go review that contract.”* S6 admits that discarding all contracts below a certain value could pose some risk. They mention that *“We don’t review contracts under [dollar value] unless the [external stakeholder] reaches out and involves this in the process. So for anything that we’re not getting involved in, we’re not looking at those. So we could be missing something potentially.”*

The presence of certain non-standard clauses can impact how much revenue is generated from the agreement, for example exclusivity clauses, auto-renew clauses, future purchases clauses and termination for convenience clauses. S3 states *“One example [of clauses that affect revenue recognition] is the commitments to future purchases. If a customer is buying, say 1000 units of [product] and then in that same contract we give them a right to buy [product] at a discount in the future, then that would be relevant to my team, since it could cause a revenue adjustment.”* Therefore, to improve the accuracy of revenue recognition, contracts containing these clauses need to be given more attention.

### 3.3 Perception of AI: some error is better than flying blind

#### 3.3.1 A preliminary tool for termination for convenience clauses

Several steps and entities are involved in the contract reviewing process. The experts work with an external team to assist them in their reviewing process. A few of the knowledge workers we interviewed were also introduced to preliminary AI tool that flags Termination for Convenience (TFC) clauses, which allows both parties to terminate the contract at any point in time without any just cause, putting potential revenue from the contract at risk. The tool flagged contracts where a potential TFC clause was detected, but does not highlight the location of the clause in the contract, the uncertainty or explanation behind the detection.

Several experts mentioned that their approach is to review all the suggestions from the tool. S6 mentioned that *“At this point, the idea is to review all of them, because it’s still pretty new.”* This highlights the lack on trust in the tool, in some cases due to its recency. Despite their over-vigilance in their usage of the tool, most participants agreed that it is useful to help them narrow their focus toward a subset of contracts. S5 stated *“I mean [the tool] might not pull information from one or two contracts. But beyond that, I mean in any case, I’m flying blind unless I go and read all those 2000 contracts, right.”* We also asked participants about how they would handle potential errors from an AI suggestion. Generally, participants highlighted that low precision is worse than low recall. S3 mentioned that *“if [external contractor] tells us something has termination for convenience and it’s not there and we review it, in*

*most cases, my team will get it right. On the other hand, if they don't tell us and there is termination for convenience in the contract, we probably won't see it. And then we're more likely to have an error."*

### *3.3.2 An ideal AI assistant for contract inspection*

We asked participants to describe an ideal tool that would assist them in reviewing contracts. Several participants mentioned that a useful tool would have the ability to compare a set of similar contracts. S5 mentioned that *"Imagine we are looking at a whole category of about 5-6 suppliers of similar type, right. And we want to look at a holistic view if you want to do a comparison of the similar clauses."* Another feature that was highlighted by several participants is the ability to flag contracts to inform the user on where to allocate their attention. S4 mentioned that *"say we have a health dashboard, this gives us those red, amber and, you know, green flags. For each of these, we wanna know what if this happened?"* S5 mentioned the importance of highlighting certain key terms in high stake contracts *"So and all contracts where I do not have termination for convenience or all contracts where we have auto-renewal provision, all contracts where we do not have caps on increases for future purchase or future renewals. So that then I know where to focus on and we can plan it out and you know assign resources to go in and work on those. So any contract over [amount] where these clauses are not standard and we can say."*

## **4. Framework**

In this section, we propose solutions to integrate human-AI contract inspection into the business pipeline by considering factors raised by knowledge workers in the formative interviews. We present a human-AI contract inspection framework and propose design guidelines based on our findings and visual analytics principles. We design a prototype with our recommended features and present a case study for Termination for Convenience clauses (TFC).

### **4.1 Considerations for AI uncertainty**

In Section 3.2, we report that when reviewing contracts, knowledge workers prioritize high-value contracts which contain non-standard clauses. However, they disregard the potential presence of non-standard clauses in the lower value contracts, which could have consequences for revenue recognition and forecasting accuracy. On the other hand, as mentioned in Section 3.3.1, they inefficiently review all the contracts identified by the preliminary AI tool as having a potential non-standard clause. We posit that to establish an effective human-AI contract reviewing approach, an AI assistant should be transparent and communicate uncertainty. To conduct an optimal review, knowledge workers should consider AI uncertainty as well as risk factors, non-standard clauses and contract value. We propose an approach derived from principles of economic theory, which states that the best choice between uncertain courses of action is the one that yields the highest expected value. For this application, the highest expected value is one that minimizes error in revenue recognition and forecasting. Therefore, in a human-AI framework where the AI can detect non-standard clauses in contracts, knowledge workers should adopt a reviewing strategy where they first inspect contracts that have the biggest impact on

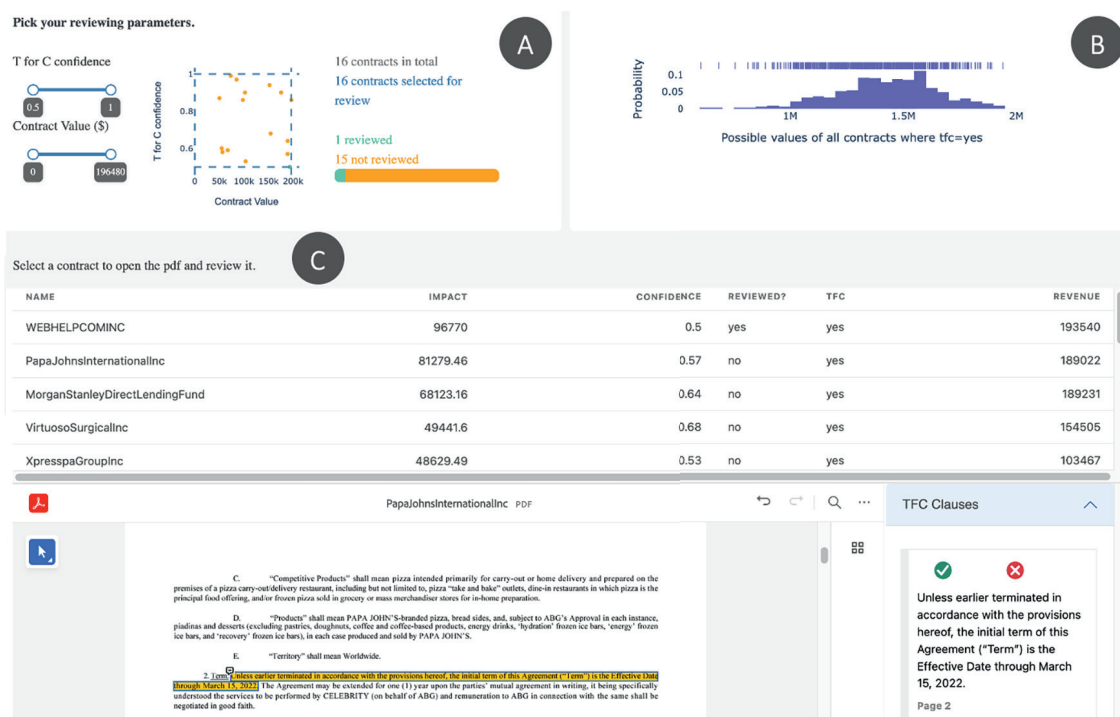
revenue recognition and forecasting accuracy, that is, contracts with low AI accuracy and high contract value.

#### 4.1.1 Prototype feature: reviewing parameters

Imagine a scenario where a knowledge worker's task is to detect the total revenue at risk due to the presence of TFC clauses. The AI is 0.95% confident that there is a TFC clause in contract A, which has a value of \$2M. On the other hand, it is 0.6% that there is a TFC clause in contract B, which has a value of \$1.5M. If the knowledge worker has to choose only one contract to review, they can weigh the AI confidence and contract value to determine the impact of each contract on revenue recognition. To effectively combine these risk factors, one approach is to compare the product of the uncertainty and the contract value such that  $iF = (1 - conf) * value$  where  $iF$  is defined as the impact (or risk) factor. Therefore, the knowledge worker should review contract B ( $iF = 600,000$ ) since the AI has a high accuracy in the presence of TFC clause in contract A ( $iF = 100,000$ ). We adopt this approach in our TFC prototype shown in **Figure 2**, where the list of contracts in (C) is sorted by  $iF$ . Since thresholds for AI uncertainty or contract value may vary, in (A) we propose that knowledge workers input these corresponding reviewing parameters and define the set of contracts that they wish to review.

#### 4.2 Communicating impact on forecasts

After reviewing a chosen set of contracts, knowledge workers then have to create forecasts and projections on data which contains uncertainty. A number of studies have



**Figure 2.** Our interface, where the user can (A) select their reviewing parameters, get an overview of the set of contracts and track their reviewing progress (B) inspect the set of possible values of contracts with TFC clauses and their corresponding probability (C) review contracts by confirming or overriding the classifier recommendation. Every time a user reviews a contract, the confidence of the label identification is set to 1 and the projections in (B) are updated.



TFC   A	TFC   B	Confidence	Value at risk (\$)
Yes	Yes	0.76	210,000
Yes	No	0.04	125,000
No	Yes	0.19	85,000
No	No	0.01	0

**Table 1.** Forecast of potential revenue at risk for contract A (\$125,000) and contract B (\$85,000).

shown that the ideal method when presented with a set of choices is to run a subset of possible analyses to investigate how decisions impact results, also called multiverse analysis [37, 38]. Kale et al. examined how researchers explore the consequences of alternative analyses when performing systematic reviews for research synthesis [39]. They found that researchers often have inconsistent rationale about their analyses and suggest shifting their attention to the impacts of decisions on the results of analysis for more optimal decision-making. In our framework, we propose to generate revenue forecasts based on the AI’s uncertainty in the presence of non-standard clauses by using the forking paths approach and showing the user the set of all possible outcomes.

#### 4.2.1 Prototype feature: visualization of the set of possible outcomes

Suppose that a classifier is 80% confident that there is a TFC clause in contract A (\$125,000) and 95% confident that there is one in contract B (\$85,000). **Table 1** shows the set of potential outcomes for the total revenue at risk across both contracts and their associated confidence. By using confidence values, we can draw a large number of outcome samples from a binomial distribution to create a probabilistic forecast. We can assume that when knowledge workers review a contract, the label’s confidence changes to 100%. More research needs to be conducted to investigate ways to capture realistic user confidence for more accurate projections. In **Figure 2B** shows the distribution of possible total revenue at risk due to the presence of TFC clause in this set of contracts. As we can see, at this point in the reviewing process, the total revenue at risk is most likely around \$1.6M. As users review more contracts, the range of possible values will get smaller and the confidence will increase. This approach allows for more detailed and transparent forecasting by communicating the range of possible values and uncertainty.

## 5. Discussion and future work

We interviewed six industry professionals to investigate their workflow, needs and challenges when reviewing contracts. Informed by our findings, we identified design considerations and developed a prototype of an interactive mixed-initiative interface that supports contract inspection and subsequent business operations. Below, we discuss some of the implications of our findings.

In our formative interviews, we found that contracts constitute an important knowledge base which is used to inform business operations. Our interviews also revealed that due the large number of contracts to review, users take a risk-based when reviewing contracts and ignore the set of lower value and lower risk contracts,

which is a black box. Finally, we identified that users would benefit from a transparent tool that allows them to filter, drill-down and review contracts in an optimal and decision-oriented manner. Based on our insights from these interviews and visual analytics principles, we designed an impact-oriented mixed-initiative interface that allows knowledge workers to have a comprehensive view of the set of contracts and insight into model outcomes and subsequent business projections—features which we posit would help them decide which contracts to review next.

While our framework and illustrated case study strives to address the lack of considerations for future business operations in contract reviewing, more research needs to be conducted to investigate how to develop a similar approach across other applications. We encourage researchers to adopt a human-centered and business-centered approach to AI research, especially when it comes to domain-specific interactions. We also encourage researchers to leverage research across other fields such as the visualization community, which has conducted extensive research on uncertainty and uncertainty visualizations. Several studies have shown that it is difficult for people to understand uncertainty information [40], despite its value in facilitating evidence-based decision-making [41, 42]. More research needs to be conducted on how to best visualize domain-specific forecasts in a human-AI framework. Moreover, future studies should be conducted to understand factors that impact trust and ways to mitigate the risk of automation bias or over-reliance.

## **6. Conclusion**

We presented our key takeaways from qualitative interviews with six knowledge workers responsible for reviewing contracts and highlight the use of contracts as a knowledge base for business operations. We propose a framework supported by a preliminary interface that strives to optimize contract reviewing strategies such that ensuing forecasting outcomes and business decisions are improved. Reflecting on our framework, we highlight the importance of human-centered and business-centered approaches in AI research. We believe that human-centered research can help develop more optimal contract inspection tools that work in collaboration with users and provide a platform to leverage the strengths of both users and AI. We encourage researchers to conduct more work to help bridge the gap between HCI and AI development.

IntechOpen

## **Author details**

Melanie Bancilhon<sup>1</sup>, Alexa Siu<sup>2</sup>, Ryan Rossi<sup>2</sup> and Nedim Lipka<sup>2\*</sup>


1 Washington University in St. Louis, USA

2 Adobe Research, San Jose, USA

\*Address all correspondence to: lipka@adobe.com

## **IntechOpen**

---

© 2024 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Lam MS, Gordon ML, Metaxa D, Hancock JT, Landay JA, Bernstein MS. End-user audits: A system empowering communities to lead large-scale investigations of harmful algorithmic behavior. *Proceedings of the ACM on Human-Computer Interaction*. 2022;**6**(CSCW2):1-34
- [2] Endsley MR, Cooke N, McNeese N, Bisantz A, Militello L, Roth E. Special issue on human-AI teaming and special issue on AI in healthcare. *Journal of Cognitive Engineering and Decision Making*. 2022;**16**(4):179-181
- [3] Williams R, Park HW, Breazeal C. A is for artificial intelligence: The impact of artificial intelligence activities on young children's perceptions of robots. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM; 2019. pp. 1-11
- [4] Horvitz E. Principles of mixed-initiative user interfaces. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1999. pp. 159-166
- [5] Rind A, Grassinger F, Kirchknopf A, Stoiber C, Özüyilmaz A. ContractVis HighLighter: The Visual Assistant for the Fine Print. In: *Proceedings of the Posters and Demos Track at the 14th International Conference on Semantic Systems, co-located with SEMANTiCS 2018*; Vienna, Austria. Aachen, Germany: RWTH Aachen; 2018
- [6] Koreeda Y, Manning CD. Contractnli: A dataset for document-level natural language inference for contracts. *arXiv preprint arXiv:2110.01799*. 2021
- [7] Hendrycks D, Burns C, Chen A, Ball S. Cuad: An expert-annotated nlp dataset for legal contract review. *arXiv preprint arXiv:2103.06268*. 2021
- [8] Wang SH, Scardigli A, Tang L, Chen W, Levkin D, Chen A, Ball S, Woodside T, Zhang O, Hendrycks D. Maud: An expert-annotated legal nlp dataset for merger agreement understanding. *arXiv preprint arXiv:2301.00876*. 2023
- [9] Resck LE, Ponciano JR, Nonato LG, Poco J. LegalVis: Exploring and inferring precedent citations in legal documents. *IEEE Transactions on Visualization and Computer Graphics*. 2022;**29**(6):3105-3120. Piscataway, NJ, USA: IEEE
- [10] Cai CJ, Winter S, Steiner D, Wilcox L, Terry M. "Hello AI": Uncovering the onboarding needs of medical practitioners for human-AI collaborative decision-making. *Proceedings of the ACM on Human-Computer Interaction*. 2019;**3**(CSCW):1-24
- [11] Fernandez R, Shah S, Rosenman ED, Kozlowski SW, Parker SH, Grand JA. Developing team cognition: A role for simulation. *Simulation in Healthcare*. 2017;**12**(2):96-103
- [12] McCall F, Hussein A, Petraki E, Elsayah S, Abbass H. Towards a systematic educational framework for human-machine teaming. In: *2021 IEEE International Conference on Engineering, Technology & Education (TALE)*. IEEE; 2021. pp. 375-382
- [13] Munyaka I, Ashktorab Z, Dugan C, Johnson J, Pan Q. Decision making strategies and team efficacy in human-AI teams. *Proceedings of the ACM on Human-Computer Interaction*. 2023;**7**(CSCW1):1-24
- [14] Ashktorab Z, Desmond M, Andres J, Muller M, Joshi NN, Brachman M, et al.



AI-assisted human labeling: Batching for efficiency without overreliance. Proceedings of the ACM on Human-Computer Interaction. 2021;5(CSCW1):1-27

[15] Bansal G, Wu T, Zhou J, Fok R, Nushi B, Kamar E, et al. Does the whole exceed its parts? The effect of AI explanations on complementary team performance. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM; 2021. pp. 1-16

[16] Houde S, Talamadupula K, et al. Investigating explainability of generative AI for code through scenario-based design. In: 27th International Conference on Intelligent User Interfaces. New York, NY, USA: ACM; 2022. pp. 212-228

[17] Weisz JD, Muller M, Houde S, Richards J, Ross SI, Martinez F, et al. Perfection not required? Human-AI partnerships in code translation. In: 26th International Conference on Intelligent User Interfaces. New York, NY, USA: ACM; 2021. pp. 402-412

[18] Weisz JD, Muller M, Ross SI, Martinez F, Houde S, Agarwal M, et al. Better together? An evaluation of AI-supported code translation. In: 27th International Conference on Intelligent User Interfaces. New York, NY, USA: ACM; 2022. pp. 369-391

[19] Ashktorab Z, Hoover B, Agarwal M, Dugan C, Geyer W, Yang HB, et al. Fairness evaluation in text classification: Machine learning practitioner perspectives of individual and group fairness. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM; 2023. pp. 1-20

[20] Bach TA, Khan A, Hallock H, Beltrão G, Sousa S. A systematic literature review of user trust in

AI-enabled systems: An HCI perspective. In: International Journal of Human-Computer Interaction. PA, USA: Taylor & Francis; 2022. pp. 1-16

[21] Collins J, Bilot C, Gini M. Mixed-initiative decision support in agent-based automated contracting. In: Proceedings of the 4th International Conference on Autonomous Agents. New York, NY, USA: ACM; 2000. pp. 247-254

[22] Mandal A, Chaki R, Saha S, Ghosh K, Pal A, Ghosh S. Measuring similarity among legal court case documents. In: Proceedings of the 10th Annual ACM India Compute Conference. New York, NY, USA: ACM; 2017. pp. 1-9

[23] Bhattacharya P, Ghosh K, Pal A, Ghosh S. Methods for computing legal document similarity: A comparative study. arXiv:2004.12307. 2020

[24] Nanda R, Siragusa G, Di Caro L, Boella G, Grossio L, Gerbaudo M, et al. Unsupervised and supervised text similarity systems for automated identification of national implementing measures of european directives. *Artificial Intelligence and Law*. 2019;27(2):199-225

[25] Tran V, Nguyen ML, Satoh K. Building legal case retrieval systems with lexical matching and summarization using a pre-trained phrase scoring model. In: Proceedings of the 17th International Conference on Artificial Intelligence and Law. New York, NY, USA: ACM; 2019. pp. 275-282

[26] Monadjemi S, Guo M, Gotz D, Garnett R, Ottley A. Human-computer collaboration for visual analytics: An agent-based framework. arXiv preprint arXiv:2304.09415. 2023

[27] Hu K, Orghian D, Hidalgo C. Dive: A mixed-initiative system supporting

- integrated data exploration workflows. In: Proceedings of the Workshop on Human-in-the-Loop Data Analytics. New York, NY, USA: ACM; 2018. pp. 1-7
- [28] Crouser RJ, Chang R. An affordance-based framework for human computation and human-computer collaboration. *IEEE Transactions on Visualization and Computer Graphics*. 2012;**18**(12):2859-2868
- [29] Kim H, Choi D, Drake B, Endert A, Park H. Topicsifter: Interactive search space reduction through targeted topic modeling. In: 2019 IEEE Conference on Visual Analytics Science and Technology (VAST). Piscataway, NJ, USA: IEEE; 2019. pp. 35-45
- [30] Endert A, Fiaux P, North C. Semantic interaction for visual text analytics. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM; 2012. pp. 473-482
- [31] Battle L, Chang R, Stonebraker M. Dynamic prefetching of data tiles for interactive visualization. In: Proceedings of the 2016 International Conference on Management of Data. New York, NY, USA: ACM; 2016. pp. 1363-1375
- [32] Brown ET, Ottley A, Zhao H, Lin Q, Souvenir R, Endert A, et al. Finding Waldo: Learning about users from their interactions. *IEEE Transactions on Visualization and Computer Graphics*. 2014;**20**(12):1663-1672
- [33] Dabek F, Caban JJ. A grammar-based approach for modeling user interactions and generating suggestions during the data exploration process. *IEEE Transactions on Visualization and Computer Graphics*. 2016;**23**(1):41-50
- [34] Ottley A, Garnett R, Wan R. Follow the clicks: Learning and anticipating mouse interactions during exploratory data analysis. *Computer Graphics Forum*. 2019;**38**(3):41-52
- [35] Monadjemi S, Ha S, Nguyen Q, Chai H, Garnett R, Ottley A. Guided data discovery in interactive visualizations via active search. In: 2022 IEEE Visualization and Visual Analytics (VIS). Vol. 2022. Piscataway, NJ, USA: IEEE; pp. 70-74
- [36] Whitworth B. Polite computing. *Behaviour & Information Technology*. 2005;**24**(5):353-363
- [37] Simonsohn U, Simmons JP, Nelson LD. Specification curve analysis. *Nature Human Behaviour*. 2020;**4**(11):1208-1214
- [38] Steegen S, Tuerlinckx F, Gelman A, Vanpaemel W. Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*. 2016;**11**(5):702-712
- [39] Kale A, Kay M, Hullman J. Decision-making under uncertainty in research synthesis: Designing for the garden of forking paths. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM; 2019. pp. 1-14
- [40] Hullman J. Why authors don't visualize uncertainty. *IEEE Transactions on Visualization and Computer Graphics*. 2019;**26**(1):130-139
- [41] Kim Y-S, Walls LA, Krafft P, Hullman J. A bayesian cognition approach to improve data visualization. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM; 2019. pp. 1-14
- [42] Karduni A, Markant D, Wesslen R, Dou W. A bayesian cognition approach for belief updating of correlation judgement through uncertainty visualizations. 2020