The Meaning in Hiring

The potential loss of self-representation in AI hiring video interview systems

Master thesis
Derek van der Ploeg

Author
Derek van der Ploeg

Master thesis
September 2021
Msc. Strategic Product Design
Faculty of Industrial Design Engineering
Delft University of Technology

Graduation committee
Chair | **Dr. ir. L.W.L. Simonse**
Innovation Strategy - Faculty of Industrial Design Engineering

Mentor | **Ir. N. Stoimenova**
Designing for AI - Faculty of Industrial Design Engineering

Company Mentor | **Dhr. Q.V.I. Dalh**
Cofounder - HelloMentor

The Meaning in Hiring

The potential loss of self-representation in AI hiring video interview systems

Master thesis
Derek van der Ploeg

Preface

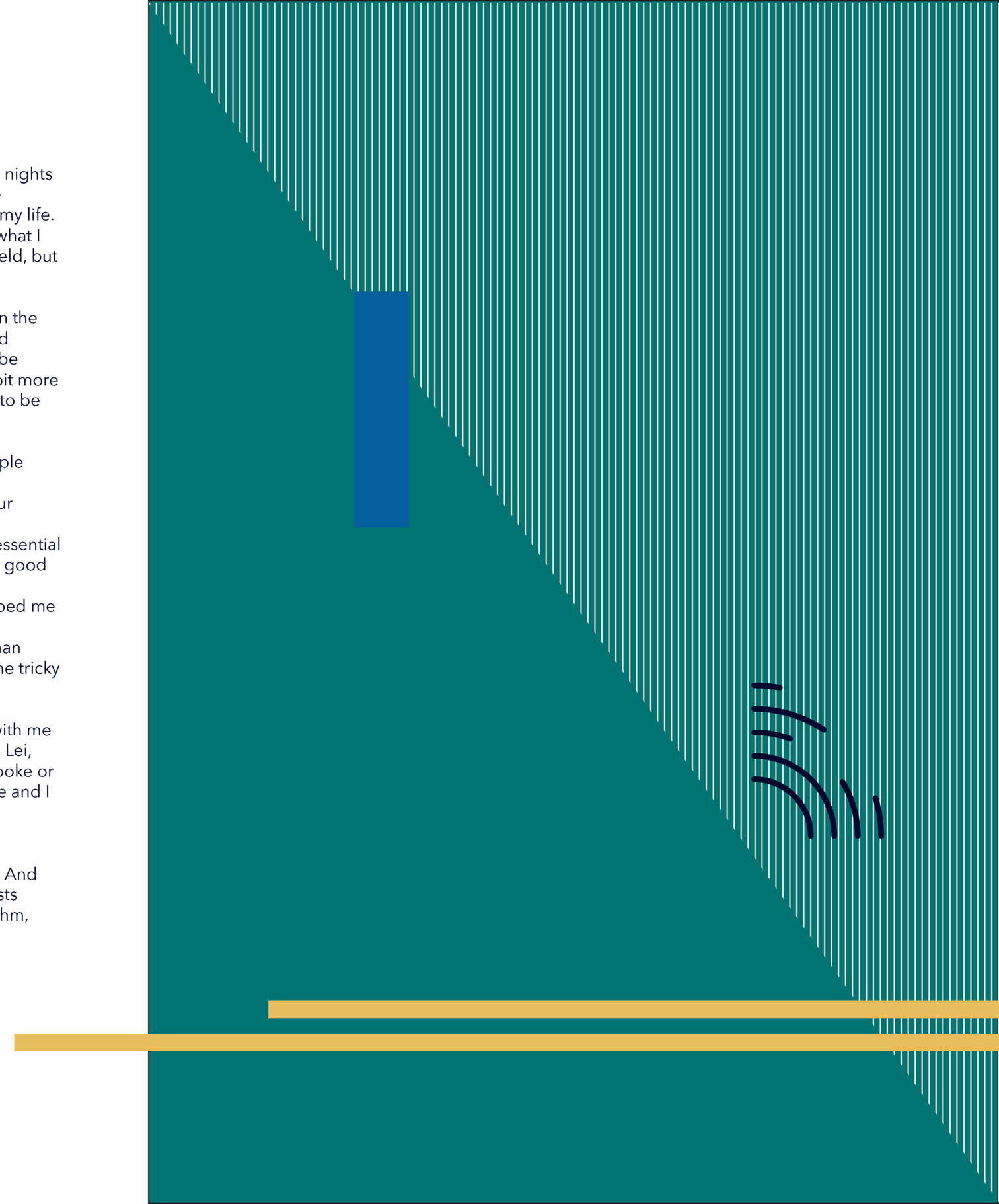
This project is the culmination of over a half a year of hard work and a lot of late nights at the library. While I cannot deny that at times I wanted to be just done with the project, this thesis has also been one of the most fulfilling things I have done in my life. Before I started my official design education in 2013 in Eindhoven, I described what I wanted to do “I just want to solve problems.” I did not care about one specific field, but about helping people live a better life, whatever that may entail.

I wanted to work on AI ethics because it is a neglected field in my opinion. Given the future importance of AI systems and potential of design to bring technology and society together, I figured it would be good to focus on how AI systems should be used in society. I truly hope that with this thesis I have been able to give a little bit more insight in how these AI systems should fit in a world where we expect everyone to be treated equally and fairly.

I would not have been able to write this thesis without the support of many people around me. I want to thank my supervisory team.
Lianne, thank you for your enlightening discussions, your moral support and your insight in what strategic design truly means.
Niya, thank you for your often razor-sharp feedback, your own work which was essential for understanding what I have found and your willingness to push me to deliver good work.
Quincy, thank you for keeping me grounded. Your insights and perspective helped me focus on something that could have an impact in the real world.
Evgeni, thank you for your inspiration through the original paper on AI and human rights, our deep and entertaining discussions and your help in understanding the tricky background behind values and technology.

I want to thank my family and friends who provided moral support and put up with me whenever I went full-on hermit. Hannah, Roberto, Agnieszka, Mitchell, Bart, Tara, Lei, thank you all (in no particular order). Additionally, thank you to everyone who spoke or worked with me during this project, your insights were essential for the outcome and I would not have come far without your help.

Finally, everyone who knows me knows that I’m a massive music fan. I estimate I have listened to 40.000 to 50.000 minutes of music in the creation of this thesis. And while variety is truly the spice of life when it comes to music, there are a few artists in particular I want to name: Bicep, Ben Böhmer, Tycho, The Glitch Mob, Nils Frahm, CHVRCHES and edIT, thank you for the good vibes.



Executive summary

Artificial Intelligence (AI) has permeated every part of our world. It discovers new molecules, recommends what to watch and informs many business decisions everyday. There it has also become part of hiring, where recruiters are in need of more efficiency. The digital age has caused the amount of applications per position to skyrocket, while organisations have noticed that in the last decades an increasing amount of their value generation is directly connected to their people. Vendors market AI systems as intelligent workers that can help human resources (HR) departments find the best people efficiently. Different systems help with analytics, writing, assessments or video interviews, where participants answers questions from an AI system on their computer.

With the implementation of AI systems, there are often ethical problems involved. Biases are often hidden in the data or algorithm, that may cause people to be unfairly treated by an AI system. But AI interviews have another ethical problem that has gotten relatively little attention: autonomy over self-representation.

During an interview, you always want to show your best side and focus on your strengths and best experiences. You know better than anyone else what you can do and therefore it is important that you are able to represent yourself. But AI interviews interfere with that self-representation because they make assumptions on what you mean before they present that information to a recruiter. Also, by nature AI systems can only work directly with quantitive data, so how are you sure that your meaning of ‘teamwork’ was properly processed?

This project uses a novel approach of Value Sensitive Design in combination with a different framework for generative prototypes to find a solution to this problem. Generative prototypes focus on generating hypotheses to further understanding. Here they were used with provocation in multiple iterations to elicit the values that people have about self-representation in hiring.

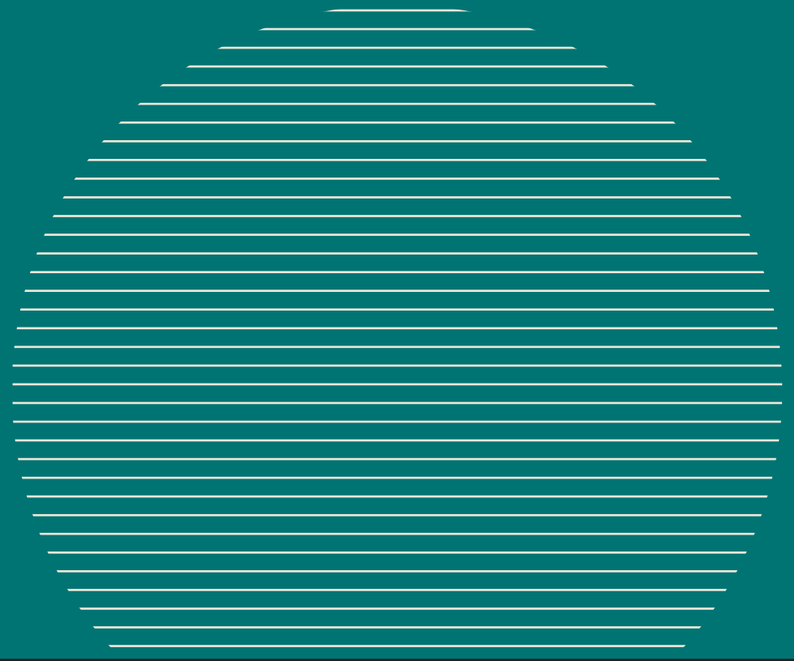
Those findings were synthesised into a new process that helps applicants maintain autonomy over self-representation through conveying feedback so they understand how well they are doing, through steering the interpretation of their answers and by keeping regular interviews to ensure the right nuance still arrives at the recruiter. This process was evaluated with another generative prototype, which informed the final three design requirements for AI interview systems in hiring: integrating feedback into the interview, managing expectations and assumptions and building in options for escalation.

In this way, the first steps are made for designing better AI systems that respect the autonomy over self-representation of applicants.

Table of Contents

Introduction

In this chapter, I introduce the context of AI interview systems in hiring and the problems that society is currently encountering in this space, namely doubts about the bias of AI systems but also the relatively under-recognised topic of self-representation and how AI systems interact with that value. Finally, I describe the different primary stakeholders for the project and what goals of the project are.



When Amazon built their AI hiring system in 2014, they intended it to become the primary way to do their recruiting. A person familiar with the effort phrased it as follows: “They literally wanted it to be an engine where I’m going to give you 100 resumes, it will spit out the top five, and we’ll hire those” (Dastin, 2018). However, already by 2015, the team noted that the system was systematically excluding women from jobs as it focused more on terms that men used more on their CV like “executed” or “captured” instead of focusing on the skills listed on the CV. These diversity problems plus general performance issues proved so persistent that Amazon disbanded the entire team in 2017 (Dastin, 2018).

Amazon’s story highlights one of the ways that AI systems in hiring can go wrong. Still, many organisations are interested in AI systems for hiring, because the value of automating hiring tasks is so great (Upadhyay & Khandelwal, 2018). Employers are willing to spend significant resources to get to a recruitment process that is more efficient or effective at finding the best people. Amazon is large enough to build their own system, but most organisations look to other companies for helping their human resource (HR) departments to automate parts of the hiring process through AI systems (Li et al., 2021; Reilly, 2019). In this thesis, these companies are referred to as ‘HR tool makers’.

The AI system at Amazon failed in ensuring diversity, but there are other ways that these systems can go wrong. For example, HR tool makers like HireVue, PredictiveHire and Retorio do automated assessments based on personality profiles for candidates. These profiles are built from the answers from applicants and the way that people give the answers. HireVue used to extract the speech and facial movements from applicants’ interviews and map them to different emotions to build these profiles. Recently, they stopped using facial analysis, but intonation and behaviour analysis is still a part of the assessment (Knight, 2021). However, Barrett et al. (2019) has shown that these systems cannot deal well with the vast range of how people show their emotions, resulting in assessments that do not accurately portray applicants and that are hard to disprove (Kim, 2017).

HR tool makers are however incentivised to present their products in a more positive light in their marketing than reality, as their business model is, of course, tied to selling these systems. Most tool makers claim that their tool has a solid scientific foundation and, in AI systems particularly, that it is more objective than humans because AI systems judge only on the psychometric data gathered. However, as an algorithm is trained on historical data, it is likely that it will reinforce current biases and it can be very hard or even impossible for outsiders (or the tool maker themselves) to understand how exactly an algorithm is coming to its conclusions on who is the best person for a position. That in turn makes it harder for companies to judge effectively whether a psychometric tool is ethical to use and how they should apply it (Dattner et al., 2019). Additionally, because these systems are often not audited by external parties, governments often cannot intervene or even know when an AI system makes a wrong decision (West, 2020; LCEF, 2020).

Biases in historical data is the focus of many working on AI ethics (Broek et al., 2019; Dattner et al., 2019; Kodiyan, 2019; Raghavan et al., 2020; Selbst et al., 2019; West, 2020), but there is an additional problem. Because AI systems can, at their most basic level, only deal with numerical data, a choice to use AI systems to check applicants for certain competences is an implicit choice to quantify these competences (Sareen et al.,

2020). But can each competence be expressed well in numbers? If the quantification of competence leads to a loss of meaning and nuance, that could limit the self-representation of an applicant. That meaning or nuance may be an integral part of the message an applicant is trying to convey during an assessment.

Additionally, the different assessments that tool makers like HireVue and PredictiveHire use as input make it very hard for job applicants to keep control over how they represent themselves: the way the personality assessment is made is opaque or even invisible for users and it partially depends on characteristics which you cannot change like how your voice sounds (Chamorro-Premuzic & Adler, 2015). This creates a situation which is unfair, as the applicant is judged on unchangeable characteristics, and which is hard to rectify, as there is no human who sees both the input and output side of the algorithm and who can ring the alarm bell if those do not align.

This thesis focuses how autonomy over self-representation within hiring can be damaged by AI systems. This autonomy is not created in isolation by the actions and motivations of the applicant. Instead, it is a culmination of the expectations and actions of the applicant, the interviewer, and society. Because the act of hiring personnel and the value of autonomy over self-representation are inherently dependent on these interactions among different stakeholders, it is important to define the different groups of stakeholders and their motivations in order to ensure they are all considered in the design process.

Stakeholders

- **Job seekers** want to find a suitable job (fitting their skills) and be treated fairly during the selection process. From interviews with recent graduates, I learned that most people do not have more than a couple of months of directly accessible savings, which makes them want to find a job quickly. Additionally, it has become normal to do a lot of applications before landing a job, so people do not want to spend a very large amount of time on every single one. At the same time, they also want a job that they believe fits their skills and interests. Within the job application process they want to be able to show what they are looking for and what they find important in a job.

- **Employers** want to find a person for their role quickly without spending a lot of resources for each position. In this project, interviews were conducted with HR professionals for their perspective. Especially for larger companies, they often apply tools from other companies to improve their hiring process, which brings a lot of additional legal, financial and ethical considerations along with it. Smaller companies rely more on less procedures and tools. Additionally, both groups tend not to be experts in the field of AI, which makes it harder for them to make good decisions in acquiring HR tools.

- **AI tool makers** want to create tools that make HR departments more effective and that increasingly look to algorithms to achieve that. They have to juggle multiple priorities: delivering high quality tools that give applicants a fair opportunity, saving HR personnel time, and getting enough sales to stay in business.

- **HelloMentor** is the client for this research project. They are an organisation that connects students with mentors in large organisations for different tech

fields. They help with discovery to eventual guidance on finding a job. Given that finding a job is becoming a process that involves more and more digital aspects, for them, it is very important to have a good idea what the impact of AI systems on their work and their students is. Also, they are bringing expertise in HR to the project.

- **Broader society, and by extension governments**, want a productive workforce, as those contribute to society in many essential ways. Overall, many of the things we want in life – good education, accessible healthcare, solid infrastructure – are less available when participation in the labor market is low. Therefore, governments have a vested interest in creating an environment where both job seekers and employers succeed.

Goals of the project

With this project, the focus of the research is to discover through empirical experiments how AI tools can infringe on human dignity by limiting autonomy over self-representation and how those problems can be mitigated without completely closing the door to AI technology in this field. These mitigations are then formulated in a potential design for the AI system, based on the empirical findings. To that end, the research questions are:

How do AI systems inhibit autonomy over self-representation in the context of hiring?

With the following subquestions:

How do organisations currently use AI systems in hiring?

How does self-representation play a role in the hiring process?

How can these inhibitions be mitigated?

A definition of AI

If there is one thing that defines AI, it is the vagueness of the term. Marketing departments have taken the term and applied it to anything that shows a modicum of adaptive behaviour. Additionally, the goalposts for keep being moved over time. The quote below highlights the AI effect: once something can be done by AI, it is argued by critics not to be intelligent (Haenlein & Kaplan, 2019). Given the importance of AI systems in this thesis, it is helpful to establish a common definition for further use.

"AI is whatever hasn't been done yet."

Douglas Hofstadter (1980)

Artificial intelligence is a field with many different applications, techniques and philosophies. The vagueness makes it hard for people understand what one is actually talking about, especially because many companies and organisations see fit to use the term as a way to hype up their product or project that is supposed to be 'smart'. This is called AI-washing (Hao, 2019; Tschopp & Ruef, 2018). For the purpose of this graduation project it is useful to establish one definition of AI to (somewhat) remove this vagueness from the discussion.

The field of artificial intelligence officially got started in 1955 during the Dartmouth Summer Research Project on Artificial Intelligence (McCarthy et al., 1955), where a small group of computer scientists came together to learn to apply the notion of machine intelligence, of making machines that are comparable to a human across the spectrum of intelligence (see Figure 1). They expected to do it in about 8 weeks. Of course, now we know they severely underestimated their goal, which still has not been achieved almost 70 years later. We have had multiple AI winters and springs where interest in AI cooled down and subsequently rose again when some breakthrough happened. Part of why the Dartmouth Workshop members were hampered in their task (and why AI is such a contentious term) is that there is no consensus on the term *intelligence* in general. There are many meanings in different fields, because we do not have yet a complete understanding of the concept of intelligence (Legg & Hutter, 2007). One definition is that intelligence is an extensive and fluid concept that encompasses

Figure 1. Several participants of the Dartmouth Summer Research Project, including organisers Marvin Minsky, Claude Shannon and Nathan Rochester.

logic, emotional intelligence, learning, creativity and many more aspects of our cognitive abilities (Levesque. MIT, 2018). This of course has its ramifications on the term artificial intelligence, as it begs the question as to what we want a machine to be able to do in order to call it intelligent.

Currently, we are in an AI hype peak buoyed by vast processing power and extensive communication networks that generate, transfer and process vast amounts of different kinds of data (Alzubi et al., 2018). This led to a focus in the technique of machine learning (ML), which uses large amounts of data to change the rules it uses to process new data (Rahwan et al., 2019). This could be seen as a form of learning and it has many useful applications in the world (Rahwan et al., 2019). So, the capability to learn, to adapt behaviour dynamically to changes in the environment is an essential part of the definition of AI for this project.

What is important to understand is that machine learning is a technique that mainly focuses on statistical correlations for inferences (Jordan & Mitchell, 2015). This makes that while ML can often be great at finding subtle relations among data points that are practically impossible or very expensive for humans for discern, it also means that the relations that an ML algorithm finds are not necessarily causal. In other words, ML is not capable of reasoning about the world to the extent a human can (although GPT-2 and GPT-3, two large scale language models, can be argued to have a surprising capacity for reasoning (Floridi & Chiriatti, 2020; Tamkin et al., 2021)). For example, AI systems are not able to properly transfer knowledge in a certain domain to another (Pan & Yang, 2010). Our definition needs to include these limits in current AI capability, because it is essential for knowing what tasks can be entrusted to current technology.

Additionally, there is an adjacent topic that needs to be taken into account. "Wishful mnemonics" are linguistic shortcuts that experts sometimes use to explain that a computer does some 'human' task (McDermott, 1976). Examples are "*The computer understands the problem instantly*" or "*The moves will help AlphaStar get to its goal.*" Such systems do not have these attributes, but these words imply that there is understanding when there is none. Mitchell (2021) shows that people can misunderstand the capabilities of an AI system when journalists for example turn a reading comprehension test result into the headline "*AI can now read better than humans, test results show.*" Because these mnemonics are so common, people who are not well versed in AI technology can get a misguided view on what exactly an AI system is able to do.

This project looks at AI systems in a larger socio-technical context. Already a lot has been written on AI systems as part of socio-technical systems, arguing that they are not just technical systems, but that it is imperative to see their impact through interactions with human society (Behymer & Flach, 2016; Coeckelbergh, 2020; Jones et al., 2013). Van de Poel (2020) argues to distinguish AI systems from traditional technologies by

"its capacity to autonomously interact with its environment and to adapt itself on the basis of such interactions. This capacity creates new opportunities for embedding values in AI systems that do not exist in traditional socio-technical systems."

I will use this definition for the entirety of this thesis, with the addition to minimise the use of wishful mnemonics.

Conclusion

In this chapter, I have defined the term AI, looking at the historical origin, technological underpinning and how the term gets used in wider circles, noting the danger of wishful mnemonics. The definition I use notes the autonomous interaction and adaptation characteristics of AI, and how that has ramifications for designing for certain values with AI systems involved.

Data insight

AI systems are unique in their autonomous ability to interact and adapt to their environment, which creates new types of opportunities to embed values in the socio-technical system in which they exist.

Method

As this project focuses on the intersection of technology and society, it is important to recognise that technology is inherently value-laden: both the development and introduction of a technology take place in a social context (Van den Hoven, 2007; Friedman & Hendry, 2019). We want a technology to do something for us and consider it successful if it hews closely to the values inherent in that goal. To that end, with this chapter I describe how my approach hinges on connecting values to technology and society.

What follows is a description of my personal values to show my personal bias and point of view. That is followed by the project approach, which is based on Value Sensitive Design (VSD), a framework developed for designing with specific values in mind in order to get a technology to do what society actually wants. Finally, I explain why towards the end of the project I brought forward a framework I will call Adaptive Systems Framework (ASF) because of limitations of VSD.

Personal values

The computer science community has in recent years recognised the impact that especially AI development can have on society and has established multiple codes of ethics like the ACM Code of Ethics and Professional Conduct (ACM, 2018) and the ICLR Code of Ethics (ICLR, 2020). Additionally, more work being done on new publication norms that focus on (potentially negative) side effects of the work researchers have been doing (Gupta et al., 2020). As my research deals directly with the societal impact of AI, I want to hold my work to the same critical stance. In particular, I will shortly describe my own values relating to the project because my personal bias *will* color this project. Additionally, in the conclusion, I will reflect on the stance I took with this project and what AI projects mean for design as a profession.

I believe strongly in liberal democracy and the innate value of a human being. All people should be treated equally and with dignity. Given the risk of large unforeseen consequences, I am quite sceptical about claims that current AI technologies should be allowed to run large parts of our society.

Instead, I see a need for a more reflective approach to applying AI technology. As humans interact with technology, we learn and modify our behaviour to that technology, which is already happening with AI technology (watching what Netflix recommends is a fairly benign example). We should lean into that interactional way of thinking and constantly adapt along with our application of AI systems. Therefore, I see more value in specific, context-dependent interventions where AI systems can be a supporting element in human-first projects.

Project approach

As this project is focused on creating autonomy over self-representation in AI hiring systems, it is helpful to work with a framework that supports designing with certain values in mind. VSD is a design approach that explicitly takes (moral) values into account during the design process of socio-technical systems. It was originally developed by Friedman in 1996 in the field of software development and has been further expanded upon in the 2000's (Borning & Muller, 2012; Cummings, 2006; Friedman et al., 2017; Le Dantec et al., 2009; van den Hoven, 2007). VSD was built on the idea that an integrative approach is needed for successful value inclusion in design projects, as values inherently involve tradeoffs in functionality such as the value of privacy and safety; perfect privacy in a messaging system leaves big problems for safety as law enforcement is not able to get access to the communications of criminals.

VSD applies three different types of investigations to fully understand the existence and application of values in a certain context. This is called the 'tripartite methodology', consisting of conceptual, empirical and technical investigations (Friedman et al., 2002). Conceptual investigations focus on the analytical or philosophically informed inquiries that are done around the central value of the investigation. Empirical investigations bridge the theoretical understanding of conceptual investigations towards the real

world and use the entire spectrum of qualitative and quantitative methods of the social sciences to understand the value tradeoffs people make in everyday practice. Technical investigations take the technological aspect of a context into consideration and look particularly at both how a technology may hinder or enable a certain value and how a technology may adapted or used in a system to enable the expression of a value. It can be described as the value affordance of a technology (Norman, 2013).

Because of the wildly different outputs of the different investigations, VSD puts emphasis on an iteration loop of doing investigations, where each investigation feeds into the next iteration. This leads to an iterative integration that is more than the sum of its parts. It is also an interactional theory, positing that technology shapes human experience and societal movement, and humans in turn shape technological advancement and implementation.

VSD has gotten multiple contributions and additions over the years. Alsheikh et al. (2011) showed the importance of accurately and precisely framing your view on the values you are researching and the culture you are designing for. Le Dantec et al. (2009) argues for putting empirical investigations first in the order of the tripartite methodology to put more focus on a grounded analysis of values found in the field instead of the already established value framework that can come if a conceptual investigation is the starting point. Additionally, more scrutiny on the idea of universal values, value lists and strengthening the voice of participants while making the inherent voice of the researcher more visible are all suggestions that Borning and Muller (2012) have made.

For this project, the tripartite investigation can be seen in Figure 2. I started with a technical investigation to understand AI systems in HR and a conceptual investigation in the value of self-representation in humans. This was followed by interviews with stakeholders to better understand how applicants use their self-representation, how HR managers see the hiring process and ethics in their work and how both parties looked at AI systems in hiring. I then used those findings to build an ecosystem map and journey map.

However, one of the findings of my interviews was that none of my interviewees had actually engaged with AI systems yet in a hiring context. When making the ecosystem and journey map, I noticed that lack of experience in the depth of my mapping: I had trouble understanding what people valued in such interactions. To better elicit the values that people have in hiring specifically when AI systems are involved, I applied generative prototypes (Lim et al., 2008) to elicit potential values from participants.

While prototypes usually tend to be used for validating hypotheses or production testing (Ruecker, 2015), generative prototypes are often used in the beginning of a design process to probe the context (Sanders & Stappers, 2012; Visser et al., 2005). They are design artefacts specifically made for design research purposes. Generative prototypes uncover latent or tacit knowledge, knowledge that is hard to articulate or that people are not yet aware of that they know. Provotypes (Boer et al., 2013) would have another good option to use, but they are built on the tenets of participatory innovation, involving stakeholders in the design process (Boer & Donovan, 2012), which is something I did not do due to practical limitations of the project.

People rely constantly on their value system on how to act and how to judge the

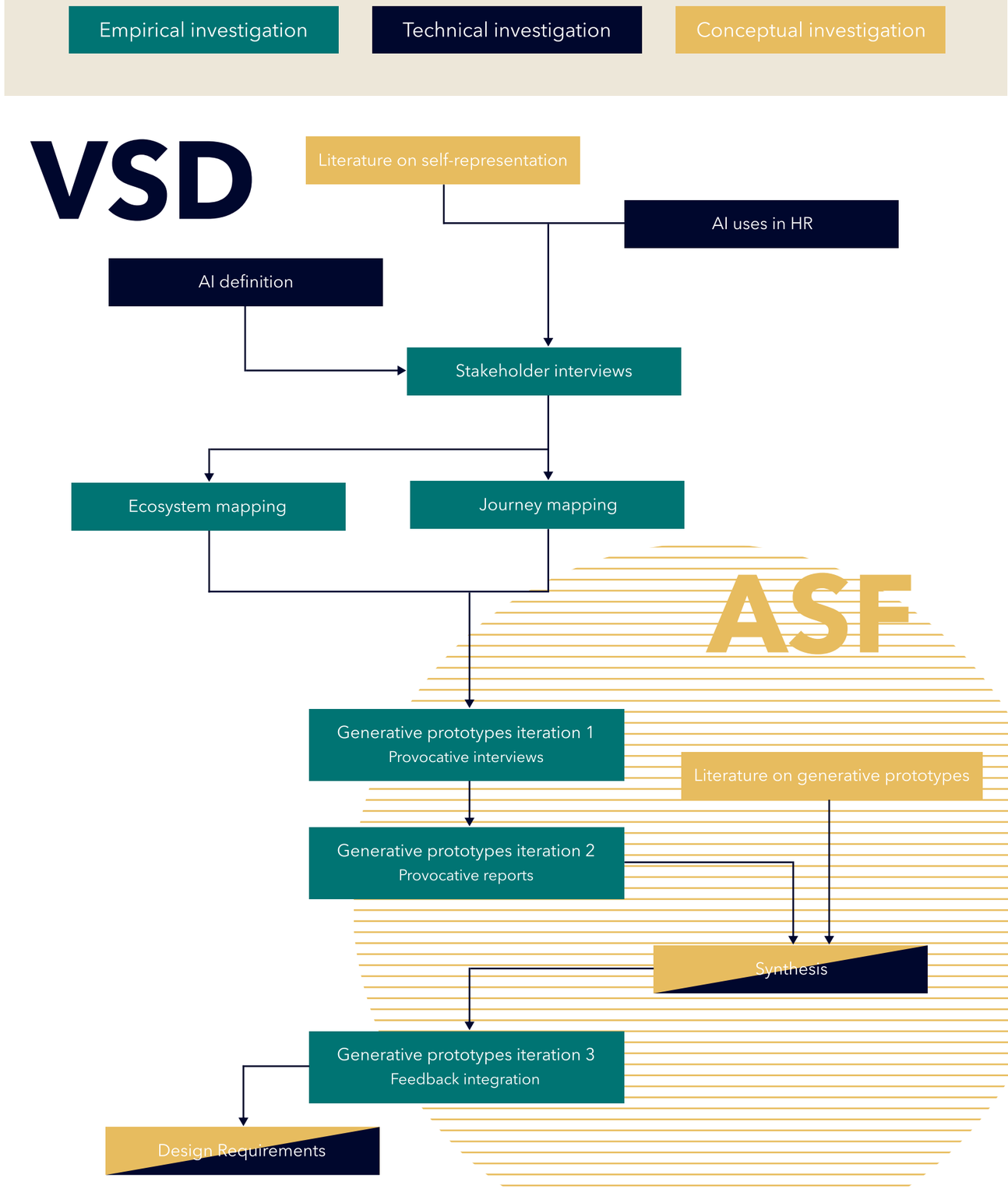


Figure 2. The many different investigations of this project. While the overall methodology follows VSD, the lined circle highlights the investigations that are part of the ASF method.

outside world, so they are applying their abstract reasoning skills to think about their values and the impact on the world. However, for many this mainly happens in the context of real life situations and not in a more formal setting that requires more explicit formulation of one's values. Therefore, this information for most people is *tacit*. I built generative prototypes that provoke to elicit this information better, hypothesising that the unusual situation caused by the provocation will help people to articulate their point of view.

However, methodologically, within VSD it is uncommon for prototypes to be used in this way (Friedman et al., 2017). Additionally, as explained in the previous chapter, AI systems are defined by the fact that they are autonomously adaptive, learning systems. As they change, the values they express and elicit also change, which is something that VSD does not account for (Van de Poel, 2021). For these reasons, I turned to a framework by Stoimenova (2021) that is specifically made to apply generative prototypes in the context of AI systems.

Adaptive Systems Framework

Adaptive Systems Framework is currently being developed by Stoimenova (2021) as a cognitive framework to help designers elicit unexpected outcomes that can happen during the development and deployment of AI systems. It proposes a set procedure for doing design iterations with an AI system, that is focused on aligning the outcomes of the system with the stated purpose by analysing whether the outcomes are positive or negative and where in the design they come from.

The framework by Stoimenova currently has no official name yet. I use the name Adaptive Systems Framework for conciseness and readability reasons. It may ultimately be published under a different name.

Table 1 shows the different elements with their description. I want to highlight in particular the element of the frame, as Dorst (2011) notes that for complex open-ended problems, framing is essential for a high quality design outcome. However, Stoimenova (2021) notes an additional characteristic of framing that is beneficial for this type of work. The act of framing forces the designer to make explicit the value they intend to create, which will help me in defining what exactly I intend my prototypes to elicit in values.

The framework can shortly be described in the following three steps: combining the purpose with data allows one to form a frame,

purpose + data → frame

the frame combined with a mode of action and an intended actuation outlines a prototype,

frame + mode of action + intended actuation → prototype

Element	Definition
Purpose	The intended outcome the AI system should deliver to its stakeholders, formulated in a solution-neutral manner.
Data	Academic research and records produced by (previous) experiments and measurements that serve as evidence for the existence or features of a phenomenon.
Frame	A cognitive act of looking at a problem situation from a specific viewpoint that informs the design of the working mechanism of the solution. It follows the structure: "IF we look at the problem situation from this viewpoint, and adopt the working principle associated with that position, THEN we will create the purpose we are striving for.".
Mode of action	The (functional) behaviour of the AI system in response to influences exerted on it from its users and environment.
Intended actuation	The intended by design action of a user that allows the AI system to function and be 'connected' to its immediate environment.
Prototype	A concrete representation of a potential AI system and its experimental setup. It supports the anticipation of unintended values the agent's actuation in a sociotechnical context might create.
Observed actuation	The intended and unintended actions a user performs in order to allow the AI system to function and be 'connected' to its immediate environment
Value	The intended and unintended outcomes of the AI system's actuation by its stakeholders. These could be instrumental and/or intrinsic values (both positive and negative).

Table 1. The definitions of the elements of ASF.

and the prototype combined with the observed actuation leads to the expression of a certain value (usually multiple),

prototype + observed actuation → value

At this point, the designer can analyse the values on whether they are positive or negative and incorporate the positive values in the purpose while explicitly designing against negative values in the next iteration of the design.
This framework allows designers to quickly and iteratively form hypotheses on how a potential AI system may work in its context. To that end, one employs generative prototypes as a means of building and exploring hypotheses, just like I have done. That

Purpose

Design a strategy for HR departments to realise autonomy over self-representation with AI systems in hiring.

is the main reason that I chose to adopt this framework, as it allowed me to analyse the results from my generative prototypes better and iterate further on them.

In the conclusion of each chapter I define what that chapter contributed to different elements of the framework. As I have now introduced the framework, I can phrase the original purpose of the project according to ASF. As ASF focuses on eliciting values through empirical means while VSD places equal importance on technical, conceptual and empirical investigations, the inclusion of ASF does put more relative emphasis within the project on the generative prototypes as primary sources of design insights. While a comprehensive inspection on the relation of ASF and VSD is not within the scope of this project, the focus of ASF on empirical work can work within the larger tripartite investigation concept of VSD, while contributing to the idea of value change that is starting to become known in VSD (Umbrello & van de Poel, 2021; van de Poel, 2021). Because the frameworks operate on different abstraction levels, I have not had significant problems integrating them into this project.

The conceptual, empirical and technical investigations I did before integrating ASF in the project lend themselves well as serving as data points to kick off an ASF iteration. Likewise, the values found in an ASF iteration can again be used for future investigations in VSD (although this project ended with an empirical ASF investigation).

Conclusion

In this chapter I have explained my method: how I use both VSD and ASF, how those relate and how the application of generative prototypes made for the later addition of ASF. VSD is the overarching methodology and philosophy of this project, with the inclusion of generative prototypes and ASF to account for both the lack of experience most people have currently AI hiring systems and the focus on value change.

Design research

In the beginning of the project an extensive conceptual investigation was conducted in order to get a more comprehensive overview of the context. By looking at the state of the art of the concepts of AI usage in hiring and autonomy over self-representation, I can get a better understanding of why and how AI systems currently get used in hiring, how they are problematic, and how that relates to the idea of autonomy over self-representation.

AI in hiring

HR recruitment is looking more and more to AI systems for both sourcing and assessment for two reasons: “talent scarcity and applicant glut” (Li et al., 2021; Paschen et al., 2020). For sourcing, AI is seen by many employers as a way to hire more diversely. Some researchers see in the application of AI systems possibilities to reduce human bias (Black & van Esch, 2020; Raveendra et al., 2020), while others warn about the specific algorithmic bias that may take its place when AI systems are implemented (Lee & Shin, 2020; Raghavan et al., 2020). For example, misuse of AI hiring systems may lead to groups in society who are practically excluded from the workforce (Fuller et al., 2021).

A major factor as to why HR departments want to use AI systems is the monumental shift in where value generation happens within an organisation. In the 1900s, the majority of value generation was based on tangible assets, such as factories and tooling (Lev, 2001). By the 21st century, 65% of value generation started being accounted for by intangible assets, including people (Black, 2014). To add to this, the increased productivity per employee aided by technology makes it all the more important to ensure employees are aligned with the organisation (Oehlhorn et al., 2020; Sahoo et al., 2011). Combined with the fact that high performers can have a productivity gap of up to 800% with an average performer in high complexity jobs (Keller & Meaney, 2017), this makes finding good people essential in order to succeed in the competitive landscape of today, forcing HR departments to spend more time and effort in finding the best performers (Black & van Esch, 2020; Hamilton & Davison, 2018). All these elements have resulted in relative talent scarcity, as talent has become much more valuable and therefore, sought after.

The other reason, application glut, is caused by the trend of job hunting went digital. Friction for people finding jobs became practically nothing, resulting in an explosion in the amount of applications per job posting (Black & van Esch, 2020; Maurer & Liu, 2007). For example, in 2017 Johnson and Johnson received over 1 million applications for 28.000 openings (McIlvaine, 2018). All these applications need to be assessed and AI systems get used in remedying this glut of applicants, primarily to increase efficiency and improve throughput of the hiring process.

AI systems used in assessments are used for a number of purposes. Table 2 on the following page breaks down a number of systems of popular AI hiring tool makers on what services they provide to their customers. The list of tools is based on a survey done under HR professionals (Li et al., 2021) plus market research with the activities determined by analysing the websites of the companies. Many tools are complete recruitment suites or integrate deeply into other applicant tracking systems (ATSs). All vendors talk about ethical AI in their promotion material but it is all focused on diversity and removing bias, not on maintaining self-representation for applicants.

Company	AI powered services	Extra info
HireVue	Analytics, Skill assessments, Video interviews	Integrates with ATS
Pymetrics	Analytics, Skill assessments, Sourcing, Video interviews	Integrates with ATS
Retorio	Analytics, Video interviews	Integrates with ATS
Bryq	Analytics, Skill assessments, Video interviews	Integrates with ATS
Entelo	Analytics, Recruitment writing, Sourcing, Video interviews	Replaces ATS
Textio	Recruitment writing	Integrates with ATS
Knockri	Analytics, Skill assessments, Sourcing, Video interviews	Focused on diversity, Integrates with ATS
Harver	Analytics, Skill assessments, Video interviews	Focused on large scale recruitment, replaces ATS
Hiretual	Recruitment writing, Sourcing	Integrates with ATS
Vervoe	Analytics, Skill assessments, Recruitment writing	Integrates with ATS
InterviewStream	Video interviews	Integrates with ATS
Traitify	Analytics, Recruitment writing, Skills assessments	Integrates with ATS
PredictiveHire	Analytics, Skill assessments	Integrates with ATS

Table 2. An overview of companies offering AI powered hiring services.

These different services all apply AI technology for different purposes.

- Analytics services provide recruiters with additional information about applicants. Algorithms will sift through data gleaned from CVs, LinkedIn profiles and other sources on the internet, depending on the tool. They offer insight on what applicants may be potential high performers or potential company fit. Autonomy over self-representation can be impacted through the analysis of information people put on the internet without intending it be used for this purpose or knowing how it will be processed.

- Assessments are intended to infer skills or personality traits relevant to the job at hand through small exercises or games that applicants have to do. They can be simple multi-choice questions, coding challenges or spreadsheets applicants need to build. AI technology is employed here mainly to grade and sort applicants. Autonomy over self-representation can be impacted through the answers that people give.

- Sourcing services are AI-powered search engines that help recruiters find potential applicants and their contact information. AI technology support recruiters by suggesting other search terms or assessing availability of applicants. Autonomy over self-representation is less of an issue here because applicants are still free to express themselves in the way they want, as they are still outside the application process of the company.

- Recruitment writing is a service supporting recruiters in their messaging to applicants to maximise engagement. Algorithms measure the engagement of each individual message sent and apply that information to suggest following messages. Autonomy over self-representation is not inhibited here because applicants are not limited by the AI system in how they can respond.

- Video interviews focus on assessing personality traits and skills based on asynchronous Q&A sessions where applicants answer preset questions. The AI system involved does not only sort and grade participants on their answer, but also may analyse facial and vocal expressions to infer personality traits. Autonomy over self-representation can potentially be inhibited in several ways: answers may be interpreted in ways applicants do not expect or decisions may be made based on facial expressions that are interpreted erroneously (Barrett et al., 2019).

As the focus of this project lies in how to maintain self-representation of applicants in the face of AI tools, video interviews tools are the most interesting activities, as there the chance that a system may inadvertently limit self-representation for an applicant is the greatest.

In the case of video interviews, AI is integrated not in the activity itself (the system does not for example alter its questions based on previous answers), but is involved through grading and analysing the results from applicants and reworking that data into overviews for recruiters so they can decide quickly on who to retain. As the AI system works as an agent to package and present information more efficiently to the recruiter, it becomes harder for the applicant to bring information to the recruiter in the way that they think is best, limiting their self-representation.

That is a problem because the repackaging of information may lead to discriminatory practices where the answer of one person may be seen as preferable to another while the meaning of the answers is the same. The meaning just got expressed differently. They did not do anything wrong, the system was just not able to deal with the diversity of expression that people have. That, as we will see in the following section, could infringe their self-representation and therefore, their human dignity (Schachter, 1983).

Human dignity and self-representation

Human dignity comes from the belief that every human being should be respected for their intrinsic worth, that they worth respecting in their "individual choices in such matters as beliefs, way of life, attitudes and the conduct of public affairs" (Schachter, 1983). As the quote from Buchanan (2001) beloww shows, it is a foundational element of human-centered design. But it is also an essential part of the Charter of Fundamental Rights (European Convention of 1999, 2012), as it is one of the guiding concepts on how to define human rights (Aizenberg and Van den Hoven, 2020).

Because human dignity is such an important concept in human rights, violations of it are worth researching and designing solutions for. According to Halbertal (2015), one of the types of human dignity violations is *humiliation*, which is by Aizenberg and Van den Hoven (2020) elaborated as: "being put in a state of helplessness, insignificance; losing autonomy over your own representation."

"Human-centered design is fundamentally an affirmation of human dignity. It is an ongoing search for what can be done to support and strengthen the dignity of human beings as they act out their lives in varied social, economic, political, and cultural circumstances."

Richard Buchanan (2001)

Autonomy over self-representation is a concept in the social sciences (Risam, 2018; Vivienne and Burgess, 2013), focused on the idea that a person has the capacity to direct the narratives in which they play a role themselves. It is a way to make their identity part of how they show themselves to the world with them acting as a direct representative of themselves for their identity, motivations, and experiences.

Our digital lifestyle has allowed us many new opportunities, such as sending medical dossiers from one hospital to another quickly, connecting with each other over social media, or booking international flights online within minutes. All this is possible due to (re)programmable artefacts for interacting with the world, aka computers. These systems separate functional logic from the physical artefact, allowing for many options to decouple functionalities (Kallinikos et al., 2013; Yoo et al., 2010). However, this decoupling also necessitates other parties to know (parts of) our identity: if you buy groceries with cash at a supermarket, you are relatively anonymous, but to enjoy the convenience of home delivery, you need to supply the supermarket with your home address and payment details, allowing the other party to collect more data on your identity.

As governments and organisations need to identify people for these purposes, our identity becomes more codified (Manders-Huits and Van den Hoven, 2008). We have less opportunities to represent ourselves because these systems are built with a certain purpose in mind, like medical dossiers or tracking airplane passengers for anti-terrorism purposes. These systems are focused on their purpose and do not record more than necessary. This has multiple reasons. First, a technical point of view: to fully process the vast amount of experiences that is a human life, let alone the seven billion people currently living, we would need sensory and computing capacity possibly only found in a Matrioshka Brain: a solar system sized computing system (Bradbury, 2000); in other words, science fiction. Second, an ideological point of view: the extensive data collection and processing would mean privacy effectively would not exist in such a system; it would be a perfect implementation of Big Brother.

Instead, all current information processing systems have a certain focus and record only limited parts of our lives. However, that does mean that the system and people using the system may end up with assumptions on who you are, based only on the information that is available, which may distorted or even false. Van den Hoven and Manders-Huits (2008) explores this idea with the concept of statistical dehumanization: the "conception of a person as it is presented in (sets of) probabilities conflicts with one's self-perception of who one is (oneself as 'self-presenter' presenting a particular identity) or what one is or should be seen as and respected for". Aizenberg and Van den Hoven (2020) argue that algorithms, through a rigid group of quantitative inputs and outputs, remove fluidity, context, and contestability (Selbst et al., 2019) from the highly complex situation where these algorithms get used, such crime prediction and job applications. These discussions show that AI systems are by their nature limited in the meaning they can convey (only data that they can quantify gets presented as the truth) and in that way they infringe on the autonomy over self-representation of job applicants.

In recruitment, there has been a great push towards using applicant tracking systems (ATS) to improve efficiency (Hacioglu, 2020; Laumer et al., 2015). These ATSs help recruiters stay on top of large amount of communications with different applicants and manage their information. Often, the systems are able to tie in to different kinds

of tests and job boards, making the job easier for recruiters by creating a more integrated experience, but it therefore also leads to less options for job applicants for self-representation: because these systems can only work with certain kinds of input and are less flexible in dealing with unexpected input, the applicant is forced to show themselves in the way the system works. For example, the Q&A style interviews most systems employ allow for less options for applicants to steer the direction of the interview and the reports the systems generate about each applicant are limited in conveying meaning by quantification effects as outlined above.

In recent years this issue has been amplified by the mass introduction of AI tools in the field (Gupta et al., 2018; van Esch et al., 2019). AI systems makes representing oneself harder because it is very hard to know in what ways the AI system will process your data. This leads to situations where you cannot know what the recruiter will see on their end, so even if a human is making all the decisions on who to hire, this problem of lessened autonomy over self-representation remains.

Conclusion

The conceptual investigation of AI in the hiring space shows that AI ethics in that context have been almost exclusively focused on removing biases and advancing diversity (Kodiyan, 2019; Mujtaba & Mahapatra, 2019). The potential loss of self-representation is not a visible issue in both industry and the scientific literature. This shows the relative neglect of this problem and reinforces the urgency of research in this direction.

Meanwhile, the reason that AI systems have been adopted so widespread in hiring in the first place is the need of efficiency and efficacy in hiring, driven by a large amount of people applying for roles and the fact that people have become over time a more essential part for value generation for organisations and individually more productive due to technology.

Data insight

The need for AI systems is driven by both efficiency and efficacy: organisations need to process applicants quicker due to volume increases and have more to gain by making less mistakes in hiring. The genie is therefore out of the bottle: the use of AI systems in hiring will be very unlikely to be reversed.

The literature study into autonomy over self-representation showed a potential danger in storage and processing of people in many systems that already happens today. Our lives have become more defined and decided by systems that do things for us but that inherently also represent us within those systems. AI systems in hiring can extend that problem by presenting their ranking of applicants as the truth with relatively little room for applicants to put things in context or contest understanding.

Data insight

AI systems have many ways to limit the autonomy of applicants over their representation through the removal of context, fluidity and contestability.

Interviews

Value Sensitive Design places a lot of importance in understanding (the relation) of human values. To that end, the purpose of the interviews was twofold. A: To get a better understanding of how stakeholders are experiencing the context, semi-structured interviews (SSIs) were conducted. I chose SSIs because they provide a lot of opportunity to ask probing questions on topics that people otherwise may not want to talk about (Adams, 2015). And B: to map the values and their interaction that the different stakeholders in the recruitment process have. The goals of the stakeholders may differ, which may lead to differing (prioritisation of) values.

Setup

In total, four interviews were carried out with recent job seekers and HR professionals (see Table 3). They were opportunistically sampled, relying on my own network to find them.

Interviewee	Field	Location
1	HR in small and medium businesses	The Netherlands
2	HR in a large technology company	France
3	Designer in a multinational consulting firm	The Netherlands
4	Researcher in a social organisation (background in energy transitions)	The Netherlands

Table 3. An overview of the interviewees.

The intention for the HR interviews was to focus on the recruitment process from the recruiter's point of view and to discuss ethics in the workplace. For the job seeker's interviews, the focus was on seeing their point of view on going through a job application and how they felt they were able to represent themselves within that process. The fact that we are talking about the recruitment and application process indicates how different the experiences and activities are for the parties, even though the effect of the processes are intertwined: getting people into jobs. All interviews were coded and then clustered using thematic analysis (Braun & Clarke, 2006) to find interesting relations, and common and unique themes.

Descriptions

Interviewee 1 is a HR interim manager, helping mainly small and medium enterprises with solving acute operational problems, like issues with function descriptions, personnel manuals, labor law, and reorganisations. Every 6 to 12 months he is on a new assignment in a different company. Interviewee 2 is head of recruiting of the EMEA region for a multinational company in IP video and access control. He is mainly recruiting sales people and developers, but also managers. The interviews showed a fairly standardised process for recruiting new employees: defining the position, advertising the opening, multiple rounds of interviews and selections (sometimes with tests) until one person is decided to be the best person for the job.

Interviewee 3 recently concluded his Master's in Industrial Design and started working as a service designer at a design studio part of a large consulting firm. Interviewee 4 had half a year off between his bachelor (in earth and sustainability sciences) and his master; started a temporary job as researcher at a social research organisation in Groningen. Both are fresh entrants to the job market, with not much experience and for who the experience of finding a job was very much also an exploration for themselves on what they find important and how that relates to the jobs they were interested in.

Results

The coding and clustering of all interviews can be found in appendix B.

Both HR interviewees stated a strong adherence to ethical handling in their work. Interviewee 1 said integrity and transparency are very important in how he feels he should do his job: "*Say what you do and do what you say.*" Interviewee 2 always tries to be an objective assessor of people, but he is also aware of the limitations of his objectivity. This came back more often, both stated that hiring mistakes are inherent to the job:

> *"I realise that it is not through three or four hours of interview that you can define someone."*

Overall, the HR interviews show that the interviewees feel they are trying to be ethical and treat everyone the same way. Additionally, it shows that the current process of hiring is very focused on the needs of the company (they define the process on their own after all) and that there is always a certain amount of risk involved with hiring new people, on whether the people will actually be a good fit. Interviewee 2 saw the current use of tools in HR as a way for employees to deflect responsibility of their own job if someone ends up being a bad hire:

> *"At least I'm not the only one involved because, see, we did some tests, we ran his CV or technical test through that kind of tool, which said it was okay, so I'm not responsible for that. We've got that chain of processes to show you we thought that it [the new hire] was a good one."*

Interviewee 3 had an extensive job search applying to over 20 companies in the span of two months. He was fairly unsure about what he wanted to do, but was certain he wanted to do something different from his focus in his master: rich interaction design. Although he did not hear from his ultimate employer for five weeks after his application, once the process started, he was offered a job within 5 working days. He did two interviews, a personality test and a competence test during the selection. He was surprised by how relaxed the interviews actually were and although he did not ace the tests, the results never came up during a later conversation.

Because he did not know well what he wanted to do, he never went into the conversations with a specific goal. Instead, he focused on simply having an engaging conversation where he tried to learn as much as possible about the company:

> *"Because I didn't exactly know where I wanted to go, I leaned on my... personal charm. I tried to have nice conversations and a good energy."*

Overall, he felt he was able to show himself well during the interviews because he felt he could be open to them about his doubts and not needing to put up a facade. Interviewee 4 had a similar experience, taking a couple of months to find an interesting job. He wanted to do research and analysis work as a preparation for his master. Because he knew it was a temporary position anyway, he was not afraid to look for jobs in places where he normally would not look. His background is in energy transitions but his current job is more in social research, of which energy transitions is only a small part. His application process was very straightforward with only a phone call and two interviews.

He felt he was able to represent himself well by doing three things: being well-prepared, showing his skills in economics and earth sciences from his education and showing his personal affinity for sustainability, also through his volunteer work.

The interviews with the job seekers show how they found a serious job for the first time. The fact that they felt they were clear on what they wanted to show and because they felt the companies were open to learn about what kind of person they actually were made them feel they could represent themselves well. This last part shows how important it is for companies to leave room for applicants to be themselves.

After the interviews were conducted, the recordings were coded and subsequently clustered. An overview of all clusters in appendix B which shows different clusters are closer in meaning and context. All clusters from both HR professionals and applicants are also in Table 4 and 5 on the following pages with a discussion of the clusters afterwards.

Cluster name	Explanation
HR needs to solve every problem that involves humans themselves	HR has a varied task list, because humans are in every part of the organisation. Therefore, they need to be flexible in their work and involve themselves in many different contexts within an organisation.
Connecting the meaning of what applicants and the org say	Almost the entire hiring process is to ensure that the applicant understands the work and the organisation and the different stakeholders within the organisation understand the applicant and their motivation. The HR professional spends a lot of their time on that.
Finding the commonalities between orgs and people, how can you align those parties?	There are many different factors that make an applicant (un)attractive to an organisation and vice versa. Recruiters are in the later stage of hiring often dealmakers or 'guardians of the process' between the two parties.
Leaning on the expertise of colleagues to get to great understanding of the work	Recruiters involve the team manager or colleagues of the open position in hiring for two reasons: first, job specific knowledge goes deeper than what the recruiter can know and they need to rely on others to understand if an applicant truly knows a skill. Second, a recruiter is ultimately successful if the people they find are good fits for the team they will work in. By involving future colleagues of the position, they can better ensure that these colleagues are happy with the final decision.
Finding people with aligned values is harder than finding skilled people	Recruiters generally have less trouble finding skilled people for a certain job than finding people that align with the organisation in how they want to work and their values around work culture.
Understanding the needs of the organisation	Recruiters do a lot of internal focused research on what a team with an open position needs. This is not only skills and team dynamics but also strategic direction of the company. Some positions may need to be filled more quickly than others for example.
Ethical treatment is important and becoming more so	Recruiters are very aware of the ethical implications of their work. They feel there is more focus in society on the social impact of businesses, which makes it even more important to act ethically.
The difficulty of assessing people properly: sometimes tests are not the answer	Assessing someone well is one of the hardest parts of the job. Although tests can help sometimes, they are not a panacea. It means connecting skills to certain tests and making sure that they actually give accurate predictions.
HR works in a demanding and very dynamic environment	HR professionals are often under a lot of pressure: every day a position goes unfilled is a day that person is not contributing to the organisation. However, a bad hire is also costly. They try to cull to a shortlist as quickly as possible while not making superficial judgements. Additionally, for some jobs people are in so high demand that HR professionals need to look for them actively to find good applicants.
Processes get dictated by laws often times	Many processes in HR are influenced or even set up by laws. HR personnel is quite aware of the legal boundaries of their work, even though they can sometimes impede progress on a project.

Table 4. The clusters plus descriptions of the HR professionals. **Continued on the next page.**

Cluster name	Explanation
Being inclusive in finding and hiring people	Recruiters want to make sure that everyone who might be interested in a job is free to apply. They'll try to cast a wide net and find interesting people in unknown places.
New ideas and tools always need to ultimately benefit the work HR is trying to do	The fundamental goals that HR is trying to achieve for the organisation has not changed that much over the years. New approaches and technology come and go, but a lot of it does not 'stick', it does not become an enduring change of the field.
Recruitment does not end at finding the perfect person. You need to make sure they land well in the organisation	Effective onboarding is essential for getting a new employee up to speed. There is a lot of integration between the people doing recruitment and onboarding to make sure that that handover goes as smooth as possible.
People are expensive: each person needs a lot of support to deliver work effectively	Given how expensive it is not only to hire someone but also give them the tools do their work (gear, licenses, space, etc.), HR professionals spend a lot of time communicating about the budget available and how to allocate resources in the most effective way.
Codes not in a cluster	3 codes could not be put into a cluster relevant to the project.

Table 4. Continuation of the previous page.

Cluster name	Explanation
Hiring is a negotiation. Applicants also have leverage and options	Although the employer is usually seen as having more power in the process of hiring, that does not mean that applicants hold no power at all. Applicants have options for leverage in when to start and what benefits they may find more important. Learning on what points the other party may give in is very helpful.
The hiring process is very opaque to the applicant. People may or may not talk to each other	Applicants generally have trouble understanding the hiring process from their side. They may talk to different people in different interviews, expecting the talks to be continuations of each other, but the people they are talking appear not to have discussed much about the candidate. Additionally, during the hiring process you do not get a lot of feedback beyond whether or not you're still a consideration.
Becoming more adept in looking and finding interest jobs comes with the process	Looking for interesting jobs is a skill of its own. The applicants became more adept at understanding what keywords they were looking for, what job boards worked best and how to spot bad positions as they applied to different companies.
Problems with tests can both ingrained in the test or originate from the organisation	Applicants experienced problems with tests in two ways: they felt they did not get (good) feedback from the test, therefore not learning anything about themselves or their standing, or the company may in fact use tests in ways that are not conducive to the hiring process. They may even be obligated to use them by top management while the actual team and relevant HR personnel do not want to use and therefore exclude them of their consideration for the best applicant.

Table 5. The clusters plus descriptions of the applicants. Continued on the next page.

Cluster name	Explanation
Conversation is a two sided play between the two parties with constantly shifting understandings	Doing an interview is for both parties an opportunity not just to get answers to questions they have, but also to gauge the other party, understand what they are after. The interviewee will constantly adjust their answers and questions based on what they think the interviewer wants to hear and adjust their idea of the organisation's motivation for the hire based on the answers. This means that both parties are continuously updating their understanding of what the other party wants and using that further on.
Being not very experienced yet in finding jobs and knowing what you want	The applicant interviewees were just done with their studies. Therefore, they did not have much experience yet in the job market, which also translates to how much experience they have in finding a job and how well they know how to go through a hiring process. One felt it was harder to self-represent yourself if you do not know yet what your work qualities are in the eyes of an employer.
Finding what part of the work you are interested in and aligning that to the wishes of the company	A job description may not be exactly what one wants it to be, but that generally does not stop people from applying. Instead, they use the hiring process to find the elements of the job that they find interesting and in that way align their interests with that what the company is looking for.
Communication takes a lot of effort, and humans can also make a lot of mistakes there	Many people are involved in the hiring process, so there are a lot of moments where information needs to be communicated. However, this is not easy and a lot of issues crop up in the communication, with people blowing a statement out of proportions for example.
Reasons to find a job is not always to find the perfect job immediately	It is very uncommon nowadays to work for one employer in your entire career and applicants are very aware of that. Applicants are sometimes looking for a job as a stepping stone towards a greater plan or they may instead be looking for a short term job while larger life plans come fruition later (e.g. find a job to bridge the time until one starts their master's).
Values are most important for applicants to match with a company	More than getting an exact match on the skillset of a certain, applicants see aligned values as a basis for a job offer as the most valuable outcome of a hiring process.
Getting a complete picture takes significant amounts of time. Many facets that need to be discussed	Even if an applicant knows what they want to say in an interview, the amount of time that sometimes is required to get across the nuance the applicant is looking for can be quite large.
Job matchmaking has become value matchmaking. Skills have become only proof that you can learn what you need for the job	Recruiters tend to not only want to understand your skillset, but also your background and what one finds important in life. These things are seen as indicators for what an applicant may find important in their job (i.e. societal impact, flexible work times or good compensation for overtime). In turn, skills have become somewhat less important because recruiters cannot guarantee a 'fit' with the company just on skills.
Valuable things to want to show to a company when applying	Applicants are quite aware that how they communicate with an organisation is important for how they come across. They will try to show commitment and professionalism in how they engage with the company to look more attractive as a potential employee.
Codes not in a cluster	5 codes could not be put into a cluster relevant to the project.

Table 5. Continuation of the previous page.

Shared clusters

The ideas of HR professionals and the job seekers very much overlapped on how much of the hiring process is concentrated around the idea of connecting the meaning of what applicants and the organisation say. Both groups spent a lot of time trying to find a 'fit of values': for applicants, does this organisation agree with my personal convictions on the actual work, my expectations on what they can offer me and possibilities to grow (vice versa for recruiters)?

The clusters also showed more clearly that hiring is a negotiation where applicants do have leverage: shortening the process of hiring someone can be very alluring to an organisation, so being available quickly can be a decisive factor even if the person does not fit the job description perfectly. To add to that, companies are aware that in reality one will never find 'the perfect candidate'. Each applicant comes with different advantages and drawbacks and applicants can use that knowledge to their advantage.

The HR interviews showed that tests mainly get done for hard skills and abstract reasoning, but the only organisation actually using tests, the design studio inside a conglomerate of Interviewee 3, did not really care about the outcomes of the test according to him. All across the board and from both sides, people were more interested in conversations than in tests. Combined with statements from HR interviewees that hard skills are relatively easiest to understand if there is a match, it shows that job finding has moved from matching hard skills as primary factor to matching values (see Figure 3). Because each job requires different things from the applicant and you cannot learn what a specific company needs exactly beforehand, *asking about skills and testing them in the hiring process become much more of an indicator that you are able to learn a similar skill than it is about the original skill itself*.

Unique clusters

Job seekers only do connecting meaning between themselves and the potential employer. Recruiters do that plus connecting meaning between the different elements of the employer. They need to have a very good understanding of what the different teams within the organisation are doing to make sure they are able to match that understanding of what the work entails with the ideas of the applicants. They note making the conceptual connection does become easier over time, but it is still a significant element where meaning about what a team needs can drift between different people in the organisation. Job seekers do not have to take that into account. This shows an additional factor that AI systems looking to help in recruitment need to handle: the meaning of the work they are testing for should preferable come directly from the people doing the work, not the people managing the recruitment process.

One thing that became very clear from the clustering is that the process for HR professionals is a much longer one than for job seekers. Next to the understanding the work within the organisation, they also spend significant amounts of time making sure the onboarding goes well after someone is hired. Also, they are often discussing the resources an organisation needs to deliver to support an additional employee. In that way, they can really be seen as one interviewee put it: the guardians of the process. In a process where AI systems are at play, that may become their role even more so.

On the side of the job seekers, the process of being hired is seen as very opaque. People do not hear back for weeks and are often left in the dark on how they

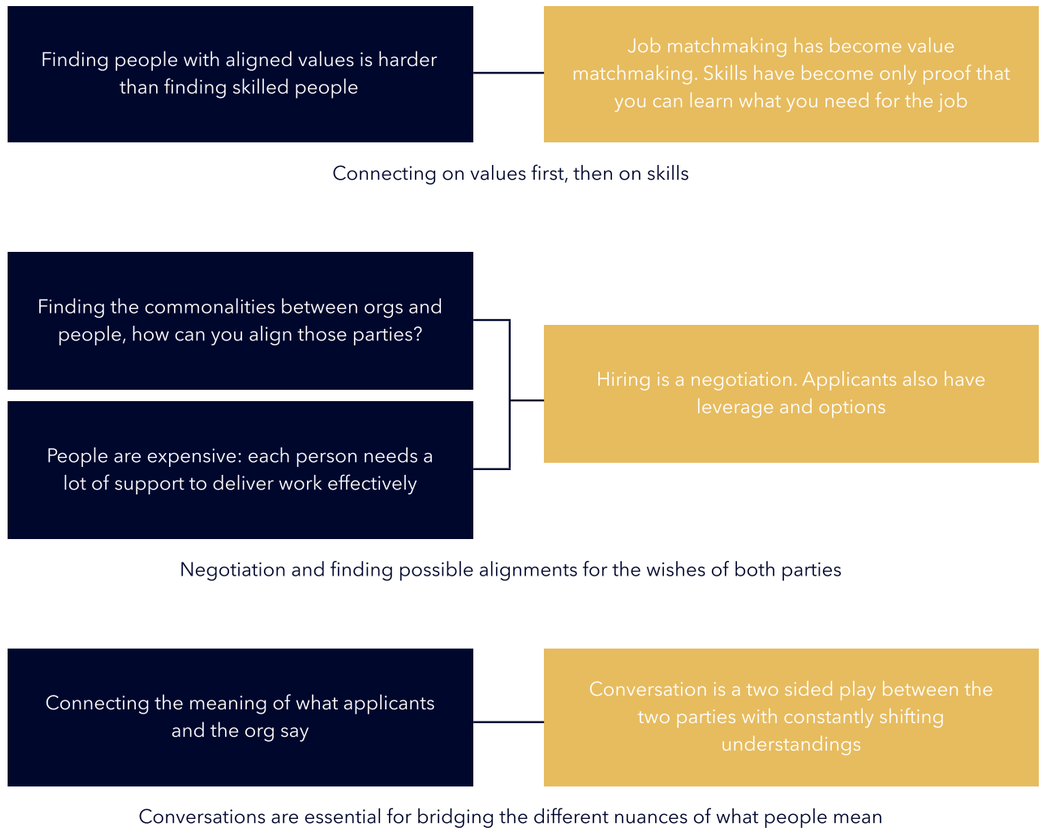


Figure 3. The shared clusters.

performed. This is an advantage for organisations, but people would like to have more feedback on how they perform in general. This may be an opportunity for AI systems to improve the status quo.

Discussion

One of the most striking insights from the interviews was that both applicants and HR professionals were generally positive about self-representation during the process: applicants said they felt generally they could show themselves well during the interviews and HR professionals said that giving that space to really learn about the applicants (and vice versa) is immensely important. This is possibly an indicator that the applicants are happy about their ability to self-represent in a normal situation (without AI).

What is important however to keep in mind here is the sample of interviewee's: there are only four interviewees and both applicants were highly educated, with a Master's and Bachelor's diploma respectively. Likewise the recruiter for the video surveillance company is mainly looking for highly educated software engineers, who are very sought after and have a lot of opportunities. The recruiter emphasised that he has to

chase them and make sure that they feel the company fits them as much as they fit the company. In other words, the labour market is tight. Applicants have a lot of place there to represent themselves, which is why it was not so much an issue in the interviews, as these were all people working in fields with relatively high competition among employers for the good candidates.

That begs the question when self-representation is less valued. Applicants have a vested interest to represent themselves in the way that they see fit, but employers may find it less important if it is easy to find a lot of potential applicants. If one applicant does not like that the application process does not give them many opportunities to represent themselves (i.e. automated tests or talking to a chatbot), the employer can just find someone else. That in turn suggests that the people who are most likely to have limited opportunities for self-representation are applying mainly to companies that are doing high-volume hiring where the individual difference between applicants is not very big (possibly when a job has less formalised or unique requirements).

Another interesting insight is the focus that both HR professionals and applicants had on aligning values. The HR professionals shared that hard skills are the easiest to check. The hiring experience from both the HR and applicant side show that hard and soft skills are still part of it, but neither party focused much of their time on these skills during interviews. Instead, skills were tested more than talked about (if they were focused on), while both parties saw values as more suitable to talk about in actual interviews. This makes sense, as employees aligned to the strategic values of the organisation are more able to put all their productivity towards to the organisation's goals as the literature on strategic alignment has shown (Oehlhorn et al., 2020; Sahoo et al., 2011). That means that getting these values right is more important than checking off all skills for a position.

Because tests are easier to do for hard skills while the market seems to move away from hard skills as primary indicator of a good applicant, this creates a paradox: we're moving towards more AI-infused measurements, while the very things that become more important to check are inherently less easy to check with a test involving AI (like on hard skills). Instead shared values and common understanding of the meaning of the work are the primary indicators of a match, which are concepts that AI systems can have trouble with evaluating reliably.

Data insight

Both applicant and employer have moved from hard and soft skills as prime indicators for a good match to values on how the work should be done. Elements like matching ideas on work ethic, societal contribution and work-life balance are examples of things that the parties want to focus on.

Additionally, both HR and job seekers see the negotiation with the hiring process as constant realignment on values between the two parties. This is where the value of conversations are very apparent for both parties, as humans do not do negotiation often without conversation. What important to note here is that every interaction between the parties is seen as a way to understand the other better. Each conversation is for job seekers helpful to understand the organisation better, also because they have the space to ask questions themselves. As most AI systems are predefined questions that get asked to the job seekers, job seekers lose a lot of opportunity for gathering context and better understanding, but it may inadvertently also hurt organisations as they may end up with employees who do not feel they actually fit the organisation.

Data insight

The hiring process is seen by both parties as a negotiation where each interaction helps to understand what the other party finds valuable.

Maps

The empirical investigation of the interviews gave a coherent and nuanced overview of the values that the different stakeholders hold. To bring these insights back into an empirical investigation for further iteration, I made an ecosystem map and journey map to better understand the interactions and processes of recruitment. These maps help with getting a more systemic overview of the context and show why certain stakeholders act in the way they do.

Ecosystem map

Based on the previous investigations, I made an ecosystem map (da Costa Junior et al., 2019; Forlizzi, 2013; Heikkilä & Kuivaniemi, 2012) showing all the different value flows between different entities in a system where a recruiter uses testing tools to do part of the selection during the process (see Figure 4).

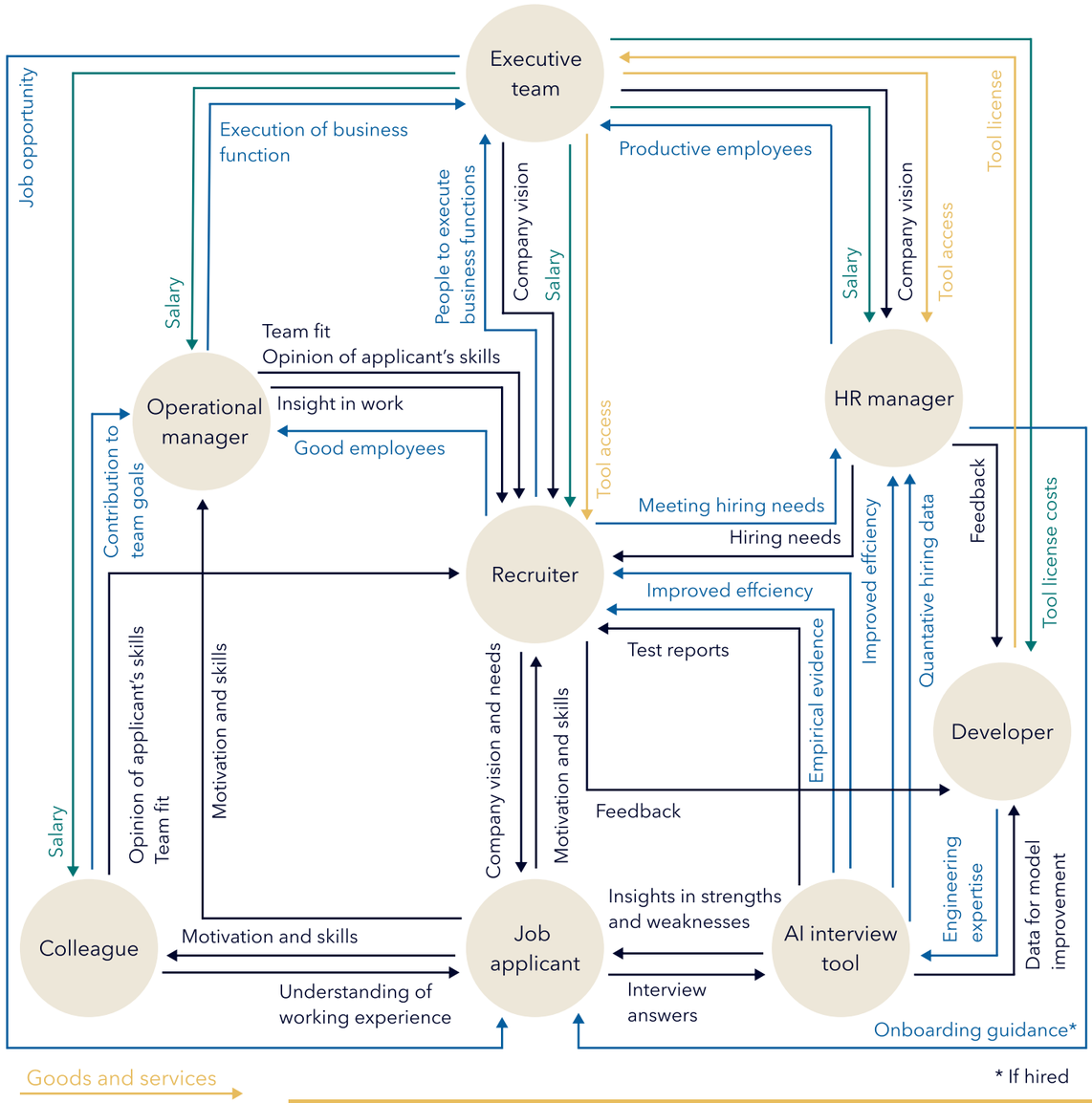


Figure 4. An ecosystem map of the hiring process involving an AI system. Depending on the AI hiring tool, different information value streams to and from the tool may exist. This ecosystem is based on an AI tool that mainly does automated video interviews and reporting, similar to HireVue and Harver.

The map shows that the recruiter is more or less the central point in the map, with the most in- and outflows of value going to/from them. As they are the manager of the process, this makes sense, but it also shows how the process is geared to optimise for the recruiter, not the job applicant.

At the same time, job applicants (in a good process) get as much exposure to as many people as possible during the application, in order to have more people get an opinion on the applicant. Applicants will often talk to recruiters, team managers and colleagues of the to-be-filled position, HR managers, and, if it is a senior position, upper management of the company. In other words, many people are involved in recruitment to give their point of view on whether a applicant is a good fit for a certain position, while also giving the applicant a perspective on the many facets of the job context. However, the AI tool is generally not equipped to facilitate all these interactions.

If we zoom in to the value interactions of the AI hiring tool, it is quite striking how large the imbalance is between in- and outflows. The tool is supposed to deliver many intangible value streams, while its only information value in-stream is interview answers. It is then required to deliver reports on those interview answers, but also deliver 'empirical' evidence for the decisions a recruiter makes (backing from a tool that a person is a good hire) and provide the hiring manager with an integrated, quantitative overview of all the applicants. It highlights how much depends on a system which is not getting that much context fed into it.

Journey map

Next, I made a journey map (Howard, 2014; Stickdorn, 2018) showing how a recruiter and an applicant deal with the different steps of filling an entry level developer position. The recruiter uses an AI test to check soft skills. Harver is an example of a HR tool that has many types of hard and soft skills assessments and was used as a basis for the journey map. Between the two journeys, there is a space that highlights the ways an AI system can impact the recruitment process. See appendix C for the journey map.

While making the recruiter journey, I noticed that her journey did not involve many problems regarding autonomy over self-representation. Because she is simply viewing the information as presented by the AI system, rather than the applicant's direct input, there is no inherent contradiction or misalignment in what she sees. Meanwhile, the applicant does not see what the AI system presents to the recruiter. This shows why self-representation is hard to ensure in a hiring context involving AI systems, because the human parties generally do not see what the other party sees or hears and do therefore not compare input and output of the AI system. There is no feedback loop between the human parties where one can intervene when they notice something being understood in the wrong manner.

What is also noticeable is the amount of times an AI system is involved in the process. It is not one involvement, but multiple with different purposes. Particularly ranking happens many times, depending on how many assessments an applicant goes through. All these individual interactions are moments when meaning of what one values in a job or how one solved a particular problem previously can be distorted by an AI system.

Discussion

The ecosystem map is very useful for understanding that while the AI tool is often a significant part of the total amount of interactions, it is only getting data from a small part of the different interactions that can happen.

Additionally, the sheer amount of value streams of recruiters in the hiring context show how many different opinions and wishes the recruiter needs to take into account. Combined with the large amount of applicants usually applying to each position, it shows why HR departments are so willing to use tools that promise to help them with the onslaught of information. This is a strong indicator that AI systems will not go away, as the benefits to HR departments are too great.

The journey map highlighted how the two parties can misinterpret each other's messages because an AI is interpreting the information for them. Particularly information towards the recruiter is filtered and repackaged. And while most information like actual videos of the applicants is usually available for recruiters to see, automation bias (Skitka et al., 1999) and the constant pressure to deliver as shown by the interviews ensure that this happens very rarely, as people trust their tools generally. As the quote from the interviews below shows, HR personnel expects their tools to work reliably and without invisible biases:

> *"If we take an AI tool to help me to find candidates, I am assuming that the database is equal [aka the dataset is representative of reality]"* - Interviewee 2

However, it is important to note that all the work done for these maps is largely theoretical. It was based on data from interviews and papers, but no first hand information from people who actually worked with these kinds of systems. This severely impacted my ability to understand better how that impacted people's values during the process.

Conclusion

The ecosystem and journey map have shown how the nature of information streams that go through the AI system in the entire hiring process can lead to misaligned meaning between the two groups and no direct way of seeing the other party's view, which is hindering the possibility of someone stepping in and saying that some piece of information is not correctly transferred.

However, a important realisation came up in the discussion. As my empirical research so far has only involved people who have not interacted with AI systems in the hiring process yet, I am still lacking data on how people *actually react* in situations of AI systems working in a hiring interview context. This interaction is very important for understanding the values that people find important in these situations. So, in the next chapter, I will describe how I made these situations and elicited these values through generative prototypes.

Generative prototypes

VSD puts a significant emphasis on empirical investigations of the context, to understand properly what people value in these contexts. In order to design for what applicants value in self-representation within the hiring context, I used generative prototypes to elicit this tacit information.

Setup

The prototype procedure consisted of five steps and were meant to evoke the situation of doing a video interview with an AI system for a job that each participant would be interested in. To that end, each instance of the prototype was individualised.

1. 1. After recruitment of the participant, they would be asked what kind of jobs they are interested in and to give a recent CV, in order to tailor the experience to their interests.

2. 2. The participant gets a set of tailored questions that evoke doing a video interview with an AI system. This interview incorporates different ways a video interview with an AI system can limit their self-representation.

3. 3. Straight after, they are given a short semi-structured interview to understand their experience of going through the video interview.

4. 4. The day after, they are sent a short report on their performance in the video interview, again with different ways to provoke their reactions based on how an AI system might interpret their responses.

5. 5. They are again afterwards given a short semi-structured interview to explain their experiences.

The video interview is done on VideoAsk, a tool that reimagines the online questionnaire by basing it on short video interactions. One records a question and people can answer using another video, audio or text (see Figure 5). Although

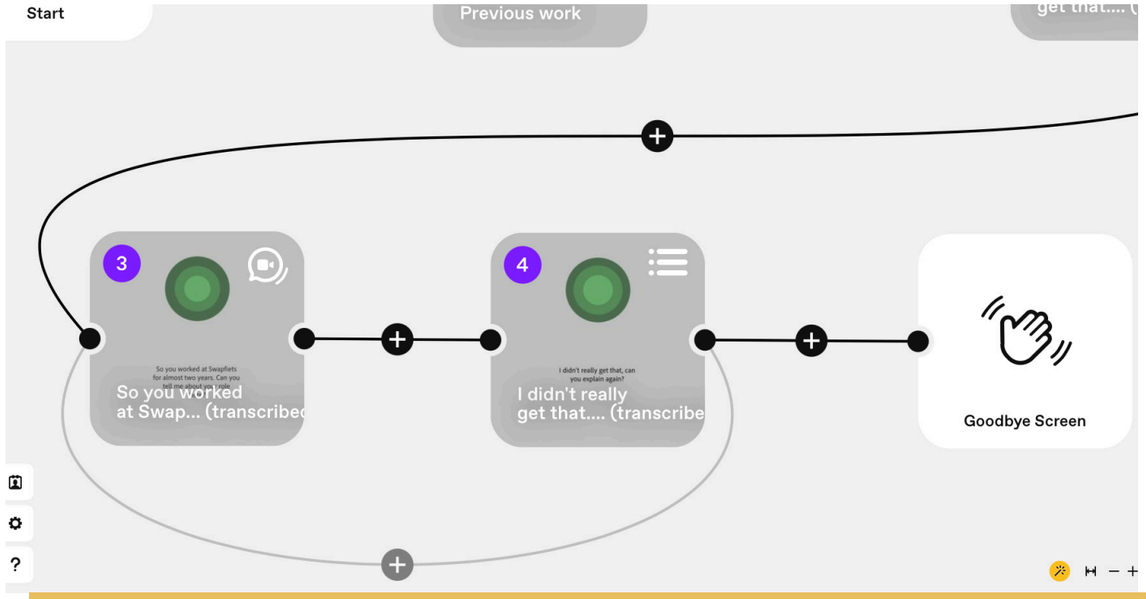


Figure 5. The interface to build a VideoAsk form. Each block represents one 'interaction moment' with a video and an opportunity for a respondent to reply using video, audio, text and buttons, depending on what is appropriate for the question. Buttons allow for conditional logic, such as loops like in the example above. For the prototype people were restricted to only give video responses.

Situation name	Explanation
One question	Participants did a 'full' interview with introduction, one question, and closing of the interview. Participants were specifically reminded of the fact that the system had enough information to decide whether they got the job based on the single answer and the video information they sent. The question was individualised to account for different jobs that participants were interested in.
CSR	Participants were asked on their position on Corporate Social Responsibility and notified afterwards that in order to limit the bias of the recruiters, both their voice and facial characteristics would be modified.
Background	Participants were asked multiple questions that would require you to take some time to properly explain and then only given originally 20 seconds to give their entire answer. Ultimately this time was lengthened to 60 seconds in order to collect more information that could be used in the report stage of the prototype experiment.
Previous experience	Participants were asked to explain one of their previous experiences as listed in their CV, but the system would act as if it did not get the response and ask for a retry of the answer. This would get participants in an eternal loop, or they could decide to forgo the question, removing the entire answer from consideration.

Table 6. The four different situations the participants were confronted with, including each provocation.

positioned as a tool for better human interaction, it can also be used to create a Wizard of Oz experience (Dahlbäck et al., 1993) where a human simulates the advanced technology (in this case, the AI analysis).

Four interview situations were written (see Figure 6), run through different digital voice generators and then combined with a visual representation made in Adobe After Effects resembling the on-screen presence of a voice assistant like Siri or Alexa (see Table 6 for an explanation of each situation). The different generated voices and visual representation were done to reinforce the idea that people were talking to an AI system instead of a human. Additionally, each situation was made to highlight one aspect of how a video interview may limit the self-representation of the participant, while pushing further than what people would see in the real world in order to intensify the reaction of people and get them to think more deeply about the situations.
The prototype were tested with seven different participants, all students of higher

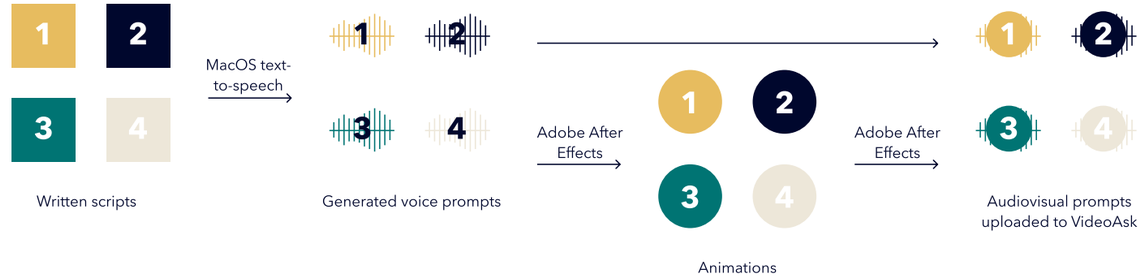


Figure 6. The VideoAsk creation pipeline. Because some situations were custom per person, they had to be done again for each participant.

education or recent graduates (see Table 7) and between the age of 24 and 27. Convenience sampling was used due to time pressure and the fact that all participants were still part of the overall labour market (also see Limitations).
At this point, the use of ASF becomes particularly clear, as that framework focuses on

Participant	Gender	Field	Preferred job
1	Female	Design	Innovation team lead
2	Male	Design	Researcher
3	Male	Design	Service designer
4	Male	Design	Technology designer
5	Female	Design	Digital designer
6	Male	Design	Social service designer
7	Female	Public administration	Policy advisor

Table 7. The participants of the expirements with the generative prototypes.

the use of generative prototypes to create hypotheses. In Table 8, you can see the first and second iteration of the framework. They are presented together below, because the second iteration happened during testing of iteration one. I did go through the iterative loop twice as the first test of iteration one showed me already clear enough how to improve on it and I would be able to gather much more information by adding the report of iteration two. Iteration 0.1 refers to the data gathering in the project before the introduction of the ASF framework.

Hi

CNCF consulting

With this report, we want to give you a little insight in your interview performance. Overall, your fit for the position was determined to be: 63%

A couple of things from your interview that we found very informative:

You came across as eloquent and thoughtful but somewhat distracted. You looked positive 44% of the time during the interview.

It was unclear what you were comparing the Research Institute of Sweden with in your background.

I've also worked as a design researcher at Rice, which is a Research Institute of Sweden, which is comparable to the Dutch. They know, I worked at a very small Department than, you know.

unknown emotion

You mentioned service designers are doing unneeded tasks

so I think being in service designer is a very interesting, but unneeded tasks, especially in the future,where our whole economy is becoming more and more service oriented

We would like to hear what you think service designers designers should be doing instead.

Unfortunately, we were not able to transcribe one of your answers on the question of what you did at Umeå, which is why we have omitted that information in your assessment.

Your Profile

From your video info plus your answers, we made a model of how you felt to us during the interview. The dot indicates what we are looking for.

	Very clear	Clear	Moderate	Clear	Very clear	
Focus your attention and energy on the outer world of people, activity and things.	E	●				I Internally focus your attention and energy on personal ideas, experience and emotions.
Take notice of your senses and what is real and tangible in the present day.	S				●	N Focus on patterns, links, connections, the big picture and future possibilities.
Make decisions using logic, facts and objective thought.	T		●			F Make decisions through feelings, values and through consideration of others
Having a plan, being organised and prepared for events	J	●				P Like to stay flexible and respond to changes as they happen

The report given to each participant was created using their answers given in the videos. The report had many different ways to diminish its value to participants and to show how the 'algorithm' had subverted or misconstrued their words and actions (see Figure 7 for an example). It included a percentage score to indicate how the participant did, but with no explanation how it was calculated. Quotes were pulled out of context or just erroneously transcribed, making the meaning of the sentences completely different. It highlighted that one answer was not received well and therefore removed from the analysis. It described how the applicant came over in a few words but nothing was actionable. It included a Meyers-Briggs personality profile, 'based on' video input and answers, but there was no real analysis or explanation, no in-depth information on how people could use this information.

Figure 7. An anonymised report. The transcription has obvious spelling errors (Rice → RISE) and participant pronounced TNO by spelling the letters in Dutch, making the transcription turn it into 'They know'. The image was chosen to be specifically unflattering.

Iteration	Element	Explanation
0.1	Purpose	Design a strategy for HR departments to work ethically with AI systems in hiring
	Data	Academic and news articles, interviews, journey and ecosystem maps
	Frame	I can create a better solution for autonomy over self-representation in hiring if I use generative prototypes to specifically elicit tacit values from stakeholders
	Purpose	Elicit tacit values in hiring interviews by provoking participants
1	Purpose	Elicit tacit values of applicants in the hiring interview process by provoking participants
	Data	Academic articles on generative prototypes, interview data and insights from maps
	Frame	If I create generative prototypes that simulate an AI interview and that are tailored to make people interested, participants will be provoked
	Mode of action	Let participants answers specific interviewing questions that limit their autonomy by restricting time or amount of input, changing their appearance or pretending not to understand their answer
	Intended actuation	Participants record answers to all four interview situations
	Prototype	An interviewing platform with tailored questions to collect information on participants
	Observed actuation	Participants recorded answers to all four interview situations. E.g., complaining about the lack of reciprocity in the conversation, not having a way to know what the other hiring party is looking for
	Values	E.g., confusion on how participant's data will be used (participants noted they had very little idea of how the AI system was interpreting their answers, but it was for them simply invisible, not provoking). Also participants did not enjoy the interview because they had received very little information from the AI system, they could not ask questions back for example (conversation reciprocity).
2	Data	The collected data during the previous iteration, interviews and elicited values
	Frame	The participants will be more provoked if I show them a report after their 'interview' to show how the AI interpreted their answers
	Mode of action	Create personalised overview of the participant's (non)acceptance that specifically misrepresent how they went through the interview
	Intended actuation	Participants would check the information and and get provoked by it. In turn, they articulate the tacit values of what they value in the context of hiring
	Prototype	Give participants an AI generated report with information on why they were (not) accepted for the job based on their answers, but the system misrepresented their answers by wrong transcriptions, nonsensical descriptions, unflattering screen captures and vague personality models.
	Observed actuation	E.g., lamenting the lack of actionable information, incredulity and confusion on how their answers were interpreted and how that affected their standing, rationalisation of the entire report if they could recognise a small part of the analysis, anger at wrong transcriptions or questions that were perceived not to be asked
	Values	E.g., missing self-representation by not understanding the mechanism of how their information were used, confusion of vague statements ("'You came across as superficial yet engaging.' That would stay in my mind for the whole day.") Getting dispirited by the feedback because it made them feel they did a bad job.

Table 8. The several iterations done up to this point in the project.

Results

All interviews were coded and structured in clusters. The complete clustering and interview codes can be found in appendix D, with an overview in Table 9 and a discussion below.

Overall, both the interview and the report succeeded at provoking the participants. All participants felt unease at the interviewing situation with the lack of reciprocity in the conversation, the lack of depth of the questions and the inability to steer the conversation as points that came back in almost every interview. One complaint that was often noted was that participants disliked the fact that before you started a response you could not see how much time you had for it. This is a quirk from VideoAsk and was not done intentionally.

Another surprising factor that people mentioned was the comfort that a human interviewer can give you. As you may get uncomfortable during an interview, another human is able to see that and react accordingly, but that was missing in these AI interviews.

Regarding the report, a large majority of participants lamented the lack of actionable information and strongly preferred interviews with humans versus getting a feedback report after the fact. This hints at the relative importance for the participants of these conversations.

What was interesting to see was that the participants did not get particularly provoked during the video interview because they did not know how to react. They were confused, but at that point they did not see yet what the impact was of the AI system. Unfamiliarity with the situation, AI systems, and job interviews (as the participants were all still students or only recent graduates) were all mentioned as causes of this.

Cluster name	Explanation
Questions being of interest to the applicant	Questions can be more or less interesting to an applicant to answer. An applicant will be happier to answer a question if that also helps them learn something about the company.
Holding ownership over the direction of the conversation	People want to have some level of control over the conversation they are having. If they have no control, it turns into an interrogation. It helps them to focus on concepts that are of particular interest to them and informs them too on where to draw the line on how explaining is necessary on a certain subject.
Knowing what the other wants to hear	Many participants noted that a conversation is a great for them too to understand the other party. Through answering and asking questions, they get a better feel of what the other party wants to hear, which in turn helps them.
Meaning is hard to convey with AI	Participants noted many different ways that the AI system made it harder for them to communicate the meaning they wanted to convey. The meaning of their words may get lost in the transcription, in an interpretation that actually reverses the meaning and how they can judged on many independent elements instead of as an holistic being.
Opportunities for AI conversations	More customisation of the standard interview was seen as a positive aspect of doing an AI interview. However, one participant did not understand why, if the questions are already tailored to them, they were still talking to an AI system instead of a human being.
Issues during an AI interview	AI interviews give applicants many different problems. E.g. inability to get into depth about a subject, inability to 'steer' the conversation, not knowing how your data is processed, no way to correct errors and potentially being less familiar with a language (one participant had not spoken English for over a year and had trouble to adjust herself).
Being comforted is important to people during interviews	Many participants noted that they did not feel comfortable during the AI interview. In a usual interview, the interviewer can see if you're nervous or anxious and adjust their manner of communication to comfort the applicant. The AI system does not do that.
Moments when to use tests and AI systems	Two participants said they would be fine with having an AI system do the first interview, as long as afterwards they can talk to a human. However, it would only be appropriate in a long hiring procedure, as they felt it would not add much value for the company if the hiring procedure is not many steps.
People want conversations, they know better how to deal with that	The AI interview is more akin to a Q&A session than a real conversation. The need for a real conversation was brought up many times as participants felt that conversations were easier and more valuable for them (and also the interviewer) on creating deeper understanding.
Getting feedback vs having conversation with real human	When asked, all participants noted that the after-interview report was not as valuable to them as having a normal interview for their understanding of the job. Still, most did say that they did like getting more feedback on how they did and what they could improve.

Table 9. The clusters and their explanations of the generative prototypes. Continued on the next page.

Discussion

Limits to self-representation were seen as missing ownership of the conversation, not knowing what the other party wants to hear and the idea that meaning is hard to convey with an AI system. These three clusters fit well with Hamilton & Davison (2018), which shows that the lack of control over the process is connected to how negatively people experience the process.

Many participants (5/7) brought up the concept that conveying meaning and their personal values was particularly hard in a AI interview. One particularly frequent complaint was that the AI system gives very little in the way of feedback and other information on what the organisation wants to hear. This lack in conversation reciprocity was a key factor that put the participants in a disadvantaged position.

Cluster name	Explanation
Issues with the report	The report gave applicants many different problems. E.g. information that was not actionable, seeing a answer that you know you did not give is very frustrating, the amount of feedback was effectively very little, missing references to questions made conflation of questions possible, which added to the confusion of the report.
All focus on the applicant for better and worse	Two participants noted the fact that you can see yourself. One liked it because they could see how they would come across (technically not true because they reasoned from a human perspective and an AI system does not judge like a human). The other did not like it, calling it distracting and likening it seeing yourself in the corner of your screen like during a Zoom call.
A job application contains a massive amount of context, which steers everything you do	The prototype was tested without much context to improve the provocation. However, many participants noted that usually, you already get a lot of information from the job posting, the website of the organisation before you've have ever talked to them, which in the case of the prototype made it harder for them to give a convincing answer because they knew practically nothing about the job they were applying for.
Value of AI interviews for companies	Applicants generally have a good idea of what the value is for companies in using AI interviews. Two mentioned the need to filter out applicants, while another noted the bias of interviews and how companies want to remove that.
Humans remember through coherency, AI cannot	The participants always pointed out errors in the report through cohesion. They knew that it does not make sense if they argue two opposite points in the course of a single sentence. Meanwhile, the AI does not look at the meaning of the sentence but only what the individual words meant and constructed its own meaning out of that, which is why the reports were so incoherent.
Seeing patterns in feedback that are not based on correct data	Many participants would, once they found a piece of information that they could agree with, reason from there to try to understand the complete report, even seeing patterns that did not exist in the real data.
Codes not in a cluster	3 codes could not be placed in a cluster relevant to the project.

Table 9. Continuation from the previous page.

Data insight

The lack of conversation reciprocity and nonverbal communication was highly problematic for the participants. Not having a way to ask questions highly diminished the freedom the participant had while not having someone to 'read' gives them very little feedback on how their answer is being interpreted, which makes it harder to adjust the following answers.

Two people mentioned the fact that your own face is front and centre during the video interview. One hated it, as it was a distraction for her and made her uncomfortable, while the other liked it because it gave him an additional way to see how he would come over to the other party. This shows that people can have very differing priorities regarding self-representation during an interview: some may want to minimise their discomfort in order to come across as more confident, while others prefer seeing as much information as possible in order to get a better idea how the other party might see something.

It was also clear that the report was essential in the prototype. While the interviews made people somewhat uncomfortable and unsure how to proceed, the reports truly provoked the participants, as they saw how 'the AI' had mangled their answers or gave them nonsensical feedback. Most reacted sad or angry as it became clear how they would ultimately come across to someone else.

Data insight

The report, as a way to bring back information to people on how they were perceived, was very important in sparking the provocation. Without it, an AI interview feels like talking to a wall with very little information being released to the applicant.

In connection to the reactions to the report, there was another interesting and potentially dangerous phenomenon. As participants actively tried to make sense of the report, many were rationalising the data they were seeing, as some data was recognisable and plausible from their own perspective. This went to the point that one participant felt she could somewhat see herself in the report, even though the entire report was specifically made to be nonsense. This shows that the confirmation bias and pattern recognition skills of humans can work against them, as they will try to rationalise and justify data that can be simply wrong. So even if something is represented well, it can still happen that one will gloss over it due to them reasoning it away.

Overall, the generative prototypes were very effective at eliciting reactions and value statements from people. For more esoteric values such as self-representation in the context of hiring, generative prototypes can be a valuable method to understand stakeholders.

Designing for self-representation

At this point of the project, it was time to combine the information gathered in the conceptual, empirical and technical investigations: synthesis. This contextualising of the different insights found is already framing the problem. So, through constructing a new design frame, I can find the new purpose and intended actuation that will lead to a better solution.

First, the many problems that the participants had with the AI interview and report stem mainly from two elements, lack of information during the interview on how they are doing and how to adjust and their inability to steer the conversation and ask their own questions. The problems have much to do with nonverbal communication and transferring meaning through conversation which are hard AI problems for which there are no good technical solutions yet. This means that if one wants to ensure autonomy over self-representation would not be harmed in any way, *it is best to currently not use AI systems in hiring*.

However, second, from the literature study and interviews with HR personnel, I learned that HR is under a lot of pressure to find good candidates quickly and with little mistakes. People are nowadays very valuable for companies and digital job boards have caused massive amounts of people applying to each position. Therefore, the benefits that AI systems can have in this context mean it is very unlikely that they are ever going away.

Additionally, third, perfect autonomy over self-representation does not exist and has never existed. If one has complete autonomy over how they self-represent, that means having control over how someone else interprets your representation. However, never have people had direct control over what others thought of them, also not before AI systems were used in hiring. It has always been a balance between how much one can show and how much room an other party has on interpreting that 'showing'. Therefore it is this balance between the two parties that needs to be maintained.

Fourth, from the interviews, both applicants and recruiters have moved from looking more for a fit of skills to a fit of values. AI systems have a lot of trouble ensuring this value match happens as seen in the generative prototypes. This means that AI systems (at least for the foreseeable future) will not be able to completely replace regular interviews in the hiring process, even if the communication problems applicants have with them are solved.

These insights led me to the creation of a new purpose and frame of how AI systems should maintain autonomy over self-representation for applicants in hiring.

Purpose

The AI system should shift the power imbalance between applicant and recruiter towards more power to the applicant to make sure that they maintain (partial) autonomy over self-representation.

Insights

Autonomy over self-representation is always imcomplete. Otherwise, interpretation by another party would be impossible.

Applicants don't gain needed information from AI interviews on how they are doing and they cannot steer the conversation well.

Both parties are looking towards value fit over skills fit. Because assessing values is hard for AI systems, human interviews should remain.

AI systems are not going away. Therefore, we must look at minimising bad impacts while maintaining efficiency gains.

Maintaining autonomy over self-representation means maintaining the power balance between applicant and recruiter

More information needs to go applicants so they know how to better represent themselves

Outcome

Applicants need a better feedback loop on how their message gets interpreted by the AI system. This way the applicant understands better how their words are seen by the other party and can adjust their representation to fit what they want show

Applicants need to get the possibility to ask questions and steer the direction of the conversation more. Without the possibilty to redirect focus of the conversation, they are stuck answering the questions a hiring manager came up with, instead of highlighting their strengths in a way that works for them

Figure 8. The solution flow based on the four primary insights. The outcome shows the first steps of a future mode of action already.

If autonomy over self-representation is a balancing act, then an AI system should fix an imbalance it caused. The remedy does not have to be the original way that the AI system caused the imbalance: the AI can give power back to the applicant in a different way. Figure 8 shows how this thinking can lead to giving applicants more information and ways to act on that information as a way to maintain the power balance between applicant and employer. This would require 'radical transparency' where applicants get to see more of how they are viewed by the AI system than a recruiter would show normally in order to be able to make better decisions on how to show themselves further. The frame that fits this purpose is as follows:

Frame

Both applicant and employer have moved from hard and soft skills as prime indicators for a good match to values on how the work should be done. Elements like matching ideas on work ethic, societal contribution and work-life balance are examples of things that the parties want to focus on.

This frame led to the following process (see Figure 9 on the following page) where the applicant would get to see the interpretation of the AI system of the interview that was just done and adjust where the meaning was not conveyed well before the AI system uses the data to score and grade the applicant. The adjustments could be used to let the system learn better representations of what people mean while applicants get a better overview and more control over how their information gets interpreted. Then, regular interviews still happen with a recruiter to make sure that the value match actually happens. Afterwards, recruiters can compare their opinions of the applicants with what the AI system recommended as an additional method for feeding back data to the AI system to improve. The mode of action is thus as follows:

Mode of action

The AI system shows its interpretations of the answers of the applicants to them, in order to get feedback on what the applicant meant and how good its interpretations were. The feedback gets incorporated in the grading of the applicants and used for further improvement of the system.

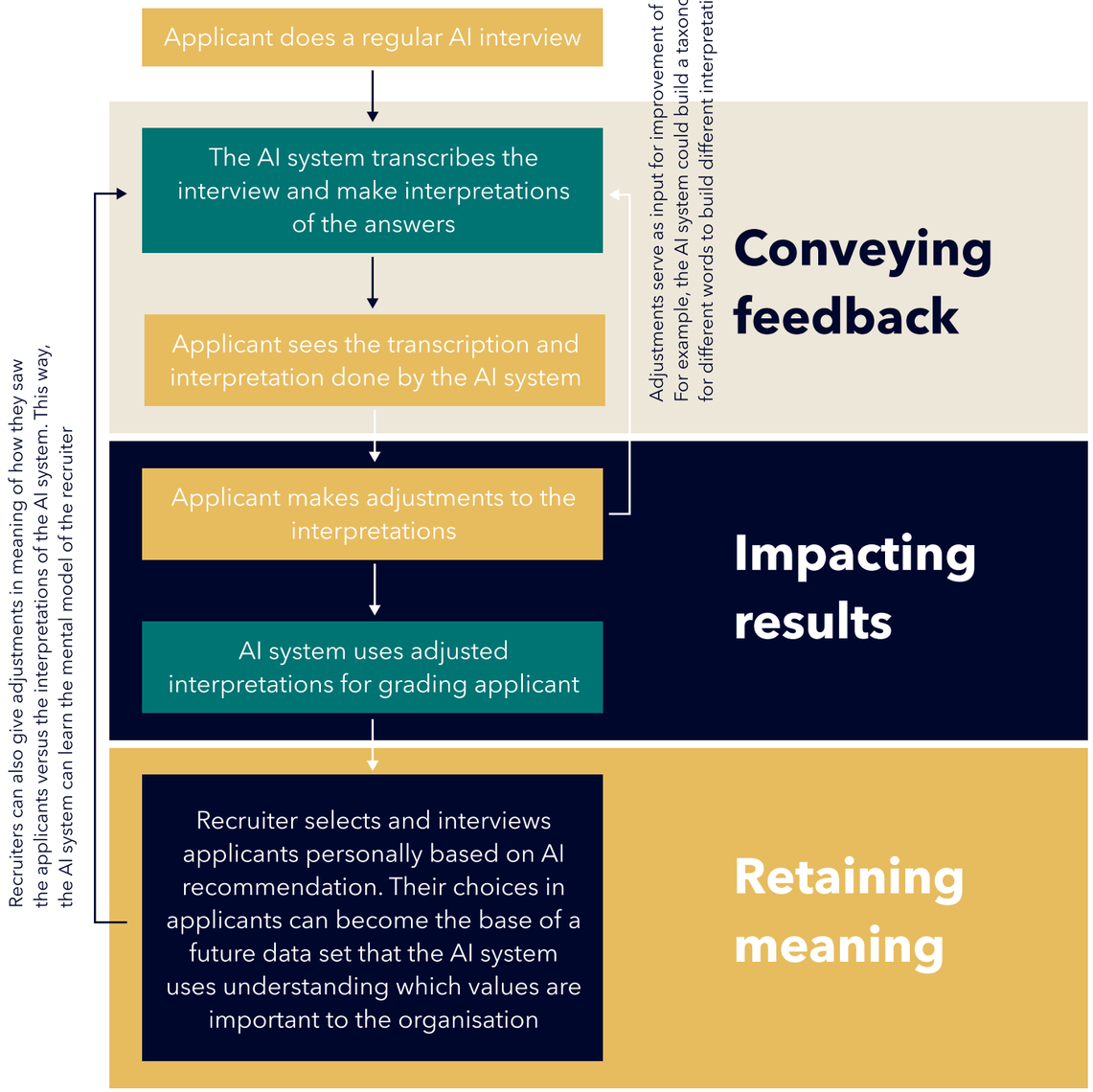


Figure 9. The process workflow for a future AI hiring interview system. The extra steps ensure that feedback about their performance is relayed back to the applicant, that they have more concrete ways of steering the meaning of what they want to say, and that the true meaning of their answers are still retained through regular interviews. Simultaneously, both applicants and recruiters have an opportunity to judge the interpretations of the AI system.

This process has certain assumptions connected to it. Without these assumptions, the entire system cannot work.

- The AI system should only interpret words themselves, not intonation or expressions for example. Otherwise the system will combine different elements and produce an interpretation that is partially based on intonation and expressions where it is very hard to understand for humans how an AI system came to its conclusions.
- Recruiters only uses AI systems for screening. Afterwards, you will always still need interviews as necessary to validate the interpretations of an AI system, particularly because the value fit people and companies are looking for is particularly hard to do for AI systems still.
- The difference in meaning can be described in a loss function. A loss function (or error function) is a fundamental element of AI systems that describes the mathematical relation between the output of an AI system and its desired target (Raschka & Mirjalili, 2019). It give the system a concrete number to optimise for as it shows the 'cost' of using a specific model, which it can compare to different iterations and select the best performing model. An example for the context of this project could be asking participants "How close is the meaning of the interpretation?" which they can answer as "Very close [1], close [2], far [3] or very far [4]". By representing the 'closeness' of the meaning of one's answer as a number, an AI system can optimise its model by lowering the discrepancy over many iterations.

Conclusion

In this chapter, I have combined the different insights gathered from all the investigations and used them to create a new frame on how to view the situation. I have sketched out how the process should look. It should focus on conveying feedback to applicants, giving them options for impacting the results of the interview and retaining meaning through doing regular interviews. Finally, I have outlined the assumptions that need to be true in order for this system to work.

Towards a better situation

To evaluate this iteration, I created a generative prototype that looks a lot like the prototypes discussed before. However, the key aspect of provocation is not a part of this iteration. Because I am now focusing on finding solutions instead of exploring the problem, I am now more interested in finding the issues with the envisioned purpose, frame and mode of action which requires a more realistic approach than provoking people with an absurd situation.

Setup

The prototype was set up as follows:

1. The participant was sent a job description for an entry level design position to give some context to the participants on what kind of company and job they were applying for.

2. The participant first does a short video interview in VideoAsk, similar to that of the provotypes. This time however, the situations are not meant to provoke, but to simulate typical questions for an AI interview (see Table 10). The questions were made to require the participants to give answers that were rich in context and nuance, as opposed to questions that are more based on facts (e.g. "How long did you work at your previous employer?).

3. Next, their answers are taken from the transcript made by VideoAsk and put with a plausible interpretation in a Google Forms form. Participants could see how the AI system transcribed their answers and how it was interpreted by how the system rephrased their words. They could adjust the text of both by changing the text accordingly.

4. Then, a short semi structured interview was done to understand their experience with the combination of the AI interview with feedback mechanism.

Question
Can you tell me how you would handle a disagreement with someone on your team?
How would other people describe your work ethic?
Can you tell me about a time when you had to make a decision without having all the information necessary?

Table 10. The questions posed to the participants, inspired by a document on typical HireVue questions (Duke University, 2018)

The iteration was evaluated with three people (see Table 11), all design students between 25 and 26, two of which were also involved with evaluating the earlier iterations. To save time, the questions were not customised to each participant as in the previous iterations.

Participant	Gender	Involved before
1	Male	Yes
2	Male	Yes
3	Female	No

Table 11. The participants of this iteration

Results

Like with the previous iterations, the interviews were coded and clustered. After an explanation of each cluster in Table 12, I will describe the relations between the clusters. The complete clustering and related interview coding can be found in appendix E.

Discussion

With this iteration, it became quite clear that while the idea of delivering feedback and giving people control over the interpretation of the AI system was a positive value, the current implementation did not appeal to the participants. The change of modality from spoken language accompanied by body language in the video responses to written language in a transcription caused all participants to see the text as lacking a significant part of the nuance they were trying to convey. However, the participants found it very hard to say what exactly was missing. One participant commented on his own words as follows:

*"So if you read that, that nonsense, because it is nonsense, it is like 'Oh my god, this guy is f*cking crazy', right?"*

Cluster name	Explanation
People express differently through text or verbally	Multiple participants noted after seeing their own words in the review that they do not express themselves in the same way verbally versus in writing. Basically, the semiotic modality has an impact how well different types of information can be conveyed. For example, the modality of writing allows for easier refining a concept because you can easier change what you have written before. Once you have spoken a sentence, you cannot go back and change it, you can only add more sentences to change the meaning of what you are saying. This meant for the participants that they saw fundamental differences between what they meant and what could be inferred from the transcript.
What writing language assessments add to the interview process	Writing in the hiring process is particularly useful to see one's skill in cohesively and succinctly articulating one's point. It allows one to highlight their precision in their thinking.
On what topics should interviews focus?	Conversely interviews are more focused on thinking on your feet. One can show their comprehension about a subject by giving good answers to questions, parrying arguments as they come up.
You always need to work together with others	Conversational and social skills are very important when people are working together. Participants felt this aspect of working could not be properly expressed because they were not able to show their social skills without interacting with another human. Also, one noted that some people rely more on these social skills than others in getting a job.
Setup of the video interview inherently limits applicants	Even though this was not a prototype meant to provoke through the interview, participants still noted the issues they had with doing the interview through an AI system. For example, there is no system to correct a participant if they answer the question in a way the system did not anticipate and feedback during the interview was lacking. Also, the setup of the interview did not allow participants to ask questions and steer the conversation. One noted they did not even consider the fact that they could not ask questions.
People communicate a lot non-verbally	Seeing the transcripts made the participants feel that a lot of what they were trying convey was hard to see in 'just the text'. They pointed to their nonverbal communication as an important element for what they were trying to express. One participant mentioned they were fine with doing online, asynchronous interviews, as long a human would actually review the footage, as he felt that the transcripts were not able to articulate his point well.
Lack of context limits applicants	All participants mentioned a distinct lack of context was hampering them in doing better in the interviews. They did not know how well they were doing, did not understand how one should answer a question, or how questions should be related to the actual job. Although some of this effect is likely caused by the fact that this was not a real job interview where you can find more information about the job and organisation, there are some concrete elements here on how improve applicants interactions with AI interviewing systems (see Discussion).

Table 12. The clusters and their explanations of the new iteration. Continued on the next page.

They deeply felt that just transcribing the words they said did not accurately represent them. Additionally, getting the feedback after the interview caused the participants to not know how to adjust their presentation during the interview. They explained not knowing how to answer a question, or how to relate a question to the job at hand for example. These elements point to a need for a feedback system that works on the same semiotic modality (Chandler, 2017) as the input of the applicants, spoken word plus non-verbal communication, and that is able to deliver this feedback directly after a response of an applicant. Incidentally, this description could used to describe a conversation between applicant and AI system.

Value insight

Applicants care about getting feedback about their performance and answers during the interview because it helps them to shape following answers.

Another element that was quite pronounced was the impact that expectations of the participants had on their performance. For example, they would have prepared for a different type of question, more focused on understanding the personality and behaviour of a person instead of questions about the skills of the applicant. Not knowing what to expect precisely from an interview and consequently not knowing how to prepare help participants back in articulating in the best way possible what they wanted to say.

Cluster name	Explanation
Aligning expectations for interview with reality	People go into an interview with certain expectations and assumptions about the process. The type of questions was sometimes different and many did not know what to expect in terms of how to behave during the interview, because they are unfamiliar with being evaluated by an AI system. Adjacently, people's expectations of computers is that they do not make mistakes. That means that people do not behave around computers in a way that accounts for the potential for a computer system to make a mistake. Particularly for AI systems, because of their basis in probability, this is an attitude that needs to change.
Strengths and weaknesses of AI systems	Some of the participants described where they felt what the strong and weak points of AI systems are and what they would prefer AI systems therefore to focus on. One said they would like AI systems to be only used in checking elements that can be backed up by factual records, for example amount of years worked at a certain employer. Another focused on the fact that adjustments are always necessary in the review and surmised that AI systems must therefore not be good in conceptual interpretation of a video interview.
Conversation is harder with a computer	Multiple participants mentioned the concept of conversing with a computer as being somewhat problematic. One saw an interview as a way to 'put a face on the company', while another mentioned the lack of a realistic avatar as making it harder for them to respond as easily as they would when talking to a human.
Information conveyed during interview has a different impact than after the fact	All participants mentioned that checking and improving their answers after the interview was not the same as getting feedback during the interview. They got the sense the system interpreting their words wrong, even though they found it hard to articulate what the exact problem in nuance was.
Better experience through AI video interview	Participants mentioned several elements in their experience which were improvements to how they generally experience regular interviews. The option to have multiple tries to nail one's answer to a question and the built-in opportunity of the platform to take notes to help one stay on target with their answer are examples that participants mentioned as empowering them.
Better outcome through AI video interviews	Even though self-representation was limited by the AI interview as mentioned in clusters above, the participants also mentioned unique benefits in the outcome of an AI interview. One mentioned that for the problems he saw with AI interviews, he saw the assessment still being better than having a hiring manager look at your CV for five seconds and judging one on that. Another mentioned the opportunity to get concrete and specific feedback as a very positive aspect of the system. One participant mentioned the fact that speaking to a computer did not reinforce the traditional hierarchy between applicant and recruiter. Especially for them as a novice job seeker, this was seen as a positive development where they felt more free to speak on their own account.

Table 12. Continuation of the previous page. Continued on the next page.

Value insight

People do not know well what to expect from an AI interview and are therefore not able to prepare themselves well for the experience, which is important for them in order to represent themselves well.

Cluster name	Explanation
Positive elements of the review	Overall, participants found that the feedback system got the gist of what they were saying and that having the chance to see how one's words are interpreted, assess one's own performance and re-articulate if necessary were valuable aspects that helped their feeling of having autonomy over their self-representation. This indicates that having a feedback system can help with retaining autonomy over self-representation.
Approaches to interviews	Several participants spoke of practices they used to stand out in interviews, like using small talk and specific conversation skills as ways to create a connection with the recruiter in order to improve their chances of getting assessed positively.
Codes not in a cluster	Two codes could not be placed in a cluster relevant to the project.

Table 12. Continuation of the previous page.

One participant felt they were not able to represent themselves well because they saw their social skills as a large part of their identity. Their approach to interviews rested for a significant part on making small talk and making a connection with the interviewer, both of which they felt was impossible with an AI interview. While this seems antithetical to the idea of selecting applicants based on merit, they argued that given the social, collaborative nature of work, interacting with other humans and being social is an essential part of being a good worker. This is a plausible idea, and while regular interviews do indeed inherently asses both merit and social connection, AI system do not, as they are often explicitly meant to remove human bias from the assessment. That does not mean that checking if one has a fit with the team should not be part of the hiring process, only that an AI interview is not the place to do that. Helping applicants understand what specifically the AI interview is meant to do may help in this aspect.

That brings forward another point. The three clusters on how the participants had a better experience and outcome with the AI interview and review show that there were many elements of the experience that they felt were better than what happens in a regular interview. For example, the possibility to have multiple tries to get your answer right was a positive element in the eyes of all participants and that for novice job seekers a computer is not as intimidating as an experienced recruiter, which can make people feel more at ease. This shows that AI systems also have unique opportunities for self-representation that may not even have been possible before with regular interviews. AI interviews are different, not necessarily bad. It is prudent to lean into the possibilities these systems provide, as they do have strengths that are unmatched in humans, such as scale and processing speed. These differences do require a more thorough analysis of the impact of each part of the interview, as these systems do have the potential to remove autonomy over self-representation if they are not implemented and controlled carefully.

That connects to the final point of this discussion, namely how the setup of the interview and the AI system can change the ability of applicants to self-represent. The Q&A style of the interview made a participant not even aware of the fact that they were in no position to ask questions and when one participant answered a question in the 'wrong' way (on a question on work ethic, they described themselves as wanting to work in an ethical way), the AI system does not nudge them and explain the question better, it just parses the answer and judges that. This is a concrete example of how the way the system works directly impacts self-representation. Both in implementation and in development, every assumption should not be taken at face value, as it can be an avenue for loss of autonomy over how one represents themselves.

Conclusion

Overall, the last iteration shows well the primary values that people have in hiring interview situations. Applicants care to get in the moment information on how they are doing and do currently not know what to expect from AI interviews. Some may expect interviews to rely on social skill, which is not a very important factor in AI interviews currently. However, there are also positive avenues for autonomy over self-representation in AI interviews, such as having multiple tries to perfect your answer to a question.

It remains important to remind ourselves that AI technology has given us a choice we did not have before. Before AI systems were commonplace, humans were the only option for making decisions in hiring processes. Now, with two actual options, it has become an ethical choice to work through humans or AI systems, which burdens us with finding out how and when which option is the better choice. In the next chapter, I will dive deeper into what different stakeholders should do to make these choices.

Design requirements & recommendations

In this chapter, I will go over the final design requirements for ensuring autonomy over self-representation in AI video hiring systems, and what the implications of those design requirements are for both AI system makers, employers and applicants. Considering these implications is important because the problems caused by AI interviews cannot be overcome just by designing the system well, as the adaptive nature of AI systems can result in unknown behaviour in unforeseen situations. Instead, they will require constant iteration and vigilance by different stakeholders to make sure that autonomy over self-representation remains ensured for applicants.

From the prototypes carried out within this project, it has become clear that the value of self-representation is an important factor for applicants in interviews. They tend to look for opportunities to show their personality and specific combination of skills to companies. However, current implementations of AI video interviews do not accommodate this value well, for multiple reasons.

First, both rounds of prototyping showed the distinct lack of context that was impeding applicants. While the prototyping rounds in this project were necessarily less context rich than real life situations, answers on how applicants learned about the companies indicated that there are also other reasons. Applicants use all the context they can get during an hiring process to update their mental model of what they think will give the best impression on the assessing entity. The quick back-and-forth of regular interviews is normally a good example of that. However, this feedback is much less available during AI interviews (see also Guchait et al., 2014). Not knowing how to answer questions, not knowing how answers get interpreted and assessed and not knowing what information might be relevant for the company to know are seen prime elements of how the lack of context is removing autonomy over self-representation for applicants.

Second, applicants hold very little options for steering the conversation or putting emphasis on what they want show. Because current 'interviewing' system work more like Q&A than a true interview, which tend to be more conversation-like, applicants have no options to ask questions of their volition or rephrase a question to put emphasis on a concept that they think is more important. This is problematic for autonomy over self-representation because applicants may have other ideas than the recruiting organisation on what they find important in the job, or on what they see as their strengths in relation to the position. Without the ability to refocus the conversation, applicants have much less room to show these different interpretations of their experiences and their view of the work.

Third, because applicants are not talking to a human, many assumptions and expectations of how you should do your interview become uncertain. For example, AI systems are still more brittle than humans and are less able to adjust to problems with changing light conditions or low quality audio. Yet applicants do not know about this. Given the inexperience most people have with AI interviews, many lean on what they know from interviews with humans, even though some of those ideas may actually be not applicable in an AI interview. They may try to lean on their social skills, while that may cause their answers to not fit the answering model of the AI system.

AI these problems need to be addressed before autonomy over self-representation can be valued within an AI interview. However, there may also be other problems that were not surfaced in this project (this is explained more after the requirements). Still, applicants must be given the room to focus on their strengths in interviews and be able to see if the meaning they are trying to convey is actually getting to the other party. The design requirements are based on these problems.

The feedback system needs to be integrated into the interview itself

As the generative prototypes have shown, feedback after the interview is still much less useful because the applicant does not get small nudges during the interview on whether they are giving good answers or on whether they understood the question well. In regular interviews, the interviewer gives away much more information on how the applicant is doing and how they understand a certain answer. This continuous loop is necessary for applicants to iteratively learn what a company is looking for and adjust their answers to that. This requirement can be (partially) fulfilled by AI systems being able to hold a meaningful conversation, which is something we see in a limited fashion currently in assistants like Siri, Google Assistant and Alexa. Being able to hold a meaningful conversation means also to be able to react and adapt to questions and new directions in the conversation that a Q&A style interview would not allow for. Conversations makes sense because humans are very well attuned to reacting to that modality. However, other ways of developing this feedback system should be researched further too, because it may uncover ways for feedback that may be easier to implement in current AI systems.

Expectations and assumptions need to be managed at all stages

In all iterations, people expressed opinions on how they thought the system worked, but most were based on what they know about human interviews, which is not necessary relevant or even true for AI interviews. AI interviews currently tend to focus on behavioural analysis (Suen et al., 2019), which is not something people tend to be aware of. Having a better understanding of how AI interviews work would help applicants, in knowing what regular interview axioms do not apply for AI interviews for example, but also employers, as they need to know how reliable the AI system is. For example, an AI system grading applicants can provide a reliability score on its own assessments, showing how trustworthy it thinks its own grading is. For example, Bansal et al. (2019) shows how communication about AI error boundaries may help humans understand when to accept or override a recommendation from an AI system.
An important part of this effort should also focus on making people aware that AI systems are ultimately computer systems and that people should not inherently trust data because it is generated by a computer. Training for users of the system, and warnings about the probabilistic nature of the recommendations built into the UI are examples of how users can be guided to maintain a healthy dose of scepticism about decisions made by the AI system.

Clear options for escalation when something goes wrong should be built into the socio-technical system of which the AI system is a component

A large part of the frustration that people felt during the interviews was due to the fact that if something went wrong, there was no way to alert or involve a human to take a closer look and intervene when an interaction between applicant and AI system goes wrong. This means that these escalations can built into the AI system itself, or they can be part of the larger hiring process. For example, the AI system can have the option for human (recruiter) review of a particular interpretation put into the UI or part of the hiring procedure may be to ask the applicant during the regular interview if they had any trouble with the AI interview. However, because not everyone will be given the opportunity for a regular interview, it is important to look for escalation pathways already during the AI interview or directly after.

Combined, the requirements help with ensuring autonomy over self-representation for applicants by helping them learn about what is different about AI interviews and how they can prepare for them, by giving them feedback during the interview on how they are doing, and by giving them options to escalate if something does go wrong.

However, there is one aspect of AI systems assessing people in hiring that was not very observable in my prototypes but that I did come across in my literature work: as AI systems ultimately need to quantify people in order to assess them, it brings forward the question whether (aspects of) job qualifications can quantified without losing part of their meaning (Delandshere & Petrosky, 1998; Van den Hoven & Manders-Huits, 2008). This question strikes at the crux of how AI systems handle information and make decisions and how quantification in itself is an act with associated ethics (Saltelli, 2020; Sareen et al., 2020). The Wizard-of-Oz setup of my research, with the researcher ‘playing’ an AI system, misses the essential component of an AI system assessing the participants. As a human agent, our information processing capacity is inherently different from an AI system (Logan & Tandoc, 2018). Because the quantification and information handling process are so alien for a human, the research setup is not conducive to discovering information on what impact quantification of information has on the self-representation of applicants. However, it is a very important to have this avenue of questioning researched further, as it is so fundamental to how AI systems process information and can therefore have a large impact on the autonomy over self-representation of applications.

Impact on stakeholders

These design requirements put the largest onus on AI tool makers. Making a feedback system that is able to communicate both verbally and non-verbally with applicants on how they are doing requires advances in natural language processing, communication of concepts and non-verbal communication. That last element may be easier to improve by letting the AI system emulate a human, for example through an avatar that you are speaking to during the interview. However, caution should be taken when doing this, as anthropomorphising an AI system may result in people attributing human characteristics to a system that does not have these characteristics (Salles et al., 2020; Złotowski et al., 2015). Future research could see what other possibilities exist for meaningful nudges that inform applicants without leaning into AI systems acting as humans. Additionally, AI tool makers can build options for escalation if something goes wrong into the AI system, where it automatically gets logged and the tool maker can learn on how to improve the system so the same fault will not happen again.

However, providing options for escalation could also be integrated in the hiring process itself, with employers explicitly including moments where they ask applicants if they are having problems. While this may increase the workload of the employer, it will still be significantly less than the time saved by including a good AI video interview system in the hiring process, while maintaining autonomy over self-representation. Employers also need to work with the providers of their AI systems on managing expectations, both for themselves and for applicants. For themselves, to learn the capabilities and limits of the system and how to use the system in the way that creates the lowest chance of errors. For applicants, to train them in the unique differences of an AI interview, like the relative uselessness of social skills in AI interviews, and the telltale signs of limitations of self-representation in AI interviews, like not having the opportunity to refocus an interview.

Applicants need to learn more about how work well in an AI interview: currently, people assume too much that because it is called an interview, that it will work like an interview with another human. This is largely a cultural shift, one that can already be seen in the many resources on the internet helping one through a HireVue video interview for example (CareerVidz, 2020; Primal Career, 2019). However, general advice only goes so far, and applicants should ask employers on the specifics of their specific hiring process to make sure that they understand in what way they will be assessed.

Informing applicants on how the AI system works is a two-sided problem: on one hand, employers benefit if applicants are able to understand the system and represent themselves well because applicants can show better who they truly are. On the other hand however, this also helps applicants to ‘game the system’ and learn how to improve their chances over other applicants, which diminishes the value of the entire system. This is a very context dependent problem, relating to the specific AI system in use, the competences that need to be assessed and the field and employer that wants to use the AI system. Therefore, finding the balance between these two conflicting interests requires testing and iterating on the hiring process in a collaboration between employers, tool makers and applicants.

The role of designers

Given the impact of these requirements on stakeholders, what should be the role of designers in the rollout of AI interview systems that respect self-representation?

First, the requirements give guidance on a high level, focusing on explaining the playing field that is necessary for good autonomy over self-representation. But the implementation of the requirements will differ from organisation to organisation. Just as how each organisation has a unique hiring process, the integration of an AI system in the process will be unique. Different vendors have different algorithms, features, and capabilities that can profoundly change what impact an AI system will have. But also choices on the level of integration, the goal of the system, the job field, and the level of digital literacy in the organisation are all factors that are part of the implementation process that need to be worked in order for an AI system to be used effectively. This implementation process requires a lot of iteration as the context is so complex that it is very hard to get right on the first go. Going through these iterations while staying focused on the goal of the project is a place where designers can lean on their iterative and collaborative approach (Norman & Stappers, 2015).

Additionally, a significant part of the requirements is focused on helping people understand the respective AI systems and its capabilities and limitations. Designers can have a specific role in making information understandable to different stakeholders, pulling from their knowledge about the technology and their focus on humans to bridge the gap between the two entities. Their focus on the user and willingness to co-create with them helps them understand what users are looking for and what the specific problems are they come across, while their technical knowledge helps them discuss with technologists on what would be feasible solutions to those problems. Finally, they can use their skills in assessing the key sensitivities in rolling out the AI system and how that may impact different groups in the organisation, like how onboarding should adjust to hires who have yet talked much less to different people in the organisation (Calabretta et al., 2016).

Limitations and reflections

Creating meaningful work is an exercise in restraint, in scope and focus. As such, there are several limitations that need to be taken into account when looking at the findings of this project. They are ordered according to project activity. Then, some steps for future research are outlined.

Finally, I have written about my contribution to the field and my personal reflection on the project.

Limitations

With this project I tried to highlight the importance of autonomy over self-representation, how AI systems impact this value and what can be done in the future to ensure that applicants can show their true self in and find a fulfilling job. These limitations help scope the findings of the project.

Before I look at the limitations of the different project activities, I want to mention my own inherent bias as a researcher again. I have almost completely done this project by myself and my skeptical view on AI systems is one that can have blinded me to insights and design solutions. I have my own preferences in design, in research approach, in solutions. It is very well possible that other, better solutions are overlooked through some limitation of my work.

Limitations of the literature research

This project is focused on the intersection of AI, ethics and HRM (Human Resource Management). A literature study on these topics was done to find research gaps and create better understanding of how they interact. Each of these topics is a sprawling, complex field with new knowledge created everyday. Despite my best efforts, it is very well possible that I have overlooked essential work that could have shifted the outlook of the project.

Limitations of the interviews

With the interviews the choice was made for practical reasons to focus on people within Europe, who almost all highly educated. More research needs to be done to see how the findings from the interviews also apply to other regions and demographics. Additionally, no AI tool maker was interviewed for their perspective. Their view on the topic of autonomy over self-representation may have important ramifications that are not clear from working with applicants and HR personnel.

Limitations of the generative prototypes

One of the primary elements of for a loss of autonomy over self-representation that came out of the iterations of the generative prototypes was a limited context for the applicants to infer information from. While the insights gathered show that this is for a large part inherent to how AI systems currently work, that does not negate that applicants normally have more information about a company through the search they did, the company website, perhaps they have talked to someone already there. A more immersive experience with more context could provide more precision in this matter.

Also, the participants involved in the prototypes were all young and highly educated. It is possible that autonomy over self-representation in AI systems is easier for them than for people who have less experience with computer systems in general or who already have less unique skills that they can use as promote themselves.

Additionally, the method of Wizard of Oz as used in this project has interesting limitations on what kind of insights can be generated from the prototypes. The effects of quantification on meaning found in the literature research was hard to emulate in the prototypes because the very information processing cannot be done by a human in a quantitive way, as our brains operate mainly on qualitative information (Jonassen, 2000; Simon, 1978).

Future research

The conclusions and limitations of this project have shown multiple avenues for new research as there are many questions still unanswered by this project. The dearth of research on self-representation in AI hiring systems has shown that there is still unknown ground in this intersection between AI, work and ethics.

To further improve this conceptual work on self-representation, it is necessary to involve AI tool makers. This thesis lays out many new technical challenges and their experience and expertise will be invaluable to better understand how AI systems should be designed to respect to self-representation.

Additionally, this project only looks at the concept of self-representation as a way for people to be harmed by AI systems. More research is needed how this value can be combined and integrated with approaches that cover other problems with AI systems. Frameworks such as the Designing for Human Rights in AI roadmap (Aizenberg and Van den Hoven, 2020) or the IHRL Framework for Algorithmic Accountability (McGregor et al., 2019) may help in this aspect.

The inclusion of direct, real-time feedback systems to improve autonomy over self-representation for applicants is one of the primary findings of this project. I point to conversations as a semiotic modality that works well for humans, because we are so attuned to conversations, but that does not mean that other modalities should be ruled out up front. Particularly because conversation about abstract concepts like values and human wishes is currently quite a hard problem for the AI community, it makes sense to look at different avenues to communicate back to applicants on how they are doing and how the system is interpreting their answers.

Finally, while there is some work on autonomy over self-representation in the social sciences and engineering ethics fields, there is very little research yet in AI circles on how this autonomy should be taken into account into the design of socio-technical systems involving AI.

> "I am feeling increasingly challenged that my skillset of being very good at programming has become somewhat secondary. It's really the bigger picture understanding of 'Who would be using that? How transparent do I need to build it for it to be adopted at some point? What types of biases in the data collection and then also in the usage?' I think, in certain areas, we have societal expectations as to what is fair and what isn't."
> "And so, it's not just the provenance of that data, but it's, sort of, deeply understanding 'Why does it look the way it looks? Why was it collected this way? What are the limitations of it?' We need to think about that in the entire process, in how we document that process."
>
> Claudia Perlich and Hilary Mason (2021)

Contribution to the field

This project intends to deliver new empirical research that will push the AI and HR communities to combine their efforts in better understanding how self-representation can be a new value to take into account when designing AI systems for the hiring process. It is intended to shed a light on how self-representation is intimately connected to the practice of hiring and what kinds of problems can be encountered when AI systems are being implemented in the hiring process.

The design requirements are intended to give AI tool makers, HR personnel and organisation executives concrete steps forward to work together on improving the autonomy over self-representation of applicants in the hiring process. The project aspires to inspire others in the design, AI, and HR community to take on this important problem and make sure that each person has a fair chance at good employment.

The research value originates from the combination of the approach. The generative prototypes from ASF brought a flexibility to designing for changing values that VSD originally did not have. VSD itself is however flexible enough to be able to incorporate the iterations of ASF as empirical investigations as part of the overall VSD approach. This supplemented VSD in particular for designing for AI systems, as their adaptive nature can lead to ever shifting contexts and values.

Through the use of generative prototypes, this project is a good example how one can do research through design: the designed prototypes were integral to furthering understanding of the interaction of AI systems with the value of self-representation and the implications of that interaction for applicants.

I recently watched a conversation between two AI experts (WIRED, 2021) and they said the quote above. This encapsulates for me very well what the value of design is in this field. As AI becomes more widespread and valuable, its problems become more diverse and complex. AI technologists cannot fix all these problems, because the problems do not always lie in the technology. This project shows that where design can make a contribution to integrating AI into society, through highlighting the needs of different stakeholders and seeking creative solutions.

Personal reflection

I've have struggled a lot in my studies with figuring what I wanted to do within design. During my bachelor in Eindhoven, I needed to write a personal vision on design every semester as part of my Personal Development Plan. I remember it took me multiple years to write down something that would actually still be there the next semester, instead of hopping to a completely different interest. My interests had already somewhat coalesced into a singular idea when I started SPD in Delft: I knew I wanted to work between technology and society and concern myself with the ethics of introducing that technology. But still, I had not really found my niche yet, it was still too broad.

My graduation project was for me a milestone where I was able to connect my interests with what I find meaningful to do within design: figuring out how AI systems should be used within society. I truly feel I have found some work that feels meaningful to me, that feels worth working on. It pushed me to work harder, write better, think more. I am very happy that I was able to contribute something to this field of AI and ethics.

Still, there were many things within the project that I had difficulties with. One was scope. I originally set out to create a strategy for HR departments to make ethical choices in implementing AI systems. While adjacent, this is clearly not the topic of my thesis anymore. I think I could have done a better job of scoping the project, as I particularly struggled in the beginning to make sense of the massive body of research that is AI ethics nowadays. I remember peering through hundreds of papers in first few weeks, with little idea of what would be the more interesting ones. Better scoping could have helped significantly in this aspect.

Connected to that, I realised only after my midterm that I was not sufficiently in control of my own process. I would often get sidetracked, or stuck digging too deep in the philosophical bedrock of different design approaches, which is not something a master thesis is intended to focus on. One of my mentors said that I should not just accept every suggestion they were making, but that I needed to critically assess whether the suggestion made sense for where I wanted to take the project. From that point on, I became more focused and my work improved. I'm naturally a good listener and always open to suggestions, but that does not mean that I should not be thinking for myself. I still have a lot to learn process-wise.

Still, there are other aspects where I see significant growth in me as a designer. I've become a much better writer, through the many rounds of feedback I have gotten back from my supervisors. I feel I am more able now to articulate my thoughts into written words and bridge different arguments. Also I see myself become more pragmatic. This was a project on the bleeding edge of AI ethics. I tend to want to deliver pragmatic solutions but had to grapple with a context that did not really allow for a solution that could completely solve the problem. Instead, I learned to accept to push for the best possible solution, even if that solution is not everything I had hoped it to be. When working with ethical problems, that is truly the best you can hope for.

Project goals

Originally, I set out to use systemic design extensively throughout the project. Looking back, that did not really happen. As I was doing my literature research I noticed that systemic design and VSD are sometimes conflicting in approach. Given that I already had some experience with systemic design and none yet with VSD, I chose to focus on VSD for this project and let systemic design be for now.

Working with VSD was a very interesting experience, where it informed a large part of the approach for the project, only to encounter a roadblock in not knowing how to deal with generative prototypes well within the framework. VSD has given me a good basis for future work in ethical design, even though I feel I have only scratched the surface in terms of all the different tools that are part of the methodology.

One of my deepest beliefs about design is that you have only delivered value when your work is being used in the real world. To that end, I intended to have two companies at the end of my project working on implementing whatever the end point fo the project would be. This did not happen, for two reasons. First, my scope was too big in the beginning of the project and given how new this research field is, I am not surprised that I was not able to go to the implementation phase. However, second, I do feel I could have done more to involve HelloMentor into the project and work with them to see what my work could do in actual organisations. That is on me, and I intend to learn better how keep external stakeholders in the loop of the project.

Finally, I wanted this project to be a showcase in my portfolio of what my interests are and what I am capable of. I feel I have definitely accomplished this goal. I am very proud of this thesis and am certain that it will help in finding a job that aligns with my future plans.

Bibliography

Abowd, G. D. (2012). What next, ubicomp?: Celebrating an intellectual disappearing act. Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp '12, 31. https://doi.org/10.1145/2370216.2370222

Adams, W. C. (2015). Conducting Semi-Structured Interviews. In K. E. Newcomer, H. P. Hatry, & J. S. Wholey (Eds.), Handbook of Practical Program Evaluation (pp. 492–505). John Wiley & Sons, Inc. https://doi.org/10.1002/9781119171386.ch19

Aizenberg, E., & van den Hoven, J. (2020). Designing for human rights in AI. Big Data & Society, 7(2), 205395172094956. https://doi.org/10.1177/2053951720949566

Alsheikh, T., Rode, J. A., & Lindley, S. E. (2011). (Whose) value-sensitive design: A study of long- distance relationships in an Arabic cultural context. Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work - CSCW '11, 75. https://doi.org/10.1145/1958824.1958836

Alzubi, J., Nayyar, A., & Kumar, A. (2018). Machine Learning from Theory to Algorithms: An Overview. Journal of Physics: Conference Series, 1142, 012012. https://doi.org/10.1088/1742-6596/1142/1/012012

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. Psychological Science in the Public Interest, 20(1), 1–68. https://doi.org/10.1177/1529100619832930

Behymer, K. J., & Flach, J. M. (2016). From Autonomous Systems to Sociotechnical Systems: Designing Effective Collaborations. She Ji: The Journal of Design, Economics, and Innovation, 2(2), 105–114. https://doi.org/10.1016/j.sheji.2016.09.001

Black, J. S. (2014). It starts with one: Changing individuals changes organizations (Third edition). Pearson.

Black, J. S., & van Esch, P. (2020). AI-enabled recruiting: What is it and how should a manager use it? Business Horizons, 63(2), 215–226. https://doi.org/10.1016/j.bushor.2019.12.001

Boer, L., & Donovan, J. (2012). Provotypes for participatory innovation. Proceedings of the Designing Interactive Systems Conference on - DIS '12, 388. https://doi.org/10.1145/2317956.2318014

Boer, L., Donovan, J., & Buur, J. (2013). Challenging industry conceptions with provotypes. CoDesign, 9(2), 73–89. https://doi.org/10.1080/15710882.2013.788193

Borning, A., & Muller, M. (2012). Next steps for value sensitive design. Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems - CHI '12, 1125. https://doi.org/10.1145/2207676.2208560

Bradbury, R. J. (2000). Matrioshka Brains. 27.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Broek, E. van den, Sergeeva, A., & Huysman, M. (2019). Hiring Algorithms: An Ethnography of Fairness in Practice. ICIS 2019 Proceedings, 6, 10. https://aisel.aisnet.org/icis2019/future_of_work/future_work/6/

Buchanan, R. (2001). Human Dignity and Human Rights: Thoughts on the Principles of Human-Centered Design. Design Issues, 17(3), 35–39. https://doi.org/10.1162/074793601750357178

Buur, J., & Matthews, B. (2008). Participatory Innovation. International Journal of Innovation Management, 12(3), 255–273.

Calabretta, G., Gemser, G., & Karpen, I. (Eds.). (2016). Strategic design: Eight essential practices every strategic designer must master. BIS.

CareerVidz. (2020, March 8). HIREVUE Interview Questions, Tips and Answers! How to PASS a HireVue Interview! https://www.youtube.com/watch?v=ycG5_uccoNk

Chandler, D. (2017). Semiotics: The basics (Third edition). Routledge.

Coeckelbergh, M. (2020). AI ethics. The MIT Press. https://mitpress.mit.edu/books/ai-ethics

Cummings, M. L. (2006). Integrating ethics in design through the value-sensitive design approach. Science and Engineering Ethics, 12(4), 701–715. https://doi.org/10.1007/s11948-006-0065-0

da Costa Junior, J., Diehl, J. C., & Snelders, D. (2019). A framework for a systems design approach to complex societal problems. Design Science, 5, e2. https://doi.org/10.1017/dsj.2018.16

Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993). Wizard of Oz studies—Why and how. Knowledge-Based Systems, 6(4), 258–266. https://doi.org/10.1016/0950-7051(93)90017-N

Dastin, J. (2018, October 10). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

Dattner, B., Chamorro-Premuzic, T., Buchband, R., & Schettler, L. (2019). The Legal and Ethical Implications of Using AI in Hiring. 7.

Delandshere, G., & Petrosky, A. R. (1998). Assessment of Complex Performances: Limitations of Key Measurement Assumptions. Educational Researcher, 27(2), 14–24.

Dorst, K. (2011). The core of 'design thinking' and its application. Design Studies, 32(6), 521–532. https://doi.org/10.1016/j.destud.2011.07.006

Duke University. (2018). Typical Questions from HireVue Interviews.pdf. https://econ.duke.edu/sites/econ.duke.edu/files/file-attachments/Typical%20Questions%20from%20HireVue%20Interviews.pdf

European Convention of 1999. (2012). Charter of Fundamental Rights of the European Union. Official Journal of the European Union, 17.

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its Nature, Scope, Limits, and Consequences. Minds and Machines, 30(4), 681–694. https://doi.org/10.1007/s11023-020-09548-1

Forlizzi, J. (2013). The Product Service Ecology: Using a Systems Approach in Design. 28.

Fountaine, T., McCarthy, B., & Saleh, T. (2019). Building the AI-Powered Organization. 13.

Friedman, B., & Hendry, D. G. (2019). Value Sensitive Design: Shaping Technology with Moral Imagination. The MIT Press. https://doi.org/10.7551/mitpress/7585.001.0001

Friedman, B., Hendry, D. G., & Borning, A. (2017). A Survey of Value Sensitive Design Methods. Foundations and Trends® in Human–Computer Interaction, 11(2), 63–125. https://doi.org/10.1561/1100000015

Friedman, B., Kahn, P. H., & Borning, A. (2002). Value Sensitive Design: Theory and Methods. University of Washington Technical Report, 2–12.

Fuller, J. B., Raman, M., Sage-Gavin, E., & Hines, K. (2021). Hidden Workers: Untapped Talent. How leaders can improve hiring practices to uncover missed talent pools, close skills gaps, and improve diversity (p. 74). Harvard Business School. https://www.hbs.edu/managing-the-future-of-work/Documents/research/hiddenworkers09032021.pdf

Guchait, P., Ruetzler, T., Taylor, J., & Toldi, N. (2014). Video interviewing: A potential selection tool for hospitality managers – A study to understand applicant perspective. International Journal of Hospitality Management, 36, 90–100. https://doi.org/10.1016/j.ijhm.2013.08.004

Hacioglu, U. (Ed.). (2020). Handbook of Research on Strategic Fit and Design in Business Ecosystems: IGI Global. https://doi.org/10.4018/978-1-7998-1125-1

Haenlein, M., & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. California Management Review, 61(4), 5–14. https://doi.org/10.1177/0008125619864925

Hamilton, R. H., & Davison, H. K. (2018). The search for skills: Knowledge stars and innovation in the hiring process. Business Horizons, 61(3), 409–419. https://doi.org/10.1016/j.bushor.2018.01.006

Hao, K. (2019, December 27). In 2020, let's stop AI ethics-washing and actually do something. MIT Technology Review. https://www.technologyreview.com/2019/12/27/57/ai-ethics-washing-time-to-act/

Heikkilä, M., & Kuivaniemi, L. (2012). Ecosystem Under Construction: An Action Research Study on Entrepreneurship in a Business Ecosystem. Technology Innovation Management Review, 2(6), 18–24.

High-Level Expert Group on AI. (2019). EU Ethics Guidelines for AI. European Commission. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Howard, T. (2014). Journey mapping: A brief overview. Communication Design Quarterly, 2(3), 10–13. https://doi.org/10.1145/2644448.2644451

Jonassen, D. H. (2000). Toward a design theory of problem solving. Educational Technology Research and Development, 48(4), 63–85. https://doi.org/10.1007/BF02300500

Jones, A. J. I., Artikis, A., & Pitt, J. (2013). The design of intelligent socio-technical systems. Artificial Intelligence Review, 39(1), 5–20. https://doi.org/10.1007/s10462-012-9387-2

Kallinikos, J., Aaltonen, A., & Marton, A. (2013). The Ambivalent Ontology of Digital Artifacts. MIS Quarterly, 37(2), 357–370. https://doi.org/10.25300/MISQ/2013/37.2.02

Keller, S., & Meaney, M. (2017, November 24). Attracting and retaining the right talent. McKinsey. https://www.mckinsey.com/business-functions/organization/our-insights/attracting-and-retaining-the-right-talent

Knight, W. (2021, January 12). Job Screening Service Halts Facial Analysis of Applicants. Wired. https://www.wired.com/story/job-screening-service-halts-facial-analysis-applicants/

Kodiyan, A. A. (2019). An overview of ethical issues in using AI systems in hiring with a case study of Amazon's AI based hiring tool. 10.

Laumer, S., Maier, C., & Eckhardt, A. (2015). The impact of business process management and applicant tracking systems on recruiting process performance: An empirical study. Journal of Business Economics, 85(4), 421–453. https://doi.org/10.1007/s11573-014-0758-9

LCEF. (2020). Civil Rights Principles for Hiring Assessment Technologies. The Leadership Conference Education Fund. http://civilrightsdocs.info/pdf/policy/letters/2020/Hiring_Principles_FINAL_7.29.20.pdf

Le Dantec, C. A., Poole, E. S., & Wyche, S. P. (2009). Values as lived experience: Evolving value sensitive design in support of value discovery. Proceedings of the 27th International Conference on Human Factors in Computing Systems - CHI 09, 1141. https://doi.org/10.1145/1518701.1518875

Lee, I., & Shin, Y. J. (2020). Machine learning for enterprises: Applications, algorithm selection, and challenges. Business Horizons, 63(2), 157–170. https://doi.org/10.1016/j.bushor.2019.10.005

Legg, S., & Hutter, M. (2007). A Collection of Definitions of Intelligence. ArXiv:0706.3639 [Cs]. http://arxiv.org/abs/0706.3639

Lev, B. (2001). Intangibles management, measurement, and reporting.

Levesque, H. J. (2018). Common sense, the turing test, and the quest for real AI. The MIT Press.

Li, L., Oh, J., Lassiter, T., & Lee, M. K. (2021). Algorithmic Hiring in Practice: Recruiter and HR Professional's Perspectives on AI Use in Hiring. Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES '21), 11.

Lim, Y.-K., Stolterman, E., & Tenenberg, J. (2008). The anatomy of prototypes: Prototypes as filters, prototypes as manifestations of design ideas. ACM Transactions on Computer-Human Interaction, 15(2), 1–27. https://doi.org/10.1145/1375761.1375762

Logan, R., & Tandoc, M. (2018). Thinking in Patterns and the Pattern of Human Thought as Contrasted with AI Data Processing. Information, 9(4), 83. https://doi.org/10.3390/info9040083

Makarius, E. E., Mukherjee, D., Fox, J. D., & Fox, A. K. (2020). Rising with the machines: A sociotechnical framework for bringing artificial intelligence into the organization. Journal of Business Research, 120, 262–273. https://doi.org/10.1016/j.jbusres.2020.07.045

Manders-Huits, N., & Van den Hoven, J. (2008). Moral identification in Identity Management Systems. In S. Fischer-Hübner, P. Duquenoy, A. Zuccato, & L. Martucci (Eds.), The Future of Identity in the Information Society (pp. 77–91). Springer US. https://doi.org/10.1007/978-0-387-79026-8_6

Maurer, S. D., & Liu, Y. (2007). Developing effective e-recruiting websites: Insights for managers from marketers. Business Horizons, 50(4), 305–314. https://doi.org/10.1016/j.bushor.2007.01.002

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. 13.

McDermott, D. (1976). Artificial intelligence meets natural stupidity. ACM SIGART Bulletin, 57, 4–9. https://doi.org/10.1145/1045339.1045340

McGregor, L., Murray, D., & Ng, V. (2019). INTERNATIONAL HUMAN RIGHTS LAW AS A FRAMEWORK FOR ALGORITHMIC ACCOUNTABILITY. International and Comparative Law Quarterly, 68(2), 309–343. https://doi.org/10.1017/S0020589319000046

McIlvaine, A. (2018, June 5). Data in the Driver's Seat. HRExecutive.Com. https://hrexecutive.com/talent-acquisitions-leaders-use-ai-to-improve-hiring/

Mitchell, M. (2021). Why AI is Harder Than We Think. ArXiv:2104.12871 [Cs]. http://arxiv.org/abs/2104.12871

Mujtaba, D. F., & Mahapatra, N. R. (2019). Ethical Considerations in AI-Based Recruitment. 2019 IEEE International Symposium on Technology and Society (ISTAS), 1–7. https://doi.org/10.1109/ISTAS48451.2019.8937920

Norman, D. A. (2013). The design of everyday things (Revised and expanded edition). Basic Books.

Norman, D. A., & Stappers, P. J. (2015). DesignX: Complex Sociotechnical Systems. She Ji: The Journal of Design, Economics, and Innovation, 1(2), 83–106. https://doi.org/10.1016/j.sheji.2016.01.002

Oehlhorn, C. E., Maier, C., Laumer, S., & Weitzel, T. (2020). Human resource management and its impact on strategic business-IT alignment: A literature review and avenues for future research. The Journal of Strategic Information Systems, 29(4), 101641. https://doi.org/10.1016/j.jsis.2020.101641

Pan, S. J., & Yang, Q. (2010). A Survey on Transfer Learning. IEEE Transactions on Knowledge and Data Engineering, 22(10), 1345–1359. https://doi.org/10.1109/TKDE.2009.191

Paschen, U., Pitt, C., & Kietzmann, J. (2020). Artificial intelligence: Building blocks and an innovation typology. Business Horizons, 63(2), 147–155. https://doi.org/10.1016/j.bushor.2019.10.004

Primal Career. (2019, November 21). HireVue Video Interview: 5 MISTAKES You Need to AVOID. https://www.youtube.com/watch?v=J2VnJOw5Cd0

Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 469–481. https://doi.org/10.1145/3351095.3372828

Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A. 'Sandy', … Wellman, M. (2019). Machine behaviour. Nature, 568(7753), 477–486. https://doi.org/10.1038/s41586-019-1138-y

Raschka, S., & Mirjalili, V. (2019). Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2. http://proquest.safaribooksonline.com/?fpi=9781789955750

Raveendra, P. V., Satish, Y. M., & Singh, P. (2020). Changing Landscape of Recruitment Industry: A Study on the Impact of Artificial Intelligence on Eliminating Hiring Bias from Recruitment and Selection Process. Journal of Computational and Theoretical Nanoscience, 17(9), 4404–4407. https://doi.org/10.1166/jctn.2020.9086

Reilly, P. (2019). The impact of artificial intelligence on the HR function. In Which way now for HR and organisational changes? IES Perspectives on HR 2018 (pp. 45–62). Institute for Employment Studies.

Risam, R. (2018). Now you see them: Self-representation and the refugee selfie. Popular Communication, 16(1), 58–71. https://doi.org/10.1080/15405702.2017.1413191

Ruecker, S. (2015). A Brief Taxonomy of Prototypes for the Digital Humanities. Scholarly and Research Communication, 6(2), 11.

Sahoo, C. K., Das, S., & Sundaray, B. K. (2011). Strategic Human Resource Management: Exploring the Key Drivers. Strategic Human Resource Management, 11(2), 16.

Salles, A., Evers, K., & Farisco, M. (2020). Anthropomorphism in AI. AJOB Neuroscience, 11(2), 88–95. https://doi.org/10.1080/21507740.2020.1740350

Saltelli, A. (2020). Ethics of quantification or quantification of ethics? Futures, 116, 102509. https://doi.org/10.1016/j.futures.2019.102509

Sanders, L., & Stappers, P. J. (2012). Convivial design toolbox: Generative research for the front end of design. BIS.

Sareen, S., Saltelli, A., & Rommetveit, K. (2020). Ethics of quantification: Illumination, obfuscation and performative legitimation. Palgrave Communications, 6(1), 20. https://doi.org/10.1057/s41599-020-0396-5

Selbst, A. D., boyd, danah, Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and Abstraction in Sociotechnical Systems. FAT* '19: Conference on Fairness, Accountability, and Transparency, 10. https://doi.org/10.1145/3287560.3287598

Simon, H. A. (1978). Information-processing theory of human problem solving. In Handbook of learning & cognitive processes: V. Human information. (pp. 271–295). Lawrence Erlbaum.

Skitka, L. J., Mosier, K. L., & Burdick, M. (1999). Does automation bias decision-making? International Journal of Human-Computer Studies, 51(5), 991–1006. https://doi.org/10.1006/ijhc.1999.0252

Steen, M. (2011). Tensions in human-centred design. CoDesign, 7(1), 45–60. https://doi.org/10.1080/15710882.2011.563314

Stickdorn, M. (2018). This is service design doing: Applying service design thinking in the real world ; a practitioners' handbook (M. Hormess, A. Lawrence, & J. Schneider, Eds.; First edition). O'Reilly.

Stoimenova, N. (2021). [How to design for unforeseen values in AI design] [Doctoral dissertation].

Suen, H.-Y., Chen, M. Y.-C., & Lu, S.-H. (2019). Does the use of synchrony and artificial intelligence in video interviews affect interview ratings and applicant attitudes? Computers in Human Behavior, 98, 93–101. https://doi.org/10.1016/j.chb.2019.04.012

Tamkin, A., Brundage, M., Clark, J., & Ganguli, D. (2021). Understanding the Capabilities, Limitations, and Societal Impact of Large Language Models. ArXiv:2102.02503 [Cs]. http://arxiv.org/abs/2102.02503

Tschopp, M., & Ruef, M. (2018). On Trust in AI - A Systemic Approach. 4.

Upadhyay, A. K., & Khandelwal, K. (2018). Applying artificial intelligence: Implications for recruitment. Strategic HR Review, 17(5), 255–258. https://doi.org/10.1108/SHR-07-2018-0051

van de Poel, I. (2020). Embedding Values in Artificial Intelligence (AI) Systems. Minds and Machines, 30(3), 385–409. https://doi.org/10.1007/s11023-020-09537-4

van de Poel, I. (2021). Design for value change. Ethics and Information Technology, 23(1), 27–31. https://doi.org/10.1007/s10676-018-9461-9

van den Hoven, J. (2007). ICT and Value Sensitive Design. In P. Goujon, S. Lavelle, P. Duquenoy, K. Kimppa, & V. Laurent (Eds.), The Information Society: Innovation, Legitimacy, Ethics and Democracy In honor of Professor Jacques Berleur s.j. (Vol. 233, pp. 67–72). Springer US. https://doi.org/10.1007/978-0-387-72381-5_8

Van den Hoven, J., & Manders-Huits, N. (2008). The Person as Risk, the Person at Risk. In International Conference on the Social and Ethical Impacts of Information and Communication Technologies & T. W. Bynum (Eds.), Proceedings of the Tenth International Conference: Living, working and learning beyond technology—ETHICOMP 2008 (pp. 408–414). Tipografia Commerciale.

van Esch, P., Black, J. S., & Ferolie, J. (2019). Marketing AI recruitment: The next phase in job application and selection. Computers in Human Behavior, 90, 215–222. https://doi.org/10.1016/j.chb.2018.09.009

Visser, F. S., Stappers, P. J., van der Lugt, R., & Sanders, E. B.-N. (2005). Contextmapping: Experiences from practice. CoDesign, 1(2), 119–149. https://doi.org/10.1080/15710880500135987

West, S. M. (2020). Ethical Implications of Using Artificial Intelligence and Automated Decision Systems [Written testimony]. New York City Council Committee on Technology. https://ainowinstitute.org/ai-now-city-council-testimony-fair-shot-act.pdf

WIRED. (2021, August 18). Computer Scientist Explains Machine Learning in 5 Levels of Difficulty. https://www.youtube.com/watch?v=5q87K1WaoFI

Yoo, Y., Henfridsson, O., & Lyytinen, K. (2010). The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research. Information Systems Research, 21(4), 724–735. https://doi.org/10.1287/isre.1100.0322

Zielinski, D. (2021, May 1). What's Ahead for HR Technology in 2021. SHRM. https://www.shrm.org/resourcesandtools/hr-topics/technology/pages/whats-ahead-for-hr-technology-2021.aspx

Złotowski, J., Proudfoot, D., Yogeeswaran, K., & Bartneck, C. (2015). Anthropomorphism: Opportunities and Challenges in Human–Robot Interaction. International Journal of Social Robotics, 7(3), 347–360. https://doi.org/10.1007/s12369-014-0267-6