



Kunnskap for en bedre verden

DEPARTMENT OF INDUSTRIAL ECONOMICS AND
TECHNOLOGY MANAGEMENT

TIØ4567 - STRATEGY, INNOVATION AND INTERNATIONAL
BUSINESS DEVELOPMENT, SPECIALIZATION COURSE

Mediating the issue of "publish or perish" by requiring clear research goals

Author:

Andreas Bjelland Berg

The problem of "publish or perish" is increasingly impacting academic trustworthiness in the eyes of the public, as well as academia itself, through decreasing reproducibility of research and increasing volumes of papers being produced. This paper looks at three diverse papers within the field of responsible artificial intelligence, and evaluate them using a set of criteria related to research questions. I find that the use of research questions is limited, but see clear connections between the papers and their overarching research goals, suggesting that research questions are possibly used without being mentioned in the papers. I conclude by drawing connections between the use of research questions and paper quality, while considering alternative methods for increasing quality, such as including a structured summary of contributions in the introduction of the paper.

This paper contributes to the field of academic writing by examining the connection between research questions and paper quality, as well as suggesting paths for future work on research questions.

November, 2022

Table of Contents

1	Introduction	1
2	Methodology	2
2.1	Method for selection of papers	2
2.2	Method for selection of assessment criteria	3
2.3	Limitations	3
3	Assessment criteria	4
4	Evaluation of selected papers	4
4.1	Presentation of Barredo Arrieta et al. (2020)	4
4.2	Evaluation of Barredo Arrieta et al. (2020)	5
4.3	Presentation of Vakkuri et al. (2020)	6
4.4	Evaluation of Vakkuri et al. (2020)	7
4.5	Presentation of Ntoutsis et al. (2020)	8
4.6	Evaluation of Ntoutsis et al. (2020)	9
5	Discussion	10
5.1	Reasons for failure to include research questions	11
5.2	Research questions and quality	11
5.3	Methods to ensure maximum quality of papers	12
6	Conclusion	12
	Bibliography	13
A	Search history	15
A.1	Relevant papers	15
A.2	Assessment criteria and paper quality	15
A.3	Secondary articles	16

1 Introduction

With the treat of fake news (Allcott and Gentzkow 2017), anti-vaccine attitudes (Lindholt et al. 2021) and climate change denial (Anderegg et al. 2010) at an all-time high, researchers must do what they can to ensure maximum trustworthiness in the eyes of the public. At the same time, the world of academia struggles with the issue of "publish or perish", a requirement that researchers must continuously produce articles, books and other output in order to be able to advance their careers (Grimes et al. 2018). This has led to an explosion of published papers, in turn making both researchers and the informed public have to wade through mountains of possibly relevant papers in order to get new insights (Bornmann and Mutz 2015). Another result of this overwhelming amount of research is a strict prioritization from publishers, with high ranking journals opting towards publishing novel results rather than replicating studies (Grimes et al. 2018).

Two issues arise from this paper explosion. First, the focus on novel results, rather than replicating studies, have led to a "reproducibility crisis" (Begley and Ioannidis 2015), where published studies are hard to replicate. An example of this can be seen in preclinical research – a field where one would assume accurate, reproducible results would be of outmost importance – where Begley and Ioannidis (2015, p.1) report that an estimated 75%-90% of findings presented in high ranking journals are not reproducible. Such findings are seen across several fields, including psychology, chemistry, biology, physics, earth sciences (Baker 2016), and cancer research (Nature Neuroscience 2017). Failure to replicate studies can have a significant negative effect on public trust in academia (Grimes et al. 2018).

The second issue arising from the increasingly increasing number of papers is the general speed of research. As the number of total publications has increased exponentially since the 1600s (Bornmann and Mutz 2015), researchers entering a new field are faced with a massive wall of possibly relevant research to review, before they can contribute their own. Although several guides for undertaking this exists (i.e., Nightingale 2009; Okoli 2015; Popenoe et al. 2021), they provide aid for the symptoms of the underlying stack of articles, rather than attack the actual problem at hand. Additionally, an ever-expanding volume of papers within a field makes it hard for established researchers to keep up, potentially limiting the collective knowledge they can provide (Davies and Felappi 2017).

One suggestion to combat the issue of reproducibility, mentioned by Baker (2016), is pre-registration, where hypotheses and plans for data evaluation are submitted to a third party before the actual experiments take place. To help with the issue of the overwhelming number of papers, there should be a way for researchers to easily get an overview of a given paper, in order to quickly be able to assess whether it will be relevant for their own research. These solutions can be combined in the way of research questions. When included in a clear way, research questions represent the goals of the presented research, succinctly summarizing the research being done. At the same time, research questions can be used as an easy-to-implement source for pre-registration, giving the same benefits as presented by Baker (2016). This shows how the use of research questions has potential for mitigating the problem of "publish or perish".

With this background, this paper will look at the following two research questions:

1. How are research questions currently used, and to what degree are they clearly presented, in existing papers published in the field of responsible artificial intelligence?
2. How can active use of research questions mitigate the problem of "publish or perish"?

The problem of "publish or perish" has been widely discussed, with several attempts at finding solutions (i.e., Bornmann and Mutz 2015; Davies and Felappi 2017; Grimes et al. 2018). This paper aims to contribute to mediating the problem, by providing ways to ensure future papers are as easy-to-understand and well structured as possible. To achieve this goal, three peer-reviewed and published papers are analyzed. The papers are carefully selected to be representative of the field of responsible artificial intelligence, and are evaluated against a pre-defined set of criteria, focusing on article structure and -clarity. This evaluation is then used as background for a discussion of research question use, and how good use of research questions can help mitigate the problems

arising from an ever-expanding volume of papers. As "publish or perish" is unlikely to go away in the near future, this discussion is highly relevant, and the solution presented in this paper may be used as one of many tools to mediate the problem.

The paper will start by discussing the methods used for selecting the papers and the evaluation criteria. Then, the criteria used for evaluating the papers are presented, before each paper goes through the same process – a short presentation, followed by an in-depth analysis of the paper. Any findings are then used as foundation for a discussion, where research question use is discussed in-depth, including advice for how to ensure maximum clarity in future articles.

This paper contributes to mitigating the problem of "publish or perish" by evaluating submitted papers for their readability and use of research questions. It connects the use of such questions to the notion of paper quality, and discuss possible reasons why research questions are not as widely used as academic writing guides suggest. Finally, this paper contributes by outlining possible future research within the field of academic writing, including suggestions for methods that can ensure maximum paper quality.

2 Methodology

2.1 Method for selection of papers

As the scope of this evaluation is limited to three papers, great care was taken when making the selection, to ensure as generalizable results as possible. To ensure optimal breadth in the selection, papers were chosen to have different levels of recognition, from different authors in different years, yet in the same field of science – responsible artificial intelligence (AI). Additionally, papers were chosen to ensure both conceptual, qualitative and quantitative papers were part of the evaluation.

For the first paper, the goal was to have a relatively recent paper, with a good standing in the field of responsible artificial intelligence. A search was conducted on the Scopus digital database of papers using the terms "Responsible AI" and "ethical AI". 404 papers were retrieved, and to ensure selected papers are relevant in the field, all results were sorted from most to least cited. The paper 'Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI' by Barredo Arrieta et al. (2020), published in the journal *Information Fusion* was at the top of the list with 1447 citations. Being of a conceptual nature, highly cited and recent, the article made for a good first paper.

For the second paper, the goal was to find a quantitative study done within the field of responsible artificial intelligence. Although AI is in itself a highly quantitative field, finding quantitative papers turned out to be harder than expected, and the 29 most cited papers in the above-mentioned Scopus search were conceptual. The 30th most cited, however, was of a quantitative nature. The selected paper was 'The Current State of Industrial Practice in Artificial Intelligence Ethics', written by Vakkuri et al. (2020). The paper, published in *IEEE Software*, has 16 citations on Scopus. While this is a relatively low number, it is still of some relevancy in a relatively small field as responsible artificial intelligence, and provides a good breadth when analyzed alongside Barredo Arrieta et al. (2020).

For the final paper, the goal was to find a qualitative study within the field. The third most cited paper in the Scopus search was a literature review, and thus satisfied the requirement for qualitativensness. The third paper, then, was chosen to be 'Bias in data-driven artificial intelligence systems—An introductory survey', written by Ntoutsis et al. (2020). Published in *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, the paper has 109 citations on Scopus, placing it nicely between the two first papers.

The complete search history used to find the above-mentioned articles can be seen in Section A.1. Note that the journey was far from straight-forward, as several fields of science failed to show promising results, or included too many, too old or too narrow-focused articles to have any significant relevance.

The field of responsible artificial intelligence was chosen for two reasons. First, it is a relatively popular field of science, with frequent additions to the state of research. Secondly, it is a field closely aligned with the author's upcoming Master's thesis, aligning it with both personal and academic interests. One issue with this field, however, is the relative freshness of the research field. As most research is done relatively recently, there is a lack of spread of papers over time. This is likely to not significantly impact the results of this analysis.

2.2 Method for selection of assessment criteria

A literature review was conducted in order to find a set of well-fitting assessment criteria for the topic. The search was done using the Scopus electronic database and Google Scholar, and was filtered on articles that were either open access or available through NTNU. The search terms are listed in Section A.2.

A challenge when searching for articles discussing research quality – and by extension contributing to the selected assessment criteria – was that most of the results were either aimed at users of the research (i.e., Domholdt and Malone 1985; Mayhew 1993), reviews done on other, unconnected fields of science containing few generalizable guidelines (i.e., Biddle and Asare 2011; Polanczyk et al. 2015; Wade and Hulland 2004), or research done on image quality (i.e., Mittal et al. 2012; Wang and Bovik 2002). To mitigate this, a thorough review of each paper was done, to ensure that all assessment papers contain usable criteria for research quality or guidelines for writing or reviewing academic papers.

2.3 Limitations

There are several limitations with these methods for selecting papers. First of all, the small scope of the course and this paper provides a limit for how in-depth the literature searches can go. As an exhaustive literature review could take weeks, if not months, the depth of this review ended up being much shallower, with only a selected number of papers being reviewed. While these were selected as the most promising articles based on the available metadata, there is a significant risk that useful or interesting articles may have been skipped due to this.

The same limitation also applies to the search for criteria-relevant papers. While several guides for academic writing were found, only one evaluation guide, and no other relevant sources for criteria, was found. While it would take more time than is assigned for this course, a deeper literature search would likely provide more relevant papers, thus leading to more relevant assessment criteria. Missing these may potentially lead to some relevant assessment criteria being missed, weakening the criteria presented in Section 3.

One limitation arises from the chosen field of science. While artificial intelligence itself is a relatively established field, dating back to at least the 1950s (Moor 2006), responsible artificial intelligence is relatively new. This leads to both a narrow spread of publication dates, as all selected papers are from 2020, as well as fewer non-conceptual papers. As most articles within the field are published relatively recently, the field has barely begun to move past the conceptualizing stage, which greatly limits the amount of qualitative and non-conceptual quantitative work that can be done.

The age of the field may also lead to less generalizable results. As a relatively new field of science, it is natural to assume that it will attract younger researchers looking to make significant contributions in the field's youth. This may lead to a lower quality of papers, and different issues, compared to more established fields of science, where researchers are more likely to have more experience with writing papers.

3 Assessment criteria

The result of the literature review conducted in Section 2.2 was a selection of seven papers. Six of these (Cuschieri et al. 2019; Davidson et al. 2012; Jha 2014; Katz 2006; Lin and Kuo 2012; Rosenfeldt et al. 2000) are guides for academic writing, providing a source of guidelines that the articles could be compared against. The remaining paper (Mårtensson et al. 2016) is a guide for how to evaluate academic papers. While primarily targeted towards reviewers, especially peer reviewers, the article still provide ways to evaluate the selected papers, especially because it gives an idea of what peer reviewers are supposed to look for when evaluating papers.

The guides were searched for recommendations related to research questions. These were then combined to create a set of assessment criteria for the selected papers. These criteria are listed in Table 1, alongside the sources they are based on. As the criteria are based on guides for both writing and evaluating academic papers, these make a nice, well-rounded framework for evaluating whether the selected articles present and use research questions the way they should.

The selected topic for this paper revolves around the actual structure of the selected papers, rather than the methodology used in the papers. As such, the same assessment criteria are used for papers using both quantitative and qualitative methods.

Research questions	
Does the paper contain one or more research questions or hypotheses?	Cuschieri et al. 2019; Davidson et al. 2012; Jha 2014; Katz 2006; Lin and Kuo 2012; Mårtensson et al. 2016; Rosenfeldt et al. 2000
How are the research question(s) presented?	Cuschieri et al. 2019; Davidson et al. 2012; Jha 2014; Katz 2006; Lin and Kuo 2012; Rosenfeldt et al. 2000
Structure	
Where in the paper are the research question(s) introduced?	Cuschieri et al. 2019; Davidson et al. 2012; Jha 2014; Katz 2006; Lin and Kuo 2012; Rosenfeldt et al. 2000
How do the results presented in the paper answer the research question(s)?	Cuschieri et al. 2019; Davidson et al. 2012; Jha 2014; Katz 2006; Mårtensson et al. 2016
How do the conclusions drawn in the paper connect with the research question(s) presented?	Davidson et al. 2012; Katz 2006; Mårtensson et al. 2016

Table 1: Assessment criteria for the selected papers

4 Evaluation of selected papers

The selected papers are evaluated, using the criteria presented in Section 3, in the same order as they were introduced in Section 2.1.

4.1 Presentation of Barredo Arrieta et al. (2020)

The first paper, ‘Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI’ by Barredo Arrieta et al. (2020), is a conceptual paper explaining the field of explainable artificial intelligence (XAI). The paper starts out by clarifying concepts and definitions for XAI. It introduces the concept of *audience* to XAI, changing the definition of explainable AI to “Given an audience, an explainable Artificial Intelligence is one that produces details or reasons to make its functioning clear or easy to understand” (Barredo Arrieta et al. 2020, p.85).

The authors goes on to discuss current methods for making artificial intelligence systems more explainable, before it performs a thorough literature review in the field of XAI, splitting the current research in two parts – one focused on general explanation methods, the other focused solely on methods aimed at explaining deep learning systems. With the review fresh in memory, the paper discusses challenges and future research opportunities within the field of XAI, mentioning cases such as interpretability possibly lowering performance, the lack of consensus regarding the definition of explainability, and the current lack of methods for explaining deep learning models.

The paper goes on to introduce the concept of responsible artificial intelligence, by building upon the research done on XAI, adding fairness, accountability and privacy to the previously used definition. The paper rounds out by welcoming reflections on XAI in sensitive scenarios, finishing by discussing how XAI methods can work in privacy- and security-focused systems.

4.2 Evaluation of Barredo Arrieta et al. (2020)

While it contains novel definitions, exhaustive literature reviews and important insights within the field of responsible artificial intelligence, Barredo Arrieta et al. (2020) does not contain any concrete research questions. Instead, the aim of the research is covered in a single sentence, where it is stated that "this overview aims to cover the creation of a complete unified framework of categories and concepts that allow for scrutiny and understanding of the field of XAI methods" (Barredo Arrieta et al. 2020, p.83). Including the aims of the research as part of a body of text, rather than explicitly listed, aligns with recommendations from Davidson et al. (2012, p.359), although with a requirement that "At the end of the Introduction, the reader should have a clear idea of what the research question is, why it is important and what the investigators aimed to do in the study" (Davidson et al. 2012, p.360).

In addition to the single-sentence summary of the aims of the article, Barredo Arrieta et al. (2020) rounds out the introductory chapter with a comprehensive list of all contributions of the article, written as a numbered list. This way of summarizing the contributions of an article, rather than the goals, makes it represent the actual findings of the research. As such, it more closely aligns with the conclusion, rather than the introduction. This difference in alignment is explicitly mentioned by both Jha (2014, p.2) and Lin and Kuo (2012, p.83), who recommends to "[g]ive in the introduction only the strictly pertinent references and do not include the data or conclusions from the work being reported" and to "not include any of the data or conclusions from your study in the Introduction", respectively.

As mentioned in Section 1, the problem of "publish or perish" is twofold, split between reproducibility and an increasing number of papers. Including a summary of contributions in the introduction, as done by Barredo Arrieta et al. (2020), does not contribute to solving or mitigating the first part of the problem, as it does not aid reproducibility. It can, however, mitigate the latter part of the problem, as including the contributions in the introduction allows a reader to quickly assess whether the article is relevant for their own research. In fact, including a summary of the results and conclusions in the abstract is recommended by both Jha (2014, p.2), Lin and Kuo (2012, p.82), Cuschieri et al. (2019, p.114), Davidson et al. (2012, p.359), Katz (2006, p.115) and Rosenfeldt et al. (2000, p.85). Although a short summary is already part of the abstract of Barredo Arrieta et al. (2020, p.82), the complete summary of contributions is longer, and goes in greater detail on what the article brings to the field of responsible artificial intelligence. As most abstracts are given strict word limitations (Rosenfeldt et al. 2000, p.85), the authors may have opted to include the summary in the introduction to be able to fully point out their contributions.

The article largely sticks close to its original aim, mentioned above. Section 2 of Barredo Arrieta et al. (2020) explains categories related to explainable AI and connects definitions found in other papers to their own definition of XAI. Section 3 and 4 connects existing concepts and methods within the field of XAI to the work done in section 2, while section 5 is focused on the status of the research being done in the field of XAI, fulfilling the "allow for scrutiny"-part of the aim. The only parts of the paper that does not directly connect to the stated aim of the article is section 6, which instead focuses on principles for responsible artificial intelligence. It can be argued that this is relevant to the aim of "understanding of the field of XAI methods", something the authors do

in their introduction of the section, stating that "[section 6] presents some of the most important and widely recognized principles [of responsible artificial intelligence] in order to link XAI — which normally appears inside its own principle — to all of them" (Barredo Arrieta et al. 2020, p.103). However, as the primary focus of the section shifts from XAI to responsible artificial intelligence, the final section of the paper moves outside of the stated research goal.

The conclusion sticks relatively closely to the research goal, as it follows the same topics as discussed in the rest of the article. Most of it summarizes the work on AI, discussing the work done and key contributions within each section. While one paragraph touches on responsible artificial intelligence, thus moving slightly outside of the stated research goal, the main focus is on how XAI is a core component of any responsible artificial intelligence system, thus connecting it back with the stated aim for the paper.

While the content of the paper aligns with the stated goal, Barredo Arrieta et al. (2020) contains no explicit references to it, instead connecting their work to their goal through the overarching topic of XAI. An example of this can be seen in the introduction of section 5, where the authors state that the section will be used to "revisit [the challenges] and explore new research opportunities for XAI, identifying possible research paths that can be followed" (Barredo Arrieta et al. 2020, p.99). While the quote makes no explicit references to the research goal, it is connected through the topic of XAI. This same method of connecting through the topic can also be seen in the conclusion, where several references are made to the topic – such as "Implications of XAI in fairness have also been discussed in detail" and "[Our reflections] agree on the compelling need for a proper understanding of the potentiality and caveats opened up by XAI techniques" (both from Barredo Arrieta et al. 2020, p.108). Neither of these quotes mention the aim of the paper, but it is still clearly connected through the topic.

This way of connecting the research to the research goal goes against several of the selected writing guides. Cuschieri et al. (2019, p.115) states that "The factual answer to the research question that was described in the 'Introduction' section needs to be clearly illustrated" and "The discussion must aim to answer the research questions that were posed in the 'Introduction' section". Davidson et al. (2012, p.360) recommends to "Start the Discussion with the answer to the research question", and Jha (2014, p.2) argues to "Begin the discussion with brief recapitulation of the main findings (the answer to the research question)". Finally, Lin and Kuo (2012, p.84) argues that "the first paragraph [of the discussion] should discuss the major findings and state whether the hypotheses were supported or rejected." Although the aim of Barredo Arrieta et al. (2020) is not formulated as something that can be supported or rejected, the point still stands that the discussion should explicitly discuss the research goal. To summarize, most current guides for academic writing suggests to explicitly connect the results and discussion to the given research goal, something Barredo Arrieta et al. (2020) fails to do.

In conclusion, Barredo Arrieta et al. (2020) does not explicitly state their research questions, instead including the aim of their research as part of a paragraph in their introduction. While the aim could be expressed more explicitly or highlighted in their introduction, the authors stick to the original goal of their paper, largely aligning their work with the stated aim. The goal is used throughout most of the paper, with four of five sections, as well as the full concluding chapter, sticking close to it. While they do not explicitly connect their research to the stated goal, this connection is made through the overarching topic of XAI, ensuring an alignment between the research and the research goal.

4.3 Presentation of Vakkuri et al. (2020)

The second paper, 'The Current State of Industrial Practice in Artificial Intelligence Ethics' by Vakkuri et al. (2020), is a quantitative paper looking at how and to what extend AI ethics are applied within the AI industry. The paper starts by describing the attractiveness of artificial intelligence, and how this has contributed, and is likely to continue contributing, to an increasing number of ethical incidents involving the technology, before a brief discussion on how their survey – using respondent data to analyze the practice of AI ethics – is novel and has never before been done.

The paper contains data gathered from 249 employees from 211 companies, primarily located in the US or Finland (Vakkuri et al. 2020, p.52). Over 65 per cent of the respondents answered that they were able to influence the functionality of their systems, and of the responding companies, over half were actively developing artificial intelligence systems. The authors note that the answers were similar whether the employing company were developing artificial intelligence or not, and independent of the geographic location of the company. All responses are therefore included in the paper.

Vakkuri et al. (2020) rounds out the paper by discussing what the presented results means for other software development companies. They connect AI ethics to other ethical trends, such as data privacy and ecological issues, and briefly discuss how AI ethics can create business advantages. The authors go on to mention several tools for implementing AI ethics in a company, such as existing guidelines, but point out that there is limited access to general use, off-the-shelf solutions for AI ethics. Finally, the paper goes on to mention several anti-patterns companies looking to implement AI ethics should avoid – such as outsourcing ethics or delegating ethics work to a single individual – before rounding out by a discussion of how the evolution of AI ethics is likely to impact the software development industry.

4.4 Evaluation of Vakkuri et al. (2020)

Although Vakkuri et al. (2020) present novel insights into the current state of AI ethics in the software industry, they do so without any explicit mention of research goals or -questions. Much like Barredo Arrieta et al. (2020), the authors include the intent of their questionnaire as part of a sentence in the introduction – “To provide needed insight into the current state of practice in the industry” (Vakkuri et al. 2020, p.51). This states the purpose of the research, and as such satisfies recommendations from Davidson et al. (2012, p.359), but the lack of explicit research questions makes it harder to “have a clear idea of what the research question is, why it is important and what the investigators aimed to do in the study” (Davidson et al. 2012, p.360). In order to understand the concrete research goals of the authors, a reader is then forced to read between the lines. Doing so creates three possible sources of research goals, inspired by the three different parts of the paper where the research questions are expected to be involved – the title (Davidson et al. 2012, p.359), the abstract (Cuschieri et al. 2019, p.115) and the introduction (Jha 2014, p.2).

The title should be “relevant to the research question, and accurately describe what the study is about” (Davidson et al. 2012, p.359). As such, it, combined with subtitles throughout the text, can be used as a source for research goals. In Vakkuri et al. (2020), the title, which reads “The Current State of Industrial Practice in Artificial Intelligence Ethics” clearly indicates that the goal of the research is to learn more about how the software development industry is currently practicing AI ethics, and thus aligns with the aim mentioned above. This is further supported by the three subtitles used throughout the text – “What Is AI Ethics?”, “What Is Actually Happening in the Industry?” and “What Should Your Organization Do?” (Vakkuri et al. 2020, p.51, 52 and 54, respectively). The three subtitles are all formulated as questions, and clearly align with the aim of the survey. As such, they are used as foundation for potential research question (PRQ) 1-3:

- PRQ 1 What is AI ethics?
- PRQ 2 How are AI ethics implemented and used in the software development industry?
- PRQ 3 How does this impact other organizations in the software development industry?

Cuschieri et al. (2019, p.115) states that the abstract “needs to contain a clear statement of the research question/s and the purpose for conducting the research”, making it a potentially good source for implicit research questions. Vakkuri et al. (2020, p.50) state in their abstract that their research “provide insights into the current state of industrial practice”. This closely aligns with both the title and subtitles, and supports PRQ 2. The abstract is short, and thus contains no support for PRQ 1 or 3.

According to Jha (2014, p.2), an introduction should “state the aim, the research question, and the study design”. Although Vakkuri et al. (2020) does not explicitly state their research questions, the introduction may still be used to discover potential research questions. The authors state

that "we know little of the current state of practice of ethics in AI", following up with asking whether "the public and academic discussion in the area [have] motivated smaller industry players to develop more ethical AI" (Vakkuri et al. 2020, p.50). Combined with their stated contributions – "helping us understand where we currently are as an industry in terms of AI ethics" and "a way to benchmark where [other organizations] stands" (Vakkuri et al. 2020, p.51), this strongly supports the representativeness of PRQ 2 and 3. The introduction discusses the background of AI ethics, including reasons why it is needed, but otherwise do not indicate support for PRQ 1. The rest of this analysis will therefore use PRQ 2 and 3 as representative research questions for the paper.

The survey conducted by Vakkuri et al. (2020) asked employees about the liabilities and responsibilities of their systems, transparency towards both users, public authorities and developers, and organizational plans for handling unexpected use and potential misuse of their system. Most of these topics are closely related to AI ethics, with liability, responsibility and transparency being listed as key principles of ethical AI (Vakkuri et al. 2020, p.51). As the questionnaire is sent to practitioners within the industry of software development, it is clear that the results produced are closely aligned with PRQ 2, directly answering the question.

The connections between the results and PRQ 3 are not as clear to see. It can be argued that there is an indirect connection, as answers to PRQ 2 at least partially answer PRQ 3, due to the actions of one organization naturally impacting the expectations and behavior of others. This is also noted by Vakkuri et al. (2020, p.55), stating that: "[as] users become increasingly conscious about privacy issues, being ethical in relation to data privacy, for example, can become a selling point." There are, however, no clear direct connections between PRQ 3 and the survey or its responses.

In the conclusion, however, clear connections to PRQ 3 can be observed. As mentioned in Section 4.3, Vakkuri et al. (2020) points out guidelines and tools to use for implementing AI ethics, clear antipatterns to avoid when doing so, potential developments of the industry and regulations that can increase the pressure to implement AI ethics, and ways it can be beneficial to do so already, even before laws or regulations are enacted. The authors summarize the recommendation for other organizations in the final sentence of the paper, stating: "Acting on AI ethics today will reap future rewards" (Vakkuri et al. 2020, p.56). As such, the conclusion directly and explicitly answers PRQ 3.

This same connection between conclusion and PRQ 2 can also be observed. Vakkuri et al. (2020) perform an in-depth analysis of the responses to their survey, amongst other discussing how liability is a somewhat ignored ethics principle within the software development industry, while predictability is taken more seriously. Likewise, they observe that a large part of the companies consider complying with regulatory standards to be enough, with regards to both technical details of their system, such as responsibility, as well as for documentation of their system. These considerations and observations shows that the conclusions drawn to a large degree answers PRQ 2.

In conclusion, Vakkuri et al. (2020) state the aim of their survey, but do not explicitly state their research questions, instead requiring the reader to interpret them from the rest of the text. When this is done, however, the connections with the rest of the text are mostly clear – the survey performed in the paper directly answers one of the potential research questions, while the discussion and conclusions being drawn afterwards directly answer both the potential research questions. While not explicitly referenced, there is a clear focus on the potential research questions throughout the text, ensuring an alignment between the paper and the aim of the paper.

4.5 Presentation of Ntoutsis et al. (2020)

The third paper, 'Bias in data-driven artificial intelligence systems—An introductory survey' by Ntoutsis et al. (2020), is a qualitative analysis of existing approaches for handling bias and fairness in artificial intelligence systems. The first section of the paper starts by introducing the concept of bias, including a brief history of bias within the field of artificial intelligence, a definition of bias to be used throughout the paper, and causes of bias. The section goes on to discuss how bias is represented in data, before it discusses the definition of fairness – and related challenges.

Finally, the first section rounds out by discussing legal issues related to bias and fairness in artificial intelligence systems.

The second part of the paper discusses ways to mitigate bias. The mitigation methods are divided in three. Preprocessing approaches look at data, and how bias can be mitigated by adjusting the dataset being supplied into an AI model. In-processing approaches look at the model itself, and how bias mitigation can be made part of the core of the process. Post-processing models look at how bias can be mitigated by adjusting the outputs of the model, after it has made a decision, but before the decision is finalized. The part is finalized by a discussion of legal issues related to bias mitigation.

The final part of the paper looks at ways to handle bias in an AI system. It starts by looking at ways to proactively handle bias, by being aware of it when collecting data, concluding that it is nearly impossible to guarantee completely bias-free data collection methods. Next, methods for how to describe bias is discussed, primarily focused on the use of ontologies – formal descriptions and representations of bias in a system. Finally, methods for retroactively describing decisions made by an AI system are discussed, highlighting many of the same methods discussed by Barredo Arrieta et al. (2020), before a discussion of legal issues related to handling bias. The paper is rounded out by a discussion of the direction of bias research going forward.

4.6 Evaluation of Ntoutsis et al. (2020)

Following the recommendations from Cuschieri et al. (2019) and Davidson et al. (2012), Ntoutsis et al. (2020) include an explicit formulation of the overarching goal of their research. Introduced in the abstract, the goal of Ntoutsis et al. (2020, p.1) is to "provide a broad multidisciplinary overview of the area of bias in AI systems, focusing on technical challenges and solutions as well as to suggest new research directions towards approaches well-grounded in a legal frame." Including this in the abstract aligns with Cuschieri et al. (2019, p.115), who state that the abstract should include "the purpose for conducting the research," as well as Davidson et al. (2012, p.359), who recommend that the abstract should "succinctly describe the aim of the study." On the other hand, Jha (2014), Katz (2006), Lin and Kuo (2012) and Rosenfeldt et al. (2000) suggests moving the aim to the introduction, instead prioritizing the results and conclusions of the research in the abstract. Cuschieri et al. (2019, p.115) also recommend to include the results and main conclusion in the abstract, alongside the aim. Neither results nor conclusions are mentioned in the abstract of Ntoutsis et al. (2020).

Although they explicitly state their research goal, Ntoutsis et al. (2020) do not mention their research questions anywhere in their paper. This makes the reading process similar to the one mentioned in Section 4.4, where the author's questions must be understood from the stated research goal and the parts of the text where research questions are expected to be involved – the title (Davidson et al. 2012, p.359), the abstract (Cuschieri et al. 2019, p.115) and the introduction (Jha 2014, p.2). While the introduction of Ntoutsis et al. (2020) does not contain explicit research questions, it does contain a division of their work into three categories: *Understanding bias*, *Mitigating bias* and *Accounting for bias* (Ntoutsis et al. 2020, p.2). Much like the process described in Section 4.4, these three categories can be used as foundation for three potential research questions (PRQs):

- PRQ 1 What is bias, and how does it affect AI systems?
- PRQ 2 How can bias in AI systems be mitigated?
- PRQ 3 How can bias in AI systems be accounted for?

In order to be able to use these potential research questions throughout the evaluation, they must align with the overarching research goal, as well as the two remaining sections mentioned – the title and the abstract. The abstract of Ntoutsis et al. (2020) has already been discussed, as it contains the overarching goal of the research. As answering the presented PRQs gives a broad overview of the field of bias in AI systems, the goal – and thus the abstract – closely aligns with the presented PRQs. Similarly, the title of Ntoutsis et al. (2020), which reads "Bias in data-driven artificial intelligence systems—An introductory survey," is almost a repetition of the research goal, and thus closely aligns with the presented PRQs. We therefore continue the evaluation using PRQ

1-3 as representative research questions for the paper.

As the PRQs are inferred from the categories put forward by Ntoutsis et al. (2020), the connections between them and the results found in the paper are clear, with each PRQ being answered by a specific section. Section 2, discussing bias, aims to answer PRQ 1. The section proposes a definition for bias, thus answering the question of "What is bias," before answering the rest of the PRQ by looking at ways bias impacts AI systems. Section 3 discusses ways to mitigate bias, and thus fully answer PRQ2. Likewise, section 4 discusses ways to account for bias in AI systems, thus fully answering PRQ 3. Although no explicit mentions of the categories are made, the connection is still implicitly there, as the results align perfectly with the PRQs. The paper does not mention the PRQs in the results, which is natural, as they are not introduced in the paper, but instead developed by the reader.

The conclusion of Ntoutsis et al. (2020) is primarily focused on the path for future research. As such, it naturally focuses on how current approaches to bias mitigation and handling can be reviewed and improved, thus aligning itself with PRQ 2 and 3. The conclusion calls for "a systematic evaluation of the existing approaches" for mitigating bias and the creation of "benchmark datasets" for testing bias mitigation (Ntoutsis et al. 2020, p.9-10), both of which would further contribute to answering PRQ 2. Similarly, recommendations are made for researching bias handling and mitigation in more complicated AI systems, such as unsupervised or reinforcement learning, which would contribute to answering both PRQ 2 and 3. Finally, Ntoutsis et al. (2020, p.10) argues that people working on AI systems should "be aware of bias-related issues and the effect of their design choices and assumptions," and that "it is a key responsibility of technology creators to understand [technology's] limits and to propose safeguards to avoid pitfalls." These can both be considered preemptive bias mitigation or bias accounting, thus directly connecting with PRQ 2 and 3.

The connection between the PRQ 1 and the conclusion, however, is not as clear. Three such connections were found throughout the concluding section. First, the conclusion discusses differing definitions of fairness, and how collectively agreeing on a given definition is difficult. As bias is defined as "inclination or prejudice of a decision [...] which is for or against one person or group, especially in a way considered to be unfair" Ntoutsis et al. (2020, p.3), having a clear definition of fairness – and thus unfairness – affects the definition of bias, connecting it to the first part of PRQ 1. Second, Ntoutsis et al. (2020, p.10) stress that "everyone involved in the [AI] decision making process should be aware of bias-related issues." While this does not answer PRQ 1, it connects back to the second part of the question, i.e., how bias affects AI systems. Finally, Ntoutsis et al. (2020, p.10) state that "biases are deeply embedded in our societies," and that the problem of bias cannot be solved by technology alone. By discussing where biases come from, as well as how they are found throughout society, this connects back to both parts of PRQ 1.

In conclusion, Ntoutsis et al. (2020) explicitly states the overarching goal of their research in the abstract, thus making it easy for readers to understand what the authors are trying to achieve. Although the authors do not explicitly mention their research questions, they are relatively simple to discern from the stated overarching goal, combined with the categorization presented in the introduction. Doing so presents three potential research questions, here named PRQ 1-3. The results of the research, as well as the conclusions drawn by Ntoutsis et al. (2020), closely aligns with the research questions. This signals that the PRQs are close to the original goals of the authors, and shows a clear alignment between the stated overarching research goal and the work done in the paper.

5 Discussion

The recommendations from the selected writing- and reviewing guides are unified – "Research emanates from at least one [research question]" (Mårtensson et al. 2016, p.596), "The 'objective' section needs to contain a clear statement of the research question/s" (Cuschieri et al. 2019, p.115) and "Before writing the paper, think carefully about what the research question is, and when writing keep the paper focused on the question." (Davidson et al. 2012, p.357), to quote a few – in that research questions are essential for writing academic papers. Yet, interestingly, none of the

evaluated papers included their research questions, despite being approved by established journals. This section will discuss potential reasons for this, as well as ways to ensure future papers include their research questions, as doing so will not only ensure compliance with recommendations for academic writing, but also, more importantly, can be used as a way to mitigate the problem of "publish or perish", as discussed in Section 1.

5.1 Reasons for failure to include research questions

There may be several reasons why all three papers failed to include their research questions. In the author guidelines of the three journals the evaluated papers were published in, neither Elsevier (n.d.), where Barredo Arrieta et al. (2020) was published, IEEE (n.d.), where Vakkuri et al. (2020) was published, nor WIREs - Wiley Interdisciplinary Reviews (n.d.), where Ntoutsis et al. (2020) was published, requires authors to include their research questions when submitting papers. This leaves us with a simple explanation for their exclusion – papers have a length limit (IEEE n.d.), and including content that is lengthy but not required limits the amount of novel and interesting content that can be included. With this explanation, the responsibility for authors not including their research questions falls on the journals, as they create the requirements for the submitted papers. Future research should look at compare author guidelines in different journals, and check to what degree their differences affect the accepted papers. If the explanation is found to be right, journal editors should carefully consider whether they should make adjustments to their guidelines, such as by requiring research questions as part of the introduction, or if the current guidelines produce the best results possible.

The missing compliance between the recommendations and the submitted papers may also be due to academic writing guides being outdated. All papers included the overarching goal of their research, with Barredo Arrieta et al. (2020) additionally including a summarizing list of their contributions. The authors may have felt like this was enough to "describe the aim of the study" (Davidson et al. 2012, p.359), explain "the purpose for conducting the research" (Cuschieri et al. 2019, p.115) and show they have "[identified] and have the idea that [they] wants to communicate through the paper" (Lin and Kuo 2012, p.81), as the various guides recommend. To assess whether this holds true, future research can employ quantitative methods, to evaluate larger volumes of papers than these three. If this is the case, and modern researchers prefer other ways of representing their research goals, new guides for academic writing should be constructed. This process should ideally be conducted by experts on academic writing, in collaboration with editors of major journals, to create accurate, relevant and generalizable guides for future papers.

5.2 Research questions and quality

Although several models for assessing the quality of an academic paper has been created (i.e., Klein and Myers 1999; Robey and Markus 1998; H. J. Rubin and I. S. Rubin 2011), there is still a "lack of widely acknowledged quality standards," and "a need for determining a universal concept model for the quality of research practice" (Mårtensson et al. 2016, p.595). As such, drawing exact connections between the inclusion of research questions and quality is challenging.

Still, the included guides for academic writing are clear that research question are essential for writing good papers. Cuschieri et al. (2019, p.114) states that an author "needs to have a clear vision of the aim and scopes of [their] paper" in order to "ensure the execution of a high quality paper that is likely to be accepted for publication." Likewise, Barroga and Matanguihan (2022, p.9) states that "it is crucial to have excellent research questions to generate superior hypotheses," which will "determine the research objectives and the design of the study, and ultimately, the outcome of the research." Including explicit research questions can be seen as a formalized method for stating the aim of a paper, thus contributing to the quality of the result. Likewise, including research questions in a paper contribute to increase the readability, and increases the paper's alignment with current guides for academic writing.

5.3 Methods to ensure maximum quality of papers

Depending on which of the theories presented in Section 5.1 are found to be correct, there are several methods that can be employed to ensure the quality of academic papers is as high as possible. First, journals can change their author guidelines, to recommend – or possibly force – authors to include research questions in their submitted papers. As this paper shows a connection between research question use and readability, and thus paper quality, doing so would ensure the quality stays as high as possible.

Journals and academic institutions have the option of going one step further than this, requiring authors to register their research questions ahead of starting their research. This aligns with the suggestions put forth by Baker (2016), and thereby contributes twofold. First, pre-registering research questions could ensure researchers avoid nit-picking results, and thus increase reproducibility and lead to accurate and honest results. Secondly, such a pre-registration would ensure research questions are available when submitting papers, thus aiding the readability and quality of the submitted paper.

Finally, researchers, journal editors and experts on academic writing could consider alternative structures for academic papers. Instead of listing their research questions, Barredo Arrieta et al. (2020) finishes their introduction with a list of their contributions. While these contributions can only be seen after the research has finished, and as such are not available for pre-registering, including contributions in a structure way leads to the same readability benefits as mentioned above, and makes it easy for readers to quickly assess whether the paper is relevant for their own research. In fact, Chen et al. (2022) argues that contributions are the most important part of any research. To further explore this method, future research should try to find alternative paper structures than what is currently employed within academia, and evaluate whether such alternative structures can increase the readability and quality of academic papers.

Although the use of research questions was similar between the quantitative and qualitative papers evaluated in this paper, there could be differences between the methodologies when it comes to the optimal use of such questions. Barroga and Matanguihan (2022, p.3-4) states that quantitative research questions are "precise" and "usually framed at the start of the study," while qualitative research questions "broadly explore a complex set of factors" and "are usually continuously reviewed and reformulated." Based on this, it is natural to assume that clear research questions possibly are more important in quantitative works – where they can ensure that the researchers do not adjust their goals after seeing the data, thus cherry-picking results – compared to qualitative works, where researchers are typically looking for connections and patterns and the research questions are more related to the topic of research, rather than concrete connections. Future research should compare differences in the use of research questions, and whether such use impacts the quality of quantitative and qualitative papers differently. The results from such findings may be used to develop separate guides for the different methodologies, to ensure maximum quality for all papers, no matter the method used.

6 Conclusion

None of the evaluated papers explicitly stated their research questions, but all included an overarching goal for their research. Even without explicit research questions, all papers had a clear connection between this goal and their result and conclusion. As such, some may argue that research questions are outdated, and that it is more important to include a clear summary of the overarching goal and contributions. While there are no clear connections between the use of research questions and paper quality, including such questions could lead to better readability, and help mitigate the problem of "publish or perish" by creating a source for pre-registration, as well as by making it easier for readers to quickly assess the relevancy of a paper for their own research.

Bibliography

- Allcott, Hunt and Matthew Gentzkow (2017). ‘Social media and fake news in the 2016 election’. In: *Journal of Economic Perspectives* 31.2, pp. 211–236. DOI: 10.1257/jep.31.2.211.
- Anderegg, William R. L. et al. (2010). ‘Expert credibility in climate change’. In: *Proceedings of the National Academy of Sciences of the United States of America* 107.27, pp. 12107–12109. DOI: 10.1073/pnas.1003187107.
- Baker, M. (May 2016). ‘1,500 scientists lift the lid on reproducibility’. In: *Nature* 533.7604, pp. 452–454. ISSN: 1476-4687. DOI: 10.1038/533452a.
- Barredo Arrieta, A. et al. (2020). ‘Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI’. In: *Information Fusion* 58, pp. 82–115. DOI: 10.1016/j.inffus.2019.12.012.
- Barroga, Edward and Glafera Janet Matanguihan (2022). ‘A Practical Guide to Writing Quantitative and Qualitative Research Questions and Hypotheses in Scholarly Articles’. In: *Journal of Korean Medical Science* 37.16. DOI: 10.3346/jkms.2022.37.e121.
- Begley, C.G. and J.P.A. Ioannidis (2015). ‘Reproducibility in science: Improving the standard for basic and preclinical research’. In: *Circulation Research* 116.1, pp. 116–126. DOI: 10.1161/CIRCRESAHA.114.303819.
- Biddle, S.J.H. and M. Asare (2011). ‘Physical activity and mental health in children and adolescents: A review of reviews’. In: *British Journal of Sports Medicine* 45.11, pp. 886–895. DOI: 10.1136/bjsports-2011-090185.
- Bornmann, L. and R. Mutz (2015). ‘Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references’. In: *Journal of the Association for Information Science and Technology* 66.11, pp. 2215–2222. DOI: 10.1002/asi.23329.
- Chen, H., H. Nguyen and A. Alghamdi (2022). ‘Constructing a high-quality dataset for automated creation of summaries of fundamental contributions of research articles’. In: *Scientometrics*. DOI: 10.1007/s11192-022-04380-z.
- Cuschieri, S., V. Grech and C. Savona-Ventura (2019). ‘WASP (Write a Scientific Paper): Structuring a scientific paper’. In: *Early Human Development* 128, pp. 114–117. DOI: 10.1016/j.earlhumdev.2018.09.011.
- Davidson, A., D. McD Taylor and F.E. Babl (2012). ‘Review article: A primer for clinical researchers in the emergency department: Part III: How to write a scientific paper’. In: *EMA - Emergency Medicine Australasia* 24.4, pp. 357–362. DOI: 10.1111/j.1742-6723.2012.01569.x.
- Davies, B. and G. Felappi (2017). ‘Publish or Perish’. In: *Metaphilosophy* 48.5, pp. 745–761. DOI: 10.1111/meta.12269.
- Domholdt, E.A. and T.R. Malone (1985). ‘Evaluating research literature: The educated clinician’. In: *Physical Therapy* 65.4, pp. 487–491. DOI: 10.1093/ptj/65.4.487.
- Elsevier (n.d.). *Guide for authors*. URL: <https://www.elsevier.com/journals/information-fusion/1566-2535/guide-for-authors>.
- Grimes, D.R., C.T. Bauch and J.P.A. Ioannidis (2018). ‘Modelling science trustworthiness under publish or perish pressure’. In: *Royal Society Open Science* 5.1. DOI: 10.1098/rsos.171511.
- IEEE (n.d.). *How to get published: Author guidelines for IEEE Computer Society Publications*. URL: <https://www.computer.org/publications/author-resources/authors>.
- Jha, Kirti Nath (2014). ‘How to write articles that get published’. In: *Journal of Clinical and Diagnostic Research* 8.9, XG01–XG03. DOI: 10.7860/JCDR/2014/8107.4855.
- Katz, M.J. (2006). *From research to manuscript: A guide to scientific writing*. Springer Netherlands, pp. 1–152. DOI: 10.1007/1-4020-4071-7.
- Klein, Heinz K. and Michael D. Myers (1999). ‘A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems’. In: *MIS Quarterly* 23.1, pp. 67–93. ISSN: 02767783. URL: <http://www.jstor.org/stable/249410> (visited on 15th Nov. 2022).
- Lin, P.Y. and Y.R. Kuo (2012). ‘A guide to write a scientific paper for new writers’. In: *Microsurgery* 32.1, pp. 80–85. DOI: 10.1002/micr.20960.
- Lindholt, Marie Fly et al. (2021). ‘Public acceptance of COVID-19 vaccines: Cross-national evidence on levels and individual-level predictors using observational data’. In: *BMJ Open* 11.6. DOI: 10.1136/bmjopen-2020-048172.

-
- Mårtensson, P. et al. (2016). 'Evaluating research: A multidisciplinary approach to assessing research practice and quality'. In: *Research Policy* 45.3, pp. 593–603. ISSN: 0048-7333. DOI: <https://doi.org/10.1016/j.respol.2015.11.009>.
- Mayhew, P.A. (1993). 'Evaluating research for use in practice.' In: *Medsurg nursing : official journal of the Academy of Medical-Surgical Nurses* 2.6, pp. 496–498.
- Mittal, A., A.K. Moorthy and A.C. Bovik (2012). 'No-reference image quality assessment in the spatial domain'. In: *IEEE Transactions on Image Processing* 21.12, pp. 4695–4708. DOI: 10.1109/TIP.2012.2214050.
- Moor, J. (2006). 'The Dartmouth College Artificial Intelligence Conference: The next fifty years'. In: *AI Magazine* 27.4, pp. 87–91.
- 'Fostering reproducible fMRI research' (Mar. 2017). In: *Nature Neuroscience* 20.3. Ed. by Nature Neuroscience, pp. 298–298. ISSN: 1546-1726. DOI: 10.1038/nn.4521.
- Nightingale, Alison (2009). 'A guide to systematic literature reviews'. In: *Surgery* 27.9, pp. 381–384. DOI: 10.1016/j.mpsur.2009.07.005.
- Ntoutsis, Eirini et al. (2020). 'Bias in data-driven artificial intelligence systems—An introductory survey'. In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 10.3. DOI: 10.1002/widm.1356.
- Okoli, C. (2015). 'A guide to conducting a standalone systematic literature review'. In: *Communications of the Association for Information Systems* 37.1, pp. 879–910. DOI: 10.17705/1cais.03743.
- Polanczyk, G.V. et al. (2015). 'Annual research review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents'. In: *Journal of Child Psychology and Psychiatry and Allied Disciplines* 56.3, pp. 345–365. DOI: 10.1111/jcpp.12381.
- Popenoe, R. et al. (2021). 'A practical guide to data analysis in general literature reviews'. In: *Nordic Journal of Nursing Research* 41.4, pp. 175–186. DOI: 10.1177/2057158521991949.
- Robey, Daniel and M. Lynne Markus (1998). 'Beyond Rigor and Relevance: Producing Consumable Research about Information Systems'. In: *Information Resources Management Journal (IRMJ)* 11.1, pp. 7–16. DOI: 10.4018/irmj.1998010101.
- Rosenfeldt, Franklin L. et al. (2000). 'How to write a paper for publication'. In: *Heart Lung and Circulation* 9.2, pp. 82–87. DOI: 10.1046/j.1444-2892.2000.00031.x.
- Rubin, Herbert J and Irene S Rubin (2011). *Qualitative interviewing: The art of hearing data*. sage.
- Vakkuri, Ville et al. (2020). 'The Current State of Industrial Practice in Artificial Intelligence Ethics'. In: *IEEE Software* 37.4, pp. 50–57. DOI: 10.1109/MS.2020.2985621.
- Wade, M. and J. Hulland (2004). 'Review: The resource-based view and information systems research: Review, extension, and suggestions for future research'. In: *MIS Quarterly: Management Information Systems* 28.1, pp. 107–142. DOI: 10.2307/25148626.
- Wang, Z. and A.C. Bovik (2002). 'A universal image quality index'. In: *IEEE Signal Processing Letters* 9.3, pp. 81–84. DOI: 10.1109/97.995823.
- WIREs - Wiley Interdisciplinary Reviews (n.d.). *Author Guidelines*. URL: https://wires.onlinelibrary.wiley.com/hub/for_authors.

A Search history

This appendix includes a complete history of search terms used in the literature searches for this paper. Searches were conducted using the Scopus database (S), Web of Science (W) and Google Scholar (G).

A.1 Relevant papers

(S) TITLE-ABS-KEY ("responsible artificial intelligence")
(S) TITLE-ABS-KEY ("responsible artificial intelligence" OR "responsible ai")
(S) TITLE-ABS-KEY (ai OR "artificial intelligence" OR "machine learning")
(W) TS=(AI)
(W) TS=("Artificial intelligence")
(S) TITLE-ABS-KEY ((ai OR "artificial intelligence" OR "machine learning") AND business*)
(W) TS=(AI) AND TS=(business*)
(S) TITLE-ABS-KEY ((ai OR "artificial intelligence" OR "machine learning") AND implementation)
(S) TITLE-ABS-KEY ((ai OR "artificial intelligence" OR "machine learning") AND management)
(W) TS=(AI) AND TS=(business* OR management)
(W) TS=("Artificial intelligence") AND ALL=(management OR business* OR governance)
(S) TITLE-ABS-KEY ("artificial intelligence" AND strategy)
(S) TITLE-ABS-KEY (business AND strategy) AND (LIMIT-TO (SUBJAREA , "COMP"))
(W) TS=(business AND strategy) AND WC=(Computer Science, Artificial Intelligence)
(S) TITLE-ABS-KEY ("responsible AI" OR "ethical AI")

A.2 Assessment criteria and paper quality

(S) TITLE-ABS-KEY ("research question")
(S) TITLE-ABS-KEY ("article structure")
(S) TITLE-ABS-KEY (article AND structure)
(S) TITLE-ABS-KEY ("article quality")
(S) TITLE-ABS-KEY ("research quality")
(S) TITLE-ABS-KEY ("paper quality")
(S) TITLE-ABS-KEY ("scientific quality")
(S) TITLE-ABS-KEY ("academic paper" OR "academic article" AND guide)
(S) TITLE-ABS-KEY (manuscript AND guide)
(S) TITLE-ABS-KEY ("evaluating research")

-
- (G) evaluating research
 - (S) TITLE-ABS-KEY ("research review")
 - (S) TITLE-ABS-KEY ("write a scientific paper")
 - (S) TITLE-ABS-KEY ("literature review")
 - (S) TITLE-ABS-KEY ("literature review" AND guide)

A.3 Secondary articles

Searches for helpful articles, i.e., articles used to establish the context in the introduction, etc.

- (S) TITLE-ABS-KEY (number AND articles)
- (S) TITLE-ABS-KEY (number PRE/1 articles)
- (S) TITLE-ABS-KEY (journal AND frequency)
- (S) TITLE-ABS-KEY ("number of articles")
- (S) TITLE-ABS-KEY (bibliometric)
- (S) TITLE-ABS-KEY ("growth of science")
- (S) TITLE-ABS-KEY ("research question")
- (S) TITLE-ABS-KEY ("publish or perish")
- (S) TITLE-ABS-KEY ("publish or perish" AND student*)