# Using AI in performance feedback systems: How do ethics affect integration?

Author: Thijs ten Vergert
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands

## ABSTRACT,

***Purpose:*** AI feedback is increasingly used in new and existing businesses, using AI in feedback processes requires integration of such systems in an existing HR department. This paper seeks to find out the ethical dilemmas and considerations regarding AI in performance feedback, and how they affect the integration of AI in feedback systems.

***Design:*** Due to the subjective nature of ethical opinions, this paper uses a qualitative research design with semi-structured interviews. Interviews were conducted with one employee being monitored by AI software, two sales representatives of an AI software company and four experts in the academic field of ethics.

***Findings:*** From the interviews held, multiple interviewed parties suggested that biases will always exist within AI software. AI software can never be fully autonomous. An employee's perception towards privacy can be explained by many demographic factors. Privacy is strictly safeguarded by country laws. Right now, software developers have difficulties explaining how an AI software comes to a feedback decision, transparency is important for an AI to be ethically sound. Managers need to be accountable for the decisions an AI software makes.

***Conclusion:*** Ethical considerations affect the multi-functionality of an AI software in performance feedback the most, the accuracy and machine autonomy of an AI software are strongly inversely related to each other. No interviewed group believes that machine autonomy with the current state of AI intelligence can be achieved in performance feedback.

## Graduation Committee members:

Dr Anna Bos-Nehles

Dr Maarten Renkema

## Keywords

Performance Management, AI, Ethics, Integration, Stakeholder, Interviews

# 1. INTRODUCTION

## 1.1 Situation

Imagine a future scenario: you are a senior in a production company, many coworkers look up to you due to your knowledge in all parts of the company, and you boost team morale by encouraging new employees and showing them how to do their work best. This however comes at a cost, due to you often being stopped on the work floor to answer questions, your production rate is considered below average in computer data. Due to your company transitioning to a fully AI-managed performance system, you are being dealt with accordingly by being deemed unfit for the company's current working requirements. This results, in you being without a job. The system will consider a job well done given that the mean production numbers have gone, however, a traditional performance manager might have recognized that you as a person provide much more than just products. Such a future can occur when a company decides to implement a system that focuses solely on performance numbers. This possible future scenario paints AI in a bad picture, however, it can also have many advantages that are discussed later in this paper. way to make sure there is a positive future including AI software can be done by setting ethical boundaries through personal considerations about whether they should implement it. Through this paper, I want to explore how these ethical considerations or boundaries are affecting the integration of AI the performance feedback of businesses.

## 1.2 Objective

In recent times, more and more companies have started to adopt AI solutions in their employee feedback systems (Schrage et al., 2019). According to Johnson et al., (2020), AI systems can provide benefits in automating decision-making processes, enhancing business processes, and providing cognitive insights. Furthermore, AI systems have shown to drastically increase intake capacity(Parveen et al., 2019) and make it more streamlined for employees to see their performance. However, there is also a negative side to the implementation of these AI systems. A study from Tong et al. (2021) shows that employees who know they are receiving feedback from AI systems receive on average a 5.4% lower job performance. With AI software being primarily number based, processing large amounts of data (Jeske & Santuzzi, 2015), AI is very good at handling a large amount of quantitative data, but not necessarily qualitative data (Deloitte Insights, 2020). Concerns arise about ethical issues of AI in performance management (Alder, 1998). In the paper, Alder suggests that in the past, electronic performance management has caused privacy breaches in companies. Similarly, a paper by Tong et al., (2021b), suggests that employees feel that AI is too intrusive in their line of work, resulting in lower work satisfaction and negative perception toward AI. Interactions with humans can result in a negative feedback loop, resulting in AI giving biased outcomes. (Akter et al., 2021). An example of this is Amazon's secret AI recruitment tool. This tool has accidentally shown bias against women in the selection process for potential new employees (Julien Lauret, 2019) As seen above, there are existing papers that focus on ethics in (electronic) decision-making processes, (Alder, 1998a; Sillup & Klimberg, 2010; Tong et al., 2021b).

These papers focus on the advantages and disadvantages of implementing AI software. Less is known about how these ethical perspectives affect software developers in integrating AI software in specifically performance feedback, as well as how existing businesses feel about the implementation of AI in what ways they want to use AI to give employees feedback.

This paper seeks to find more clarity by first clarifying what the ethical perspectives of ethics, employees, and software developers regarding the usage of AI software in performance feedback are. Afterwards, I seek to explain through this paper how these ethical perspectives affect the integration of AI software in performance feedback.

## 1.3 Research question

To answer this objective, a research question has been created that discusses the current ethical perspectives from experts in the scientific field of AI-ethics, and businesses, while also trying to answer how integration of AI in performance feedback gets affected.

*"How is the integration of AI software in performance feedback affected by ethical considerations?"*

## 1.4 Academical relevance

This paper is relevant for multiple academic fields. These fields are AI software, business management, and performance feedback. To see how this paper will be relevant to each of these different fields, I will be discussing them here:

### *1.4.1 Ethics:*

This paper is relevant for the academic field of ethics as it brings new ethical perspectives on AI feedback in specific. A contribution of this paper is the ethical issue of businesses having a high probability of performing with a "black box" due to software developers not being able to explain how AI comes to feedback decisions.

### *1.4.2 Business Management:*

This paper is relevant for the academic field of business management as the results can show how businesses can introduce performance management using AI in a more ethical sound manner. This paper contributes to the field of business management as it can indicate where potential dangers are in implementing AI software for performance feedback. For instance, this paper shows that in the current state of AI, there is consensus amongst experts a business should not implement AI feedback fully autonomously.

### *1.4.3 Performance feedback:*

This paper is relevant for performance feedback research as it shows in what ways ethics can affect the effectiveness of AI in performance feedback. This paper provides new insights in the 3 specific areas of integration of AI generated performance feedback. These insights show that performance feedback in it'

## 1.5 Practical relevance

The practical relevance of this paper is to give insights on how to implement AI in their performance management in an ethically sound way, or if they should at all consider implementing AI in their business. This paper gives insights to business owners as well as software developers to see to what degree there is a need for more intelligent AI, how ethical topics affect how multi-functional an AI software can be and how interactive this AI can be. From this paper it becomes clear towards businesses looking into implementing AI software in their performance feedback systems that they must know why they want to implement AI before implementing it in their business. This paper furthermore shows that businesses must be careful regarding their security systems, as data suggests that widespread implementation of AI software is vulnerable to security breaches. This paper shows that business owners should not see AI software as a replacement for conventional

# 2. LITERATURE REVIEW

## 2.1 Performance management and feedback

To give context as to how technological advancements have been shaped by ethical approaches, it is important to look at how traditional performance management has been shaped in the past.

According to Osmani and Ramolli, (2012), traditional performance management is a collection of communication moments between managers and employees. Evaluations are done yearly or half-yearly through formal meetings. Osmani and Ramolli (2012) argue that there are four different types of performance assessments: assessment of features/attributes of personal character, assessment of behaviour, assessment of results, and self-assessment. They argue that the key performance indicators of performance management are productivity, quality of work, initiative, teamwork and problem-solving.

This performance feedback could be done previously was through sampling employees, (Komaki, 1986) where manager directly observes or inspects work. According to Jordan, (2013., p. 201) traditional performance feedback meetings are done through only line managers, traditional performance feedback outcomes are a salary increase, promotion or demotion, or development, and these events are done on a yearly basis.

## 2.2 AI in performance feedback

Nearly 80% of all businesses now use some sort of Electronic Performance System (Tomczak et al., 2018). Electronic Performance Systems use computers to collect, store, analyses, and report on information about employees' activities (Tomczak et al., 2018). Managers have quicker and easier access to a large amount of data from employees due to computers being able to process a large amount of data quickly.

In modern days AI performance systems are used. As there are many different interpretations of what AI is, this paper will use the description of AI by Haenlein et al., (2019) They state that AI is "A system's ability to correctly interpret external data, to learn from such data and to use those learnings to achieve specific goals and tasks through flexible adaptation" Regarding feedback systems this means that AI software can flexibly give advice to employees and feedback on the quality of their work, while adapting to the situation this employee might be in.

The reason this description is chosen is because it is most fitting towards performance feedback. Performance feedback AI is supposed to be very good at handling a large amount of quantitative data (Deloitte Insights, 2020), AI has the potential to: analyze employees based on predetermined parameters, identify employee characteristics to fit their personal needs, identify employees' preferred methods of working and remove human biases (Maity, 2019). Moreover, Malik et al., (2022) state that AI can give employees continuous feedback, resulting in a higher job involvement from employees.

The paper of Du and Xie (2021) suggests that three key indicators can explain the integration of AI in performance feedback, being interactivity, multi-functionality, and level of AI intelligence. These three indicators are important to performance feedback as performance feedback itself can be done through various means, as discussed in the section above. When talking about implementing AI in performance feedback, the three distinct indicators given by the paper of Du & Xie conceptualize the most important dimensions: the type of interaction and frequency of interaction an employee has with AI, the number of tasks an AI fulfills in giving performance feedback, and the degree of intelligence. The degree of intelligence is an important factor towards performance feedback as it determines the quantity or quality of the type of feedback an employee can receive.

**Table 1: Description of key indicators (Du & Xie, 2021)**

| | |
|---|---|
| Interactivity | Interactivity describes the quality and quality of consumer instructiveness. Interactivity has two dimensions, the nature (interaction interface, modality) of an interaction and the scale (quantity) of interaction. |
| Multi-functionality | Multi-functionality is the number of functions or tasks a product can perform. |
| Level of AI intelligence | With the level of AI intelligence, it is meant the scale to what degree the AI is intelligent, whether it is an AI that performs tasks below the capacity of human intelligence, or if it is an AI that does tasks that are too advanced for regular human intelligence to process. |

## 2.3 Ethics in feedback systems

We first need to understand how ethical opinions are established and what defines them as ethical opinions. According to Jalil et al. (2010), the core basis of ethics is what it means to be a good person and what to do good. This sense of right and wrong can also be described as morality (Velasquez, 2014). Velasquez argues that in a business setting, business ethics are important for long-term strategies, as they can improve long-term relationships with both external stakeholders and internal employees. Business ethics tries to solve this by applying morality to a business setting, and critically think what moral implications a business idea will have towards greater society. The paper by Wang (2020) states that the ethics of AI are the moral behaviours of artificially intelligent agents, well-implemented AI will result in ethical AI. Ethics is important according to this paper as AI can interfere with many human rights, including the right to human dignity, the right to privacy and the right to work and the right to an adequate standard of living, with bad ethical AI approaches, these rights will all be affected. All these considerations explain ethics as something related to the well-being of humans and are therefore key towards the sustainability of a company.

The paper of Sillup and Klimberg, (2010) argues that some of the key areas business ethics help within feedback systems are respecting individual employees, having a sense of mutual respect between managers and employees, ensuring fairness in performance assessment systems, and having transparency in the decision-making process of performance assessment. The paper of Alder (1998) states that there are ethical issues towards using performance feedback in the areas of fairness and privacy. As ethical considerations need to be categorized to be able to answer the research question, I went looking for a set of ethical principles to hit the most important areas of concern discussed the papers about performance feedback mentioned before. The paper of (Kieslich et al., 2021) states that there are seven guidelines businesses should adhere to regarding ethics in AI feedback, these guidelines being:

**Table 2: Description of the attributes (Kieslich et al., 2021)**

| Ethical principle | Description |
|---|---|
| Explainability | Explanation of the decision: People know why the machine does the things it does |

| | |
|---|---|
| Fairness | No systematic discrimination: No groups (ethnicity, race, gender, etc.) get discriminated by the system |
| Security | State-of-the-art security technology: The system is secured against hackers to keep data safe |
| Accountability | Full responsibility toward employees: The decisions made by AI need to be accounted for by top management towards their employees |
| Accuracy | Virtually no errors in decision-making: The software used in AI feedback should be without fault |
| Privacy | Exclusively earmarked use of data: The data of employees should be safe and secure so that the things employees do will not be leaked to the outside world |
| Machine autonomy | No human supervision: For a good working AI system, given that it is Artificial Intelligence, it needs to not have constant human supervision. |

## 2.4 Conceptual framework

To answer my research question, ideas need to be conceptualized to get an answer to the limit to which AI can be implemented ethically. As the paper of Du and Xie (2021) suggests, due to there being many kinds of businesses, there is not one core AI system that reflects all business types. Therefore, our conceptual framework must implement these dimensions to see how far the limit to AI involvement is: The paper suggests there are three factors crucial to the integration process of AI in businesses:

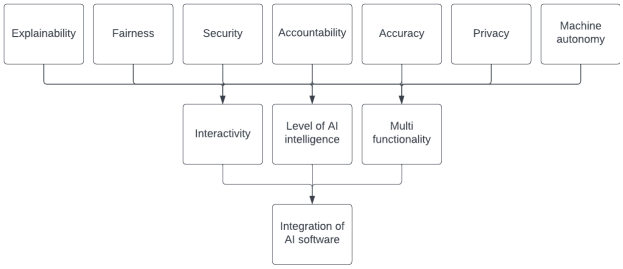
- Interactivity
- Multi-functionality
- Level of AI-intelligence

To explain how these three factors are affected in the integration process of AI in businesses, we use the description of attributes given by Kieslich et al. (2021). The attributes that will decide how high the interactivity of a software can be, how multi-functional a software can be and how intelligent the AI software is are the seven attributes listed below:

- Explainability
- Fairness
- Security
- Accountability
- Accuracy
- Privacy
- Machine autonomy

These two levels of concepts together form a conceptual framework as seen below, where each of the steps of this research is easily identifiable, and where the causal relationship of the parts of this research can be explained.

**Figure 1: Conceptual framework**




# 3. METHODOLOGY

## 3.1 Research design

Research committed by this paper is explorative as this paper tries to spark a discussion regarding ethics when it comes to AI-feedback systems. An explorative purpose follows to explore how feedback systems are affected.

The research will be a qualitative analysis based on interviews. This paper is centered around ethics means that this paper will include many opinions and personalized data, which means that there will be a need for open-ended questions (Sofaer, 1999). The data collected in this instance is qualitative, meaning that a qualitative analysis will follow.

## 3.2 Subject selection

This paper requires data from three categories of interviewees to get a substantial amount of data to get a proper result. These three subjects are a business employee, sales representatives of an AI company and individual unbiased experts on ethics in AI.

The reason these three different categories of subjects are chosen is that a sales representative represents a company to sells AI software to their customers, this means that a sales representative is responsible for voicing the company's norms and values. This means that a sales representative can give a good insight into the ethical perspectives of a company that sells AI software. An employee is chosen to get data on the employee perception of using AI software. This employee perception can be used to identify further ethical considerations taken in the decision-making process of this AI software. Unbiased experts in ethics are chosen to get further deeper opinion on ethical problems, these experts are furthermore used as a control group to

## 3.3 Criteria for subject selection

To be able to interview the selected subjects smoothly, there are key areas which every subject should comply with. These areas are chosen to be able to transcribe the data extracted from each subject. The characteristics chosen which a representative must have, are chosen to ensure that enough qualitative data can be achieved.

### *3.3.1 Sales representative*

The sales representative should be competent in speaking English. The sales representative should work for a company that is specialized in the creation of AI feedback systems and should have good expertise in what goes into the creation of said software. The software developer should be a senior within their company that has had enough time to know what their company does and to know which departments do what. The sales representative should also have a strong interest in implementing further AI software, while they also need to have their own ethical views on the software they are selling.

### *3.3.2 Employee*

The employee for our research should be competent in speaking English. The person should have experience being judged by AI software and have experience with regular feedback management as well. The employee should be active in their respective company for more than 1 year.

### *3.3.3 Expert in ethics*

The experts interviewed, as suggested in the name, should be knowledgeable about ethics. The experts should have a basic understanding of AI technology, and what some of the ethical concerns are towards AI. The experts must be fluent in either Dutch or English for fluent communication. The experts should be active in the field of ethics for at least a year and should have experience in writing papers about ethics, preferably in the domain of AI software.

## 3.4 Data collection

For this research, semi-structured interviews are chosen because semi-structured interviews fit best the criteria needed for the type of research committed. According to Dejonckheere and Vaughn (2019), semi-structured interviews are an effective method when qualitative open-ended data is needed, which explores participants' thoughts feelings and beliefs. Talking about ethics and ethical considerations involves personal opinions and explaining this choice.

As semi-structured interviews are the basis of this paper's research, questions are open-ended to spark a conversation with the interviewee. This way a large amount of data can be extracted without getting stuck on yes or no answers, and opinions and thoughts can be processed in our data. This interview will be recorded either through a textual matter or by asking consent beforehand to record it using either a microphone or video footage.

To get interviewees suitable for the data required to make this paper, I chose to use connections from my supervising professors to professionals in AI feedback system technologies, together with this, I contacted businesses related to AI feedback systems through platforms such as Linked-In. Furthermore, I contacted professors from the University of Twente.

**Table 3: Number of interviewees and their characteristics**

| | |
|---|---|
| *Representatives* | 2 subjects interviewed<br>Both subjects have extensive work experience with sales and development of AI software. |
| *Employee* | 1subject interviewed<br>Employee works in a business that sells AI performance feedback software.<br>Employee has been active in the company for more than a year. |
| *Experts* | 4 subjects interviewed<br>All subjects are researchers on the subject of ethics. Two of which are specialized in the field of AI software. |

## 3.5 Data analysis

The data gathered in the collection phase is transcribed to be easily applicable to our research. As the data used in this paper will be semi-structured, qualitative interviews, large amounts of data will be available to use during the data analysis phase, deductive coding is used to specify the specific themes talked about in the interviews. The data is coded by transcribing the interviews held into text, and by selecting the data most representative for each ethical consideration. For the findings, the data is first coded to contain the most important information expressed by the interviewees, afterwards each specific consideration is discussed with all the data sourced from the interviews.

**Table 4: Coding for interview data**

| **Ethical considerations regarding AI** |
|---|
| 1. Explainability |
| 2. Fairness |
| 3. Security |
| 4. Accountability |
| 5. Accuracy |
| 6. Privacy |
| 7. Machine autonomy |

# 4. RESULTS

In the results section, first, a summary of the findings from the interviews will be given using a table. Afterwards, this paper will go into more detail as to what these findings mean.

**Table 5: Coded interview data**

| **Ethical considerations regarding AI** |
|---|
| 1. **Explainability:** "Software developers themselves often cannot explain the reason behind an AI software's decision making." |
| 2. **Fairness:** "Avoiding bias has to be an intentional goal that can be minimized but never fully eliminated" |
| 3. **Security:** Security within AI is location-based, and a company responds to these locations differently. There is risk involved when a business has multiple different facets. Secure cloud providers provide suitable security measures. |
| 4. **Accountability:** Depending on the business, top management do find themselves accountable to explain to an employee how an AI came to conclusions. |
| 5. **Accuracy:** A system cannot pick up nonverbal cues, making it possibly inaccurate. Accuracy can be guaranteed, but the interpretation of the projected information will always need to be supervised by a person. |
| 6. **Privacy:** "Anybody who's supervised people know that some people want really direct feedback. Like tell me start doing this, stop doing that. Other people you kind of have to sugarcoat it because you're going to make them cry." |
| 7. **Machine autonomy:** "For AI to give fully autonomous feedback, a miracle must happen" |

## 4.1 Ethical considerations regarding AI:

### *4.1.1 Explainability*

In the first part of the interview, it was suggested by the sales representative, that when they must explain what an AI can do for a possible client, rather than the client knowing what they want to do with an AI, the client should not be interested in the AI software that is being sold. This suggests that the sales representative is implying that an employer should have a key understanding of the AI and its usage, and explainability cannot be achieved if an employer cannot understand the software themselves.

The experts explain that it is important for business owners to be able to explain AI decision-making in a performance feedback context as ethics deal with human emotions. They argue when a person gets performance feedback, the AI does not have freedom of will, the AI does not reason, and the AI computes. Therefore, the feedback given by said AI needs to be explained by a human. The experts indicate that reasons for a certain feedback decision should be provided, in a way it is transparent enough for the user to understand how AI came to a feedback decision, is so that the user can select the data that is relevant for them to perform better.

One of the sales representatives stated that some of the people developing the AI have a difficult time explaining how the technology came to a certain decision. These developers frequently hide behind the assertion that they cannot explain the algorithm or process due to it being a trade secret. This rhetoric results in little explanation to business owners and ultimately the employees.

The experts explain that the reason why software developers might have difficulties with explaining how an AI comes to feedback decisions is because there is a distinct difference between the mechanics that are used for an AI to conclude, and the actual process an AI makes to reach a certain conclusion. They say many software developers can explain the mechanics that are used as these are concepts developed by software companies themselves. The experts state that the key area where developers are currently struggling to explain their software is the explanation path. This results in a black box situation as illustrated by the sales representatives. There is a consensus amongst ethics experts that such a black box is problematic, as it will hurt the transparency and explainability in the end.

### *4.1.2 Fairness*

The ethics experts state that AI-generated feedback is valuable in the way it can provide an employee with an objective view on their performance, and that AI software can give more in-depth insights on specific performance indicators of employees. However, as AI is still produced by software developers, the experts say there is usually a bias, which is an ethical concern of AI software. The experts say this bias comes from humans writing the AI software in the first place.

Similarly, the sales representatives explain that biased software cannot be avoided as a person developing AI software has unconscious biases themselves. They say this can be reduced to a minimum as according to the sales representatives, many software developers employ specific teams that tackle the possibility of software having biased outcomes. This way, the amount of bias in software can be reduced to a minimum, however, the sales representative believes that it can never be fully eliminated.

The experts explain that bias is an ethical concern as it may result in unfair and systematic discrimination. They argue that biased software can result in AI growing habits of discrimination towards a certain gender, age group, racial group, etc. This is an ethical concern they

In the interview, the sales representatives mentioned that AI software should be used for coaching and to make people better, rather than to use it for compliance, control, and constraint. The sales representative referred to a scenario in sports, where an AI can either act like a referee or a coach. Quote: *"Does anybody like referees on the planet, right? I don't know. I don't think so. And they're constantly blowing their whistle and calling a foul and saying stop doing that. You can't do that. I'm gonna punish you, put you in the penalty box.*" The sales representative explains that according to him, most people do not like the concept that is a referee. They are seen as someone that judges their every move and something that can punish an employee on something an employee can deem unfair. People, however, often enjoy the possibility of getting coached, as it will make them a better person at their job. It is implied that employees are more likely to perceive coaching as something less likely to be unfair, as coaching is meant to build upon an employee's strengths.

Another ethical concern regarding fairness according to the experts is the possibility of employees seeing AI software as something more than software. They argue that AI tricking employees into being more than software can leave an employee vulnerable to being cheated out of human-on-human interactions.

### *4.1.3 Security*

According to the sales representatives, there will always be risks involved with using AI software. When a business has multiple departments, the software needs to be sent from one department to the next, as this is data that is vulnerable to being exposed when a leak gets found, businesses that have multiple departments will always be prone to these weaknesses. Larger businesses with larger departments that span over multiple different countries are more likely to be vulnerable as this data is moved through the cloud.

According to the sales representatives, software developers could provide their own data centers, but due to cost restraints and economies of scale, they explain that it is not feasible for every software provider or business to do so. Therefore, most companies use existing large-scale and secure cloud providers.

### *4.1.4 Accountability*

According to the employee, in their specific business, top management does explain to employees how the AI comes to certain conclusions and on what they can improve according to the AI. This means that there is an accountability of top management towards the employees. Due to the sample size of this question, 1 person, it cannot be said that every single business does this. The employee states that they do formal meetings with top management to observe the data collected by AI evaluating them. Here, top management takes accountability of the decisions made by the AI, by doing these meetings.

The experts say that AI should reach an understanding between employees and managers. The more responsible a manager is with the decisions and recommendations given by AI feedback, the better an employee's perception of said manager will be. To be responsible is where an ethical issue lies as it requires sufficient control of AI software by a manager, insufficient control will lead to a lack of accountability.

### *4.1.5 Accuracy*

The sales representatives stated that in a future scenario, accuracy can be guaranteed by AI software. They also stated? that in the current state of art, this is not possible yet. They therefore suggested that accuracy can only be guaranteed with human involvement. The reason why accuracy is so important, according to the experts, is because inaccuracy is harmful to an employees' wellbeing.

It is suggested that the level of accuracy is dependent on the type of job environment. The sales representatives stated that in a job environment like manufacturing, it is simple to measure someone's performance. As it is a linear process, AI software is less dependent on variables to accurately determine an employee's performance. The sales representative refers to these jobs as blue collar jobs (manual labor): *"It's the white collar and the no collar roles that are going to be truly interesting to see how those evolved over time and how people collect data and use that to define whether or not somebody's performing well."* With this the software developer implies that right now, white collar (office work) and no-collar jobs (creative work) are harder to accurately determine the performance of an employee.

### *4.1.6 Privacy*

From the interview with the employee, it became clear that depending on the location of a business, the business may respond differently towards the amount of data used from their employees that can be collected. The employee's business, being housed in Germany, meant that according to the employee, there were more laws that regulate the flow of data being collected. This is further backed by the sales representatives, who stated that their specific company had much stricter privacy laws to coincide to, compared with other countries. They stated: *"Germany or France, the unions, the Labor Council, they are really, really strict over there"*

The sales representatives furthermore stated a second important factor that affects a persons view towards privacy. This factor being the type of employee that is being judged by AI software. Quote: *"Anybody who's supervised people know that some*

*people want really direct feedback. Like tell me start doing this, stop doing that. Other people you kind of have to sugarcoat it because you're going to make them cry."* There are multiple reasons to why a person might behave differently towards AI. The experts on ethics believe the most important factor towards an employee's perception towards privacy is gender, they argue that male employees are more competitive on at work and are more eager to get their performance non-stop appraised by an AI software.

Another ethical concern regarding privacy according to the experts is the awareness of employees towards them being monitored. When an employee is being monitored without their knowledge it indicates a lack of trust from the manager's side. According to the experts, trust-relationships are crucial towards the privacy of an employee within a company, and without such relationships an employee can experience paranoia on the workfloor. The experts mention this phenomenon is also called the *chilling effect.* I

### *4.1.7 Machine autonomy*

Similarly discussed in the accuracy section, the sales representatives stated that accuracy cannot be guaranteed by AI software. This means that to have an accurate software, no full machine autonomy is possible in the current state of art.

The level of machine autonomy can be partially decided by the job environment of a business, as similarly discussed in the accuracy section. To define different job environments, the sales representative uses the example of collared jobs, he states that blue collar jobs are easier to account for as they are more linear jobs. He states that white collared and no collared jobs will be more and more interesting in the future in the way it determines if someone is doing well or not. An ethical issue of machine autonomy according to the experts is that these job environments shape a certain standard an AI software thinks an employee is supposed to perform like. They argue that a job environment can disregard an employee's opportunity to present themselves the way they want.

This argument of white-collared and no collared job being harder in the current state of art to be fully autonomous is further explained by the employee. They state that the way they get feedback through AI management is by using a self-service to request data. All this data gets send to her managers, who still do an in-person review of her performance bi-annually.

## 5. DISCUSSION

**Table 6: Most important findings**

| | Interactivity | Multi-functionality | Level of AI intelligence |
|---|---|---|---|
| Explainability | Unexplainable software can lead to undesirable interactions. | "Black box" forming is more likely when a software tries to be highly multi-functional | To achieve explainability, the amount of intelligence an AI can has would be limited. |
| Fairness | AI cannot cheat a person | Multi-functionality suffers when the AI is made to be a coaching software rather than a controlling one | Highly intelligent software is more likely to be biased. |
| Security | Higher scale of interactions results in more vulnerability | Security will be at risk if a software is made to be functional across many departments. | |
| Accountability | | The more functions an AI software has, the harder it is for an owner to take ethical accountability for it. | |
| Accuracy | | Accuracy will be lower when a software is highly multi-functional. | To get high accuracy, a software cannot be too complex |
| Privacy | The scale of interactions of a software is affected by privacy laws for multi-national companies. | | |
| Machine autonomy | For most companies, machine autonomy will require supervision which affects the scale of interactivity. | When the AI cannot be autonomous, the multi-functionality suffers from it. | A more intelligent AI software can achieve a state closer to autonomy. |

## 5.1 Effects on interactivity

Interactivity is affected on multiple of the ethical areas discussed in the previous section. The data implies that interactivity is affected by the explainability of AI software in the nature of the interactions. The data suggests that many developers of AI do not know how an AI comes to a certain conclusion. This poses a danger for employees as the interaction employees have with AI might not be able to be accounted for.

Regarding fairness, the data implies that interactivity gets affected by ethical concerns in the area where it cannot cheat a person into thinking they are talking to an actual human. This affects the interactivity in the nature of the interaction. For AI to be implemented ethically, this would mean that the means of giving performance feedback would be done on a less interactive basis. This supports the literature of Tong et al., (2021) which stated that employee satisfaction was lower when using AI software to give performance feedback.

As discussed in the security section, when there is a higher scale of interactions between departments, the security of said AI software becomes more vulnerable. This would influence interactivity in the scale of interactions that could be done while giving feedback.

From the privacy section it becomes clear that privacy has a big effect on the interactivity of an AI. For a multi-national company, the scale of interactions will be limited, as some countries have privacy laws that do not approve of AI interaction. According to Du & Xie, (2021), AI is interactive when it is contingent, synchronous, participative, modality-rich, and anthropomorphic.

By reducing the level of participation in hopes to safeguard employee privacy, interactivity is affected.

The data suggests that there is currently no way to get a fully autonomous AI working for a business, this would affect level of interactivity an AI can have with a person. The nature of interactions in this case will be altered as there will be human supervision required to ensure that the decisions made by AI are according to the wishes of its business owners. Effects on multi-functionality

If a business wants to ethically implement their AI software, the explainability will influence multi-functionality. As a multi-functional software is more complex, with more functions to consider, the software will be more difficult to understand which can create a higher probability of a "black box" forming. This will result in software developers and consequently managers unable to explain how the AI came to a feedback decision. If a business is incapable of doing so, the explainability of a highly multi-functional feedback system cannot be reached.

According to the sales representatives, the way a software can be used in a fair and ethical way is by making the software a coaching software rather than a refereeing one. With coaching software sales representatives referred to software that gives employees areas they can work on and software that will automatically assign employees a course which can help them achieve a better output. This impacts the multi-functionality of a software. When a software can only be primarily used as a coaching tool, the options of functions that can be implemented are reduced.

As discussed in the security section, the more different departments an AI software has, the more sales representatives of a software provider believe it hinders the security of a company. From this information, it can be argued that this will have an impact on the multi-functionality of a software. When a business wants to employ a software with a high level of multi-functionality, according to software developers, they are at risk of making the data being handled less secure.

This category coincides with the explainability of AI software, as when an owner cannot explain the decisions an AI makes, they cannot account for those decisions towards the employees within their business. This will negatively affect the opportunities in functions an AI can have in their business, were they to follow an ethical implementation.

As discussed in the accuracy section, different job environments require different types of AI integration. From the data given, it is implied that a software will be less accurate when it is more complex. This means that a software with a higher multi-functionality, is more likely to be less accurate. This implies that to get a high accuracy, AI can only have a limited functionality.

Depending on the functions a multi-functional AI wants to bring, privacy concerns can pose to be a challenge. As the data suggests, privacy is something personal and a governmental regulated human right. A business that spans borders requires highly diversified AI to account for the local differences between people. Depending on the laws within different countries, this can mean that privacy rules limit the options a business can have within that specific area.

From the data it is shown that currently there is no fully autonomous AI possible within businesses. This affects the multi-functionality of an AI as no autonomy means that there must be someone to supervise the answers given by AI software, this limits the functionality an AI software can have.

## 5.2 Effects on level of AI-intelligence

It is stated that many AI-developers do not know how their AI came to a certain conclusion. It is also stated that the owner of a business should have a fair understanding of what they want to achieve with the AI software. To implement AI in a way where there is a high explainability, this would suggest that an AI cannot be too complex, as business owners will not be able to explain and understand what an AI software is doing. This means that when a high explainability wants to be achieved, the level of AI-intelligence is limited.

As the software developers suggested, AI is prone to be biased in some way. According to the experts, highly intelligent AI that uses feedback loops to come to a feedback decision are more likely to be biased.

As discussed in the accuracy section, to get a more accurate AI in the current state of art, it depends on the type of job this AI is set. If a high accuracy wants to be achieved, the business should reduce the level of intelligence, and the expectations the business can have on the AI being used.

Depending on the interpretation of intelligence, the experts say that AI will never have intelligence since they believe AI does not have inherent intelligence but has intelligent coding to run its systems. Other experts were of opinion that an intelligent algorithm in fact can lead to a higher level of machine autonomy.

# 6. CONCLUSION

We draw the following conclusions based on the interviews and analysis. The multi-functionality of an AI software is most affected by ethical principles, with the level of AI intelligence being the least affected by ethical considerations.

The next conclusion is that accuracy and machine autonomy are strongly related to each other in the way they affect the implementation of an AI software. The data showed that these two ethical principles inversely affect each other.

A third conclusion is that under no circumstance, any of the interviewed groups believe that AI in the current state of art can give feedback fully autonomously. While some experts were optimistic of the future, many suggested that a miracle would need to happen for AI to be implemented fully autonomously in performance feedback.

# 7. LIMITATIONS

This paper contains several limitations. First, due to the niche category of people available to talk about AI being implemented in their businesses, and the limited number of responses towards interview inquiries, only a limited number of participants were able to be sourced. This resulted in for example the employee giving little qualitative data. Without the option to look at more employee data, this resulted in a limited amount of qualitative data being used from employees.

Another limitation is the fact that people interviewed were people that are employees from businesses that sell AI software for performance management. These employees have a certain standard to uphold for the company they work for, therefore these employees might give biased opinions that paint their work or the software they work with in a brighter picture than it is. These biased opinions can affect the questions regarding their opinions on AI software and the effectiveness of their implementation.

# 8. ACKNOWLEDGMENTS

I want to thank my professor dr. Anna Bos-Nehles and dr. Maarten Renkema for the guidance during the process of writing this paper, for arranging interviews, checking work, and arranging the procedures for graduation. I want to thank both

Nikita Kononov and Hanna Riecken for helping with arranging interviews, for helping with formulating interview questions and for checking intermediate work. I want to thank the interviewees for their answers and time, without their input this research could not have been possible.

## 9. REFERENCES

Akter, S., McCarthy, G., Sajib, S., Michael, K., Dwivedi, Y. K., D'Ambra, J., & Shen, K. N. (2021). Algorithmic bias in data-driven innovation in the age of AI. *International Journal of Information Management*, *60*, 102387. https://doi.org/10.1016/J.IJINFOMGT.2021.102387

Alder, G. S. (1998a). Ethical Issues in Electronic Performance Monitoring: A Consideration of Deontological and Teleological Perspectives. *Journal of Business Ethics 1998 17:7*, *17*(7), 729–743. https://doi.org/10.1023/A:1005776615072

Alder, G. S. (1998b). Ethical Issues in Electronic Performance Monitoring: A Consideration of Deontological and Teleological Perspectives. *Journal of Business Ethics 1998 17:7*, *17*(7), 729–743. https://doi.org/10.1023/A:1005776615072

Du, S., & Xie, C. (2021). Paradoxes of artificial intelligence in consumer markets: Ethical challenges and opportunities. *Journal of Business Research*, *129*, 961–974. https://doi.org/10.1016/J.JBUSRES.2020.08.024

Haenlein, M., Kaplan, A., Tan, C. W., & Zhang, P. (2019). Artificial intelligence (AI) and management analytics. *Https://Doi-Org.Ezproxy2.Utwente.Nl/10.1080/23270012.2019.1699876*, *6*(4), 341–343. https://doi.org/10.1080/23270012.2019.1699876

Jalil, M. A., Azam, F., & Rahman, M. K. (2010). Implementation Mechanism of Ethics in Business Organizations. *Www.Ccsenet.Org/Ibr International Business Research*, *3*(4). www.ccsenet.org/ibr

Jeske, D., & Santuzzi, A. M. (2015). Monitoring what and how: psychological implications of electronic performance monitoring. *New Technology, Work and Employment*, *30*(1), 62–78. https://doi.org/10.1111/NTWE.12039

Johnson, R. D., Stone, D. L., & Lukaszewski, K. M. (2020). The benefits of eHRM and AI for talent acquisition. *Journal of Tourism Futures*, *7*(1), 40–52. https://doi.org/10.1108/JTF-02-2020-0013

Jordan, J. (n.d.). *Traditional performance appraisal versus 360-degree feedback.*

Julien Lauret. (2019). *Amazon's sexist AI recruiting tool: how did it go so wrong? | by Julien Lauret | Becoming Human: Artificial Intelligence Magazine*. https://becominghuman.ai/amazons-sexist-ai-recruiting-tool-how-did-it-go-so-wrong-e3d14816d98e

Kieslich, K., Keller, B., & Starke, C. (n.d.). *Artificial intelligence ethics by design. Evaluating public perception on the importance of ethical design principles of artificial intelligence*. https://doi.org/10.1177/20539517221092956

Komaki, J. L. (1986). Toward Effective Supervision. An Operant Analysis and Comparison of Managers at Work. *Journal of Applied Psychology*, *71*(2), 270–279. https://doi.org/10.1037/0021-9010.71.2.270

*Machine learning and qualitative data analysis | Deloitte Insights*. (n.d.). Retrieved May 18, 2022, from https://www2.deloitte.com/us/en/insights/focus/technology-and-the-future-of-work/machine-learning-qualitative-data.html

Maity, S. (2019). Identifying opportunities for artificial intelligence in the evolution of training and development practices Evolution of training and development practices. *Journal of Management Development*, *38*(8), 651–663. https://doi.org/10.1108/JMD-03-2019-0069

Malik, A., Thevisuthan, P., & de Sliva, T. (2022). *Artificial Intelligence, Employee Engagement, Experience, and HRM*. 171–184. https://doi.org/10.1007/978-3-030-90955-0_16

Osmani, F., & (Ramolli), G. M. (2012). Performance Management, Its Assessment and Importance. *Procedia - Social and Behavioral Sciences*, *41*, 434–441. https://doi.org/10.1016/J.SBSPRO.2012.04.052

Parveen, N. A., Palaniammal, V. S., Of 3 Rd Author, N., & Phil Management Scholar, M. (2019). A STUDY ON ARTIFICIAL INTELLIGENCE IN HUMAN RESOURCE MANAGEMET TODAY AND TOMORROW. *International Journal of Research and Analytical Reviews*. www.ijrar.org

Schrage, M., Kiron, D., Hancock, B., & Breschi, R. (2019). *Performance Management's Digital Shift RESEARCH REPORT In collaboration with*. https://sloanreview.mit.edu/big-ideas/performance-management

Sillup, G. P., & Klimberg, R. (2010). Assessing the ethics of implementing performance appraisal systems. *Journal of Management Development*, *29*(1), 38–55. https://doi.org/10.1108/02621711011009063/FULL/PDF

Sofaer, S. (1999). Qualitative methods: what are they and why use them? *Health Services Research*, *34*(5 Pt 2), 1101. /pmc/articles/PMC1089055/?report=abstract

Tomczak, D. L., Lanzo, L. A., & Aguinis, H. (2018). Evidence-based recommendations for employee performance monitoring. *Business Horizons*, *61*(2), 251–259. https://doi.org/10.1016/j.bushor.2017.11.006

Tong, S., Jia, N., Luo, X., & Fang, Z. (2021a). The Janus face of artificial intelligence feedback: Deployment versus disclosure effects on employee performance. *Strategic Management Journal*, *42*(9), 1600–1631. https://doi.org/10.1002/smj.3322

Velasquez, M. G. (2014). *Business Ethics Concepts and Cases Manuel G. Velasquez Seventh Edition*. www.pearsoned.co.uk

Wang, K. (2020). Artificial Intelligence (AI) Ethics: Ethics of AI and Ethical AI. *Journal of Database Management*, *31*(2), 74–87. https://doi.org/10.4018/JDM.2020040105

## 10. APPENDIX

**Interview Guide (Businesses)**

**General questions:**

- In short: how does AI work in performance management?
- Which data is being collected within the performance management process? How is this data collected?
- What happens with this data after collection?
- What is the overall goal of AI-generated performance feedback? What do you want organizations to achieve with it?

**Questions regarding the implementation:**

- How are line managers reacting on new AI performance software? Is the integration successful?
- What are issues that line managers and employees experience with using this software?
- How would you rate line managers capability to realize the full potential of the software?

**Questions regarding ethics:**

- How important do you think it is for employers to be able to explain the AI's decision towards their employees?
- Do you think AI can give fully accurate results every time when judging someone's performance?
- How can a business achieve no bias in their AI software?
- Do you think a future with fully autonomous AI can be achieved?
- Will government intervention be needed further in the future to safeguard employees privacy?
- Do you believe AI can be implemented ethically giving more qualitative results, or will a middle man always be necessary?

**Questions regarding integration:**

*Talk about paper of Du & Xie*

3 key factors: interactivity, multi-functionality, level of AI intelligence.

- What do you believe will affect the multi-functionality of AI from the matters discussed earlier? (privacy, fairness, machine autonomy, security, explainability)
- Can a highly intelligent AI behave on a higher level autonomy than a lower intelligence, or is it similar?
- Do you think a high level of interactivity be achieved if a company places high value on privacy?

## Interview Guide (Experts)

**Introduction interview:**

**I am currently writing my bachelor thesis in International Business Administration on how ethical considerations affect the implementation of AI-based performance feedback. I already gathered some data from an interview with an employee and 2 sales representatives of a software development company. An interview with you will provide me with valuable data from unbiased experts in the field, whose opinions and thoughts can greatly benefit the contributions of my thesis.**

**What is meant with performance feedback in this scenario?**

**This thesis addresses performance feedback and feedback management from a HR perspective. Performance feedback is the feedback that an employee receives about his/her performance. Based on performance indicators, customized feedback can be provided to employees about their current performance. Due to technological developments and Artificial Intelligence (AI), more data can be collected (content) and the provision of feedback can me automated (process). AI-generated feedback can provide opportunities to give continuous and on-demand feedback to employees. This is why more and more businesses are implementing AI in their performance feedback systems. I want to know in what way to implement AI-generated performance feedback in an ethical way.**

**If you agree, I would like to record the interview to transcribe the interview. The (audio) documents will solely be used for research purposes and your name and function will never be mentioned in communications about the research (data anonymity). If you agree I will start the recording now and ask you a couple of questions . Afterwards you also get the chance to ask a question.**

**Questions regarding AI-generated feedback in general:**

- **-We know from software providers that artificial intelligence will be used to provide employees performance feedback. What do you think about that?**
- **What are in your opinion ethical issues with AI-generated performance feedback?**
- **How valuable is AI-generated performance feedback in your opinion?**
- **What do organizations (software providers/employees) need to effectively implement AI-generated performance feedback?**

**Questions regarding the 7 ethical principles used in my paper:**

- UNDER WHICH CONDITIONS COULD AI PROVIDE EFFECTIVE PERFORMANCE FEEDBACK? WHY?
- DO YOU BELIEVE THAT SOFTWARE DEVELOPERS ARE NOT ABLE TO EXPLAIN THE DECISION MAKING PROCESS OF THEIR OWN AI SOFTWARE?
- IN WHAT WAY COULD THE DECISIONS/FEEDBACK BE EXPLAINED IN AN ETHICAL WAY?
- IN WHAT WAY CAN AI EFFECTIVELY EXPLAIN THE DECISIONS MADE AND FEEDBACK PROVIDED?
- HOW IMPORTANT IS IT FOR BUSINESS OWNERS TO BE ACCOUNTABLE FOR THE FEEDBACK GIVEN BY AI SOFTWARE?
- TO WHICH EXTENT CAN AI GIVE EMPLOYEES FEEDBACK WITHOUT INVADING A PERSONS PRIVACY?
- WHAT IS THE MOST IMPORTANT INDICATOR TO EXPLAIN EMPLOYEE'S BEHAVIOR TOWARDS AI GENERATED FEEDBACK? (DEMOGRAPHICS)
- WHAT HAS TO HAPPEN FOR AI TO GIVE FEEDBACK FULLY AUTONOMOUSLY?

**Questions regarding the implementation of AI software:**

**In what way is the intelligence of AI software related to the level of autonomy an AI can have?**

**To what degree should an AI be multi-functional enough where it facilitates a manager, instead of replacing the manager?**

**Should AI feedback be interactive where it replaces human feedback? Is this even possible in the current state-of-art?**

**Do you have any questions for me?**