

AI4RFQ: Exploiting User Annotations Towards Automating the Extraction of Information and Requirements from Specification Documents

RENE KAISER, Know-Center – Research Center for Data-Driven Business & Big Data Analytics, Austria

HERMANN STERN, Know-Center – Research Center for Data-Driven Business & Big Data Analytics, Austria

A common task in business-to-business relationships is the analysis of specification documents within a request for quotation (RFQ) process. For non-standard products or services, one company with a need provides detailed specifications to potential suppliers and invites them to make an offer. In order to create an offer that fits with the specification, a supplier company needs to carefully inspect the documents to extract any requirements that are relevant for designing the contents of this offer and calculating its price. In a research effort focusing on HCI and natural language processing aspects, we seek methods and technology to support such RFQ processes. We have created a software tool that supports the reading and requirements identification process with a set of intelligent features that are expected to speed up the process significantly. Whenever users are extracting information, they are asked to annotate all text passages that contain relevant information. This annotation data is used to train machine learning models that provide suggestions in the future. The degree of automation increases continuously as more example data is collected. We strive to participate in the 'Automation Experience at the Workplace' workshop to discuss this change in work practice for our target users in terms of a technical as well as a user experience dimension. We would like to especially reflect on users' motivation to cooperate with the system and provide useful annotations (sharing their precious work expertise), the question of responsibility when dealing with automatic suggestions, as well as the fear of being replaced by 'an Artificial Intelligence' in the near future.

CCS Concepts: • **Human-centered computing** → **HCI design and evaluation methods**; • **Information systems** → **Search interfaces**; • **Applied computing** → **Document searching**.

Additional Key Words and Phrases: RFQ, NLP, automation, decision support, user experience, document-centric work, requirement extraction, information extraction, trust, responsibility, concentration

1 INTRODUCTION

Request for quotation (RFQ) processes are an essential part of many business-to-business (B2B) relationships. One company is in need of a custom good, service or product, and another company may be able to supply it. The latter obtains specification documents from the former and needs to check them carefully in order to decide if they can meet the demands and thus make an offer, work out the offer details, and define a (competitive) price. To be able to do that, requirements, norms and other essential information need to be extracted from the specification documents. Because those specification documents may consist of hundreds of pages and reference dozens of attached files, this is a labour-intensive process.

In our research understanding, the analysis of specification documents is a very particular case of knowledge work. The task is complex and requires more than the recognition of typical '*shall be / shall have*' requirement phrases. The ability to extract all relevant information including the surrounding context within the documents requires in-depth

Workshop proceedings *Automation Experience at the Workplace*

In conjunction with CHI'21, May 7th, 2021, Yokohama, Japan

Copyright © 2021 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

Website: <http://everyday-automation.tech-experience.at>

knowledge and experience concerning the domain, the products, the offer and its (e.g., technical) details and peculiarities. Requirements may be contained in natural language paragraphs, bullet point lists, tables, and even diagrams. There are interdependencies between requirements which need to be taken into account. Contradicting information may appear which needs to be noticed and clarified. In certain cases, the information sought is not directly contained in the specifications at all, but needs to be inferred based on background knowledge and even the gut feeling of experts mastering these tasks for several years.

For offer engineers whose primary task is to work on RFQ cases on a daily basis, this task of skimming, reading and processing information can be described as cognitively challenging due to its monotonous nature. We understand this especially holds for the reading and skimming parts which may last several hours in one go. Users need to maintain concentration in order to spot relevant information, not miss any detail, and avoid mistakes processing this information. They need to carefully decide, with the help of experience, which parts of the specifications to read carefully and what to skim only, in order to save time for the most relevant parts. There may be significant time pressure due to *i)* the cost of the offer process itself, getting more expensive the longer it takes, and *ii)* the deadline to place an offer which is very short-term in many cases.

From a company's perspective, several risks are involved in the RFQ process, in particular to miss individual requirements, contradictory information or showstoppers which prevent the company from placing a competitive offer. In many instances, the price of the offer is binding. Placing an offer at too cheap a price may hence result in a substantial financial loss. Eventually not winning an offer at all typically means that all the work put into the offer creation is not directly remunerated – offers usually can't be reused. A balance needs to be found for how much time is invested in the offer process, weighing these risks strategically.

Even though the basics of RFQ processes and the challenges involved are well-known, on a detailed level, the process can obviously differ significantly from company to company and from offer engineer to offer engineer. In some cases, the offer engineer is asked to scan the specifications for any requirements that may appear, for example in software companies that are happy to develop any individual software solution with any feature the customer requests. In other cases, a company may offer physical devices that can be adapted to customer needs to a certain extent. In these cases, the offer engineer may use a predefined form with a list of parameters and is primarily looking for any choices that express how the product shall be customised. Besides this example, there are numerous other distinctions in RFQ processes.

While research has looked at document-centric knowledge work in general, and published findings we can build on (e.g., [3, 5, 7–9, 13]), HCI research on technology supporting the specific task of RFQ processing is scarce. This is surprising given the fact that this is a very relevant process and cost factor for many companies. In terms of understanding, assessing and designing for user experience, naturally there are overlaps and takeaways from other HCI research efforts that can be considered and adapted to this particular situation. While our research can even be informed by methodological aspects and learnings from research on everyday automation experiences [2], there are profound differences to our context as we are concerned with a professional work task with very particular characteristics.

We have looked at supporting concentrated work over periods of multiple hours from an HCI perspective before, but in a surveillance context where the user's predominant tasks are passive monitoring [6]. We understand that apart from basic overlaps, supporting concentration in active reading and information extraction tasks requires very different software capabilities. In our current work, we continue a long-lasting line of research on work-integrated support, automating tasks by applying advances in information technology (see e.g. [11]).

In the realm of a socio-technical intervention (cf. [1]) involving company partners in multiple research projects, we are developing a prototypical RFQ system which includes several advanced capabilities. Two main groups of features are combined:

- *Supporting the reading process and highlighting relevant parts of the document:* based on our insights, users are not reading every part of a specification with the same level of diligence. Based on the ability to understand the structure of a document and what certain parts in the section hierarchy are supposed to contain, some parts are read carefully while other parts are only skimmed. We have been told by professional offer engineers that the ability to spot relevant information while only skimming can be learned with experience. The intelligent support our tool offers in this regard is that the users' attention is steered towards what the system considers relevant – matches to predefined keywords, phrases, combinations of numbers and units, etc. This is currently achieved by a colour highlighting overlay within an integrated document reader. In parallel we investigate alternative methods for users with visual impairments such as colour blindness.
- *Providing suggestions and hints:* the second group of intelligent support features provides suggestions for handling requirements or relevant parts of the documents which the system automatically detects. To be able to do that, data from previous cases is exploited, as a natural language processing component trains detectors and classifiers in the form of machine learning models. The system shall automatically detect requirements or text snippets which contain information that are relevant for a certain requirement. It shall spot cases that warrant clarification with the customer, for example contradicting information, or deviations, i.e., requirements that cannot be (entirely) fulfilled or at least need to be clarified with colleagues or the customer. In the RFQ process, the offer engineer may even decide to completely refrain from reading the specification documents start to end, and instead only iterate through the list of automatically generated suggestions, especially in cases with very tight deadlines. This workflow obviously requires a level of trust towards the tool.

The key idea behind the 'intelligence' of our system is that offer engineers use it as a tool to create offers in a work-integrated way. Most of the document-centric tasks (reading, identification of requirements, linking references, creating deviation lists, exports, etc.) are offered within the system. The new knowledge created by finding, rating, commenting and linking requirements etc. is fed back into the system and helps it to adapt to new requirements, contexts, relevant terms and so on. Machine learning models are iteratively re-trained based on this historic data, learning from a set of examples that continuously covers more and more cases.

From the point of view of designated users, our system includes 'AI' capabilities – *artificial* or *assistive* intelligence, depending on the definition of the term. Users were involved in our design process via interviews and observations following the *contextual inquiry* method (cf. [4]). We are aware that the stakeholder users involved in our development understand that a new tool with artificial intelligence features may change their tasks in the RFQ process significantly. As the process becomes more efficient, their job descriptions will change. We argue that partly automating standard tasks gets them more time to focus on the harder tasks requiring human expertise and knowledge. However, it is also clear that some of those jobs will disappear. The stakeholder users involved in our development process can be expected to be aware of that. As part of our research, we focus on a human-centred perspective not only with respect to how to support such document-centric concentrated knowledge work, but also the impact of ever-increasing automation on the users personally.

The subsequent Section 2 describes the vision of future technology support for RFQ processes in the form of a *design fiction*. It describes how a fictitious employee experiences working with an AI system. Section 3 lists a set of aspects and questions concerning the impact of increasing automation in RFQ processes from a user experience perspective. We would like to contribute and discuss these questions at the Automation Experience at the Workplace workshop and follow-up activities.

2 DESIGN FICTION: EXPERIENCING COLLABORATION WITH AN EVER-IMPROVING AI SYSTEM

In order to describe the vision of how a future version of our tool shall support the user in RFQ processes, and to explore which impacts an increase in automation capabilities may have, we utilize the *design fiction* [10, 12] methodology. The following describes a fictitious usage scenario in a distant future 8 years from today.

Paula is an employee at a large company that produces solar panels. In her role as an offer engineer, she is handling one RFQ case after another for the majority of her working time. In principle, Paula is happy with the infrastructure at her workplace, and she especially appreciates the two large adjacent screens on her desk which help to reduce eyestrain as she keeps reading and processing specifications for hours. A major downside of her spot within an open-plan office is that she feels frequently distracted by acoustic cues caused by colleagues moving or talking, as well as certain electronic devices. She wouldn't prefer to work in an isolated office by herself either as that would limit social interactions and prevent informal knowledge exchange. Hence, Paula accepts that at certain times of the day maintaining concentration is challenging and tries to mitigate this problem proactively by scheduling certain kinds of tasks for these phases. To interact with a computer, in addition to a traditional keyboard, Paula extensively uses gesture interaction. For the most part she refrains from using voice control features in order not to distract colleagues, even though she prefers this interaction method in her private life.

Three years ago, a software system was introduced at her company that changed the RFQ process in fundamental terms. As a veteran in the RFQ department, Paula was deemed a stakeholder user and was involved in the development and configuration of the system. She enjoyed this process as the research partner she collaborated with valued her expertise and opinion, and also it was a nice break from the usual work routine.

The new RFQ tool attempts to automate many aspects of Paula's work. The cooperation between the user and the system for the most part follows a semi-automatic paradigm, which translates to a workflow where the system provides suggestions and the user is reviewing them in order to take the final decision. The system's suggestions are useful with great accuracy. The longer it is used, the better it can recognise rare exceptional cases. Most RFQ cases can now be completed in a single working day compared to 6–10 days without the tool. Two thirds of the colleagues in her department have moved to other roles within the company since the tool was introduced. Initially, among the designated users there were substantial concerns and fears towards the new system which at that time was marketed to include 'AI' capabilities, a hype buzzword that is no longer *en vogue* by the year 2029. Today, most former members of the RFQ department are content with how their role and work routine has changed. They are engaged with fewer monotonous routine tasks and instead focus on more cognitively challenging knowledge work.

When the new RFQ tool was eventually rolled out, Paula took training courses that not only explained in a very transparent manner what the system does, can and cannot do, and what data is recorded, but also basics on artificial intelligence technology and its social and ethical impacts. Paula is genuinely proud of her 'AI literacy' certificate. She framed it and displays it on the wall next to her work desk.

The RFQ tool applies machine learning techniques that require data to train on. Any information the users enter is stored and contributes to the continuous improvement of suggestion and automation models. It could even be reconstructed for how long each page of the specifications has been looked at. However, Paula's company has always denied doing or even planning this. Initially, Paula felt uneasy about this, felt watched, felt pressure. After some time she got used to it. Paula rarely thinks about the data that is captured by her inputs and what could be extracted if analysed. This holds as well for the device interpreting her gesture interactions. She trusts that the data is only used for algorithmic purposes and not for example for work performance comparisons among colleagues. None of her colleagues ever reported any issue that occurred within the company in that regard.

The tool guides Paula through the process and provides suggestions based on what it detected automatically. This includes any requirements that are found, an assessment on the company's ability to fulfil them, estimations of associated work effort and cost, etc. The tool further spots any potential issues that warrant Paula's attention, such as contradicting requirement information or requirement details that have never been encountered in historic cases.

The system even predicts the chances of winning the offer based on historic data. When this value is low, the managers of Paula's department instruct her to invest less time. In that case she will not review each of the system's suggestions and just accept a certain share blindly, and take the risk of making assumptions rather than clarifying each open question with the customer in detail. While Paula's intrinsic motivation is to process every RFQ case in full detail and best possible quality, she understands the company's strategy to invest less time in some cases and to accept the risk of an offer that is not competitive or simply too cheap. In these cases, she feels less responsible for the offer she creates.

Whenever Paula enters information, she annotates and links all the text passages within the specifications that are relevant. In turn, Paula can look at all text passages that led the tool to suggest something automatically. It is important that automatic suggestions are explainable. Due to this, Paula feels as if she were collaborating with a virtual agent whom she supervises. Paula is especially impressed when the system spots any issues stemming from logical interdependencies of multiple requirements. The system may detect such cases particularly when it concerns the basic parameters to customise solar panels which must be provided for each and every offer. Regarding these predefined parameter requirements, the system has a lot of historic example data at its disposal.

Paula never feared that the tool will replace her completely. Despite all the intelligent features, there are still cases where she needs to take decisions based on her comprehensive background knowledge and apply multiple layers of logical inferences in order to come to a conclusion. Such cases will never be automated, she assumes.

3 AUTOMATION EXPERIENCE AND SUCCESS FACTORS

The successful introduction of AI-based workplace systems raises several questions in multiple dimensions. This not only contains typical change management issues because a new system also changes work practices. The 'AI' buzzword also brings many unanswered questions since some claim that AI can "do everything" and will be able to replace employees. So why would users be willing to speed up that process by "feeding" AI systems and contribute to getting themselves replaced in the future? Many employees still do not really know what AI is and what it can – and cannot – do.

In the realm of a research project, we investigate success factors for creating and introducing human-centred AI-based systems. Our focus is on a human-centred approach to AI applications and in particular on quality control applications. The RFQ process discussed in this position paper is one of the use cases being worked on in this project. In interviews with (research) companies, end users and work councils we compiled a list of success factors, raising questions that we seek to exchange on, in the realm of the Automation Experience at the Workplace workshop at CHI 2021. The following

discusses these aspects in the form of questions which we would like to contribute to and discuss at the workshop. Questions are clustered along 5 aspects.

- *HCI and user-centred design – how to best support the user and design collaboration with the system?*
 - How can we best support RFQ processes as a socio-technical intervention, concerning both work processes and technology?
 - How can we evaluate the user experience of users and understand if subgroups exist that have different needs, compared to others?
 - User-generated annotations are key to our approach: we need the user to collaborate with the system and tell it what exact snippets of text in the specifications contain relevant information. A risk in that regard is that there is some degree of freedom for how these annotations are created. The resulting data may either be useful, useless or even detrimental to the automatic suggestion models. How can we ensure the collaboration of users with an intelligent system becomes a fruitful partnership?
 - Currently, extracting information from specifications is a monotonous task. Can the prospect of a more interesting job routine help to win the support of employees despite the obviously negative consequence of jobs being lost?
 - How can we best integrate support and automatic suggestions such that the user is not distracted in lengthy phases of concentrated reading?
 - What degrees of freedom on how a task is worked on does a new intelligent system need to offer in order to maximise the productivity of users whose work tasks and processes are significantly changing? What shall be mandatory, what shall be optional in that regard? Are there different ways of working and thinking that need to be supported separately?

- *User experience and concerns*
 - An intelligent system logging user actions and interactions records in great detail how employees work. It makes their work measurable and comparable, in principle. What far-reaching issues are implied by this circumstance?
 - Is there any sensitive knowledge that is currently protected but would become unprotected when an AI system starts being used, e.g., if the data it captures is connected to further available data sources?
 - What particular aspects are relevant with respect to the user experience in this context and how can user experience be evaluated?
 - In what way could users be held accountable for unrealistic expectations to an intelligent system?
 - What success criteria are relevant with respect to the willingness of the users to metaphorically teach their expertise to an AI system?
 - In how far is user acceptance related to the topic of 'AI explainability'? What impact can be observed on the user experience if users are able to understand in sufficient detail what the system is doing in terms of intelligent functionality?
 - How is an organisation dealing with mistakes in the realm of an employee using and collaborating with an intelligent system?

- What skills will be lost when the new system is used, as a risk for both the employees personally as well as the company?
- *Employee rights, data protection and privacy*
 - Which laws, rules and regulations such as the General Data Protection Regulation (GDPR) in the European Union or other country-specific worker protection laws need to be respected when introducing an AI system at a workplace?
 - Which stakeholders (e.g., work councils) are (or should be) involved in all phases of the development and roll-out process and what are their rights and possibilities to influence it?
 - How does the lifecycle of any data that is collected look like, who is responsible for deleting it, and how is a 'data protection officer' (part of GDPR regulations) involved in these processes?
 - Are employees aware which data is captured and stored – especially any data related to themselves and their work procedures? What is this data used for, how is such data analysed, what knowledge will be inferred from it? Who will have access to this data? What influence can a user have on the data that is captured?
 - Who owns the data? Could the data be made available to third parties?
 - What happens when an employee is leaving the company?
 - What ethical questions do appear in this context?
- *AI literacy – what level of understanding is required?*
 - How can users best understand what artificial intelligence can and cannot do? Do users need training on AI basics, or only on the specific tool they use?
 - Concerns of dealing 'with an AI' as a somewhat mysterious entity the capabilities of which are not really understood – how to best deal with this issue from our point of view of a technology provider and research partner? How can unwarranted fears be mitigated?
 - To what extent need users with limited IT education and training understand natural language processing (NLP) features, ranging from simple regular expression search patterns to sophisticated machine learning models?
 - How can a company that starts using an intelligent system best communicate the advantages of the change to its employees?
 - Automating routine tasks also means, in a positive sense, that offer engineers can focus on other tasks that require more skill and knowledge. How can this transition be supported?
- *AI trust, ethics and responsibility – how to define the rules of collaboration?*
 - Will users trust intelligent features and in what form does this affect the output of their work?
 - How will users perceive the increasing level of automation as the system keeps learning?
 - Will users feel responsible for results that are based on automatic suggestions?
 - In what way does a company which starts using AI systems need to change as an organisation?

ACKNOWLEDGMENTS

The Know-Center is funded within the Austrian COMET Program – Competence Centers for Excellent Technologies – under the auspices of the Austrian Federal Ministry of Transport, Innovation and Technology, the Austrian Federal Ministry of Economy, Family and Youth and by the State of Styria. COMET is managed by the Austrian Research Promotion Agency FFG. The present work has also received funding through the AK Wien Digifonds-funded project "AI for Good".

REFERENCES

- [1] Gordon Baxter and Ian Sommerville. 2011. Socio-technical systems: From design methods to systems engineering. *Interacting with computers* 23, 1 (2011), 4–17.
- [2] Peter Fröhlich, Matthias Baldauf, Thomas Meneweger, Manfred Tscheligi, Boris de Ruyter, and Fabio Paternó. 2020. Everyday automation experience: a research agenda. *Personal and Ubiquitous Computing* 24, 6 (2020), 725–734. <https://doi.org/10.1007/s00779-020-01450-y>
- [3] Richard Goodwin, Rama Akkiraju, and Fred Wu. 2002. A Decision-support System for Quote Generation. In *Eighteenth National Conference on Artificial Intelligence* (Edmonton, Alberta, Canada). American Association for Artificial Intelligence, Menlo Park, CA, USA, 830–837. <http://dl.acm.org/citation.cfm?id=777092.777219>
- [4] Karen Holtzblatt and Hugh Beyer. 2016. *Contextual design: Design for life*. Morgan Kaufmann.
- [5] Paul Jones, Shivani Sharma, Changsung Moon, and Nagiza F. Samatova. 2017. A Network-Fusion Guided Dashboard Interface for Task-Centric Document Curation. In *Proceedings of the 22nd Int. Conf. on Intelligent User Interfaces* (Limassol, Cyprus) (*IUI'17*). ACM, New York, NY, USA, 481–491.
- [6] Rene Kaiser and Ferdinand Fuhrmann. 2014. Multimodal Interaction for Future Control Centers: Interaction Concept and Implementation. In *Proceedings of the 2014 Workshop on Roadmapping the Future of Multimodal Interaction Research Including Business Opportunities and Challenges* (Istanbul, Turkey) (*RFMIR'14 at ACM ICMI*). ACM, 47–51.
- [7] Ka Ho Leung, Ching Chug Luk, KL Choy, Hoi Yan Lam, and Carman KM Lee. 2019. A B2B flexible pricing decision support system for managing the request for quotation process under e-commerce business environment. *International Journal of Production Research* 57, 20 (2019), 6528–6551.
- [8] Nishtha Madaan, Hima Karanam, Ankush Gupta, Nitisha Jain, Arun Kumar, and Srikanth Tamilselvam. 2017. Visual Exploration of Unstructured Regulatory Documents. In *Proceedings of the 22nd Int. Conf. on Intelligent User Interfaces Companion* (Limassol, Cyprus) (*IUI'17 Companion*). ACM, New York, NY, USA, 129–132.
- [9] Hugo Romat, Nathalie Henry Riche, Ken Hinckley, Bongshin Lee, Caroline Appert, Emmanuel Pietriga, and Christopher Collins. 2019. ActiveInk: (Th)Inking with Data. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) (*CHI'19*). ACM, New York, NY, USA, Article 42, 13 pages. <https://doi.org/10.1145/3290605.3300272>
- [10] Daniel M. Russell and Svetlana Yarosh. 2018. Can We Look to Science Fiction for Innovation in HCI? *Interactions* 25, 2 (2018), 36–40. <https://doi.org/10.1145/3178552>
- [11] Hermann Stern, Rene Kaiser, Philip Hofmair, Peter Kraker, Stefanie N. Lindstaedt, and Peter Scheir. 2010. Content Recommendation in APOSDLE using the Associative Network. *Journal of Universal Computer Science (JUCS)* 16, 16 (2010), 2214–2231.
- [12] Theresa Jean Tanenbaum. 2014. Design Fictional Interactions: Why HCI Should Care about Stories. *Interactions* 21, 5 (Sept. 2014), 22–23. <https://doi.org/10.1145/2648414>
- [13] Maartje ter Hoeve, Robert Sim, Elnaz Nouri, Adam Fourney, Maarten de Rijke, and Ryen W White. 2020. Conversations with Documents: An Exploration of Document-Centered Assistance. In *Proceedings of the 2020 Conference on Human Information Interaction and Retrieval*. 43–52.