

# The Value of Media Ecology for Enabling Human Rights Defenders to Advocate for the Protection of the Right to Mental Health in the Context of Deploying Artificial Intelligence Technology as part of the Decision-making Process

Tetyana (Tanya) Krupiy

Volume 4, numéro 1, printemps 2024

URI : <https://id.erudit.org/iderudit/1111624ar>

DOI : <https://doi.org/10.7202/1111624ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

New Explorations Association

ISSN

2563-3198 (numérique)

[Découvrir la revue](#)

Citer cet article

Krupiy, T. (2024). The Value of Media Ecology for Enabling Human Rights Defenders to Advocate for the Protection of the Right to Mental Health in the Context of Deploying Artificial Intelligence Technology as part of the Decision-making Process. *New Explorations*, 4(1).  
<https://doi.org/10.7202/1111624ar>

Résumé de l'article

Traditionally, human rights activists gathered evidence about violations of particular individuals' human rights to demand that states change their conduct and adopt measures to prevent further violations. Deploying artificial intelligence as part of the decision-making process creates challenges for activists to detect all sources of harm and demand that states take action to address the harms. Abeba Birhane points out that employing artificial intelligence technology can generate harmful impacts that are either difficult to detect or invisible. If harms remain invisible, then it is difficult for human rights defenders to document them. Equally, it becomes challenging to articulate why the harms in question constitute international human rights law violations. As a result, it is harder for human rights defenders to call on states to take action to safeguard fundamental rights. This article puts forward that individuals can make harms arising from the deployment of artificial intelligence as part of the decision-making process more visible by using the theoretical framework of media ecology. It demonstrates that media ecology can provide an additional tool for human rights activists to detect how using artificial intelligence as part of the decision-making process can undermine the enjoyment of a human right. The article uses the right to mental health as a case study to develop this argument. In order to contextualise the analysis, the article focuses on the employment of artificial intelligence to screen candidates for employment as a case study.

© Tetyana (Tanya) Krupiy, 2024



Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter en ligne.

<https://apropos.erudit.org/fr/usagers/politique-dutilisation/>

**érudit**

Cet article est diffusé et préservé par Érudit.

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche.

<https://www.erudit.org/fr/>



## The Value of Media Ecology for Enabling Human Rights Defenders to Advocate for the Protection of the Right to Mental Health in the Context of Deploying Artificial Intelligence Technology as part of the Decision-making Process

Tetyana (Tanya) Krupiy — Newcastle University — [tanya.krupiy@newcastle.ac.uk](mailto:tanya.krupiy@newcastle.ac.uk)

**Abstract:** Traditionally, human rights activists gathered evidence about violations of particular individuals' human rights to demand that states change their conduct and adopt measures to prevent further violations. Deploying artificial intelligence as part of the decision-making process creates challenges for activists to detect all sources of harm and demand that states take action to address the harms. Abeba Birhane points out that employing artificial intelligence technology can generate harmful impacts that are either difficult to detect or invisible. If harms remain invisible, then it is difficult for human rights defenders to document them. Equally, it becomes challenging to articulate why the harms in question constitute international human rights law violations. As a result, it is harder for human rights defenders to call on states to take action to safeguard fundamental rights. This article puts forward that individuals can make harms arising from the deployment of artificial intelligence as part of the decision-making process more visible by using the theoretical framework of media ecology. It demonstrates that media ecology can provide an additional tool for human rights activists to detect how using artificial intelligence as part of the decision-making process can undermine the enjoyment of a human right. The article uses the right to mental health as a case study to develop this argument. In order to contextualise the analysis, the article focuses on the employment of artificial intelligence to screen candidates for employment as a case study.

**Keywords:** media ecology, international human rights law, harm, mental integrity, mental well-being, mental health, artificial intelligence technology, decision-making

### I. Introduction

The United Nations established the United Nations High-Level Advisory Body on Artificial Intelligence in October 2023 to recommend how states should govern artificial intelligence

technology (hereinafter AI) on the global level (Office of the Secretary-General's Envoy on Technology 2023). This development is due to an increasing awareness that using AI as part of the decision-making process increasingly impacts people's life opportunities, dignity and the enjoyment of human rights (Sapignoli 2021, 8). International organisations (United Nations Secretary-General António Guterres 2018, 3), government authorities (Henley 2021) and private companies (Schildhorn and Ford 2022) use AI as part of the decision-making process (Sapignoli 2021, 5). For this reason, there is a potential that the employment of AI by organisations to reach decisions about people's entitlements to resources and opportunities can significantly impact the subjects of the decision-making (Ibid., 8). Furthermore, UNESCO acknowledged that there is a need to consider the broader impacts that the employment of AI brings about, such as changing how individuals express themselves and engage in social interactions (UNESCO 2021, 5).

In light of these developments, Eileen Donahoe and Megan MacDuffee Metzger put forward that norms of international human rights law can provide a global shared framework for requiring states to regulate how organisations design and deploy AI (Donahoe and MacDuffee Metzger 2019, 116). They view the employment of international human rights law as guiding the development and employment of AI and enabling states to govern AI with "wide global legitimacy" (Ibid.). Moreover, they claim that international human rights law can guide states to ensure that organisations develop and use AI in a manner that respects human dignity (Ibid.). Human rights advocates play a crucial role in informing governments of the policies and laws they need to adopt to govern AI in a manner compatible with international human rights law (Chander and Jakubowska 2021). In order to call on states to fulfil the states' international human rights law obligations (United Nations 2024), human rights activists first need to establish what harm the employment of AI creates. They then need to show why human rights norms apply to a particular type of harm. However, it can be challenging for the affected individuals and human rights activists to detect and describe all harms arising from the deployment of AI (Birhane 2021, 129).

The United Nations High Commissioner for Human Rights, Michelle Bachelet, acknowledged that the opaqueness behind AI decision-making processes makes it challenging to evaluate

the effects of such systems on compliance with human rights norms (Human Rights Council 2021, para. 20). She said that “data environment, algorithms and models underlying the development and operation of AI systems” are factors which impact on the ability of the public to understand the human rights consequences of using these systems (Ibid., para. 20). Tetyana Krupiy explains that the naked eye cannot always detect the cause-and-effect relationship between the use of AI and the effect on the individual (Krupiy 2021a, 14-15). The nature of AI as a technical system makes it hard to trace a direct relationship between the use of particular data to program AI and the impact on the applicant (Ibid.). Similarly, Abeba Birhane notes that some “algorithmic harms may be secondary effects, invisible to designers and communities alike” (Birhane 2021, 129). The challenge of detecting harmful effects leads Birhane to remark, “What questions might be asked to help anticipate these harms?” (Ibid.). The fact that the harmful impacts arising from the deployment of AI are not always immediately apparent (Ibid.) makes it difficult for human rights defenders to document the harms in question. Consequently, it is challenging for human rights defenders to call on states to uphold the enjoyment of human rights when organisations deploy AI as part of the decision-making process.

Norms of international human rights law do not make it possible to render the harms arising from the use of AI visible. They merely provide guidance on what types of harm are relevant for the purpose of establishing whether there is a human rights violation. Consider the International Covenant on Economic, Social and Cultural Rights (hereinafter ICESCR). Article 2(1) of this treaty imposes obligations on states to safeguard the protection of human rights (ICESCR 1966, Art. 2(1)). Article 2(2) stipulates that states which are party to this treaty should take measures to guarantee that individuals can exercise the rights enshrined in this treaty without “discrimination of any kind” (Ibid., Art. 2). Article 12(1) elaborates that states which are party to this treaty “recognise the right of everyone to the enjoyment of the highest attainable standard of physical and mental health” (Ibid., Art. 12(1)).

Knowing that states recognise the right to the highest attainable standard of mental health (Ibid.) does not tell human rights defenders anything about the impact that AI's screening of candidates for employment (Parker 2023) has on the candidates' mental wellbeing. Another challenge is that international human rights law norms assume that the harm is visible to an individual and that the individual experiences the harm in an embodied fashion. However, as

Birhane explains, this is not necessarily true when using AI technology (Birhane 2021, 129). Birhane's scholarship points to the fact that the human rights community needs new approaches that will enable it to detect harmful impacts arising from the deployment of AI that are not obvious or readily visible (Ibid.). This article will demonstrate that the theoretical framework of media ecology can help human rights advocates to detect some harmful impacts on the mental wellbeing of the applicants arising from the employment of AI to select candidates for employment (Parker 2023) when such impacts are not readily apparent. Individuals may not immediately feel these harms.

The article will use the impact on the candidates' mental wellbeing arising from selecting candidates for employment using AI (Ibid.) as a case study. The article uses this case study because it is not apparent to an average person how the employment of AI as part of the decision-making process affects the mental wellbeing of the candidates. Moreover, there is limited literature on this issue. Another reason for the choice of this case study stems from the fact that Articles 2(1) and 12 of ICESCR oblige states to take steps "to the maximum of available" resources in order to achieve "progressively the full realisation of the right" (ICESCR 1966, Art. 2(1)) to the enjoyment of the highest attainable standard of mental health (Ibid., Art. 12). There is thus a connection between the right to mental health and the impact of an employment screening process on the mental wellbeing of the applicants.

This article is of interest to both the human rights community and media ecologists. Human rights advocates will learn that the employment of AI as part of the decision-making process to screen the applicants for employment (Parker 2023) implicates the right to mental health in Article 12(1) ICESCR. They will become acquainted with a possible approach for using the media ecology framework to identify harms that may not be readily apparent. As a result of having more knowledge about harms that are not readily visible and which may manifest themselves over a longer term, human rights advocates can better articulate how the harmful impact in question maps onto a particular norm of human rights law. In light of this knowledge, they can launch advocacy campaigns to urge governments to address the harms in question. Media ecologists will find this article interesting because they will learn about the value of the theoretical framework of media ecology for law.

The article has the following structure. In order to provide context for the discussion, the article will explain how it defines AI in section 2. Section 3 will discuss why it is timely to scrutinise what impact the employment of AI as part of the decision-making process in the employment context (Ibid.) has on the mental wellbeing of the applicants. Currently, there is limited literature on this topic. Section 4 will introduce the theoretical framework of media ecology. It will put forward why the theoretical framework of media ecology is a promising avenue of inquiry for identifying how employing AI decision-making processes impacts on the job applicants' mental wellbeing, which may not be readily apparent. Section 5 uses the theoretical framework of media ecology to establish how the use of AI to screen candidates for employment (Ibid.) results in applicants experiencing poor mental wellbeing. Section 6 covers how the findings can inform the advocacy efforts of the human rights defenders. It distils lessons that the human rights community can learn from the present discussion.

## **II. Defining AI**

Defining what AI is at this stage is essential to set the context for the discussion. It is crucial to acknowledge that understanding what constitutes AI has been evolving (Sheng Loe 2018, 5180). Moreover, the terminology relating to AI is contested (Jones 2023). Numerous definitions of AI exist (Sheng Loe 2018, 5180). Furthermore, the labels people use to describe different types of AI are contested (Jones 2023). The Ada Lovelace Institute anticipates that the definitions of AI and the different types of AI will continue to evolve (Ibid.). This diversity of definitions of AI (Ibid.) necessitates providing examples of different definitions of AI and explaining why the present article adopts a particular definition of AI for the present analysis.

The definition of the Australian Human Rights Commission focuses on the programming tools that computer scientists use to create such systems. The Australian Human Rights Commission defines AI as “a cluster of technologies and techniques, which include some forms of automation, machine learning, algorithmic decision-making and neural network processing” (Australian Human Rights Commission 2021, 17). In order to understand this definition, it is necessary to distinguish between different terminology relating to AI. Machine learning is a subfield of study within artificial intelligence (Brown 2021). Machine learning underpins many AI applications (Gajjar 2023, 12). Systems operating on machine learning

detect patterns in the data (Brown 2021). As a result of detecting such patterns (Ibid.), the AI generates a model of the external environment (Gajjar 2023, 12). Subsequently, when one inputs new data into AI relating to an applicant, the AI maps this new data onto its model of the external environment to make predictions about people or to generate other useful outputs. (The Alan Turing Institute 2024; Gajjar 2023, 12) The definition of the Australian Human Rights Commission is sufficiently broad to encompass different applications of AI. This is the case because the definition focuses on the programming techniques, such as machine learning, which computer scientists employ to create such systems (Australian Human Rights Commission 2021, 17).

The European Union took a similar approach to the Australian Human Rights Commission when defining AI for the Artificial Intelligence Act. In 2021, the European Commission circulated the draft Artificial Intelligence Act to address the governance of AI technology in the European Union context (Council and European Parliament Draft Regulation COM (2021) 206, 1). The purpose of the draft Artificial Intelligence Act is to ensure that the use of AI technology complies with the fundamental rights and European values (Artificial Intelligence Act 2024, Art. 1). The January 2024 version of the draft Act defines AI in Article 3(1) as a “machine-based system designed to operate with varying levels of autonomy” (Ibid., Art. 3(1)). It uses input data to make inferences to generate “predictions, content, recommendations, or decisions” (Ibid.). Such systems can exhibit the trait of being adaptable after organisations begin to use them (Ibid.). This definition is very broad because it focuses on what the AI does rather than the computer science techniques underpinning such systems. What is common to the techniques used for creating AI software is that the software applies statistical techniques to the input data (Agarwal 2020) to create predictions, make recommendations and produce decisions (Organisation for Economic Co-operation and Development 2019, 1).

The United Kingdom, when defining AI in its 2023 policy paper “a pro-innovation approach to AI regulation,” emphasises the characteristics of this technology, which require a distinct approach to regulation (Department for Science, Innovation and Technology 2023, 22). Unlike the Australian Human Rights Commission (Australian Human Rights Commission 2021, 17), the United Kingdom does not focus on the programming tools the developers use to program

AI as a key characteristic for defining these systems. The whitepaper defines an AI system or technology as “products and services that are ‘adaptable’ and ‘autonomous’” (Ibid., 13). Autonomy refers to AI's ability to generate decisions without ongoing human control (Ibid., 22). Adaptability refers to the fact that AI can draw fresh inferences from new data after being trained to recognise patterns in the data (Ibid., 22). There is a similarity between the definitions of AI in the UK policy paper (Department for Science, Innovation and Technology 2023, 13) and the draft Artificial Intelligence Act (Artificial Intelligence Act, Art. 3(1)). They do not treat the programming techniques as the core characteristic of AI.

Despite the differences between the approaches that the United Kingdom and the Australian Human Rights Commission take to defining AI, there is an overlap between their two respective approaches. This similarity between the definitions of AI stems from the fact that the techniques that computer scientists employ to program AI influence how AI operates and generates outputs. The AI software will exhibit unique characteristics associated with the techniques used to program AI during its operation. The difference between the definitions in the UK policy paper and the Australian Human Rights Commission is more about form than substance at this point. While the Australian Human Rights Commission makes programming techniques specific to AI explicit (Australian Human Rights Commission 2021, 17), the United Kingdom policy paper implies that the technology relies on employing these programming techniques (Department for Science, Innovation and Technology 2023, 22).

In addition to noting that there has been a proliferation of terminology within AI (Jones 2023), it is essential to acknowledge variations between different types of AI. For example, generative AI is software that can generate video, text and other content based on user inputs, such as text prompts (Ibid). Some but not all generative AI software uses foundational models (Ibid.). A distinct feature of foundational models is that they can perform various tasks (Ibid.). In contrast, “narrow AI” models are designed to perform a single purpose or task (Ibid.). AI, which uses a foundational model, can take input in multiple formats and generate multiple outputs (Ibid.). For instance, it can generate images and provide answers to questions (Ibid.). Foundational models allow companies buying AI to customise the software to their needs by adding new data and fine-tuning the software's performance to a specific task (Ibid.).

This article uses the definition of AI from the United Kingdom Government Data Ethics



Framework. This document defines AI as systems capable of performing tasks that human beings traditionally thought to require intelligence (Central Digital and Data Office 2020). The United Kingdom Government Data Ethics Framework acknowledges that AI is evolving as a technology (Ibid.). The document specifies that AI “involves machines using statistics to find patterns in large amounts of data” (Ibid.). Such machines can “perform repetitive tasks with data without the need for constant human guidance” (Ibid.). The article uses this definition of AI because this definition acknowledges the evolving nature of AI technology (Ibid.).

Furthermore, the article uses the definition of AI in the United Kingdom Government Data Ethics Framework because the article focuses on examining the impact of using software that employs a particular set of processes in order to generate predictions about the future performance of job applicants and to produce decisions which applicants to select for employment (Organisation for Economic Co-operation and Development 2019, 1). The key feature of these processes is that they use statistical techniques (Agarwal 2020) and the logic of optimisation (Badar, Umre and Junghare 2014, 39).

It should be noted that it is beyond the scope of this article to include all possible applications of AI. Generative AI is an example of a type of AI that is excluded. The article encompasses both narrow AI and foundational models within its scope. What is crucial to the inquiry is the impact of using statistical techniques and the logic of optimisation as a central component of the decision-making process. Given this focus of inquiry, it is immaterial whether an organisation can add new data to AI after purchasing it or whether AI can take data inputs in different formats. Consequently, the article includes the employment of AI as part of the decision-making process to screen the applicants for employment (Harlan and Schnuck 2021), irrespective of whether such systems use foundational models or narrow AI.

The products enabling the human resources staff to use AI to screen applicants for employment (Parker 2023) already exist. For example, HireBee offers AI, which can identify the most qualified candidate who fits the client company’s culture more closely (Ibid.). The software achieves this by analysing the candidates’ resumes, cover letters and online profiles

(Ibid.). The developers of AI-based software Sapia.ai promise employers that this software can evaluate the candidates' soft skills, behaviour traits and cognitive ability based on the written answers that the applicants provide to a handful of questions (Sapia.ai, 2024). Companies also claim that AI can analyse recorded videos that the applicants made, draw inferences about the candidates' personality traits and evaluate the candidates' suitability for the job role (Harlan and Schnuck 2021). The discussion and findings in this article are relevant to all of these applications of AI in screening candidates for employment. As was already explained, statistical techniques (Agarwal 2020) and the logic of optimisation underpin AI (Badar, Umre and Junghare 2014, 39).

### **III. The Need to Study the Impact of Using Artificial Intelligence to Select Candidates for Employment on the Mental Wellbeing of the Applicants**

It is necessary to study the impact of the use of AI to select candidates for employment (Parker 2023) on the mental wellbeing of job applicants for two reasons. First, there is evidence that the employment of some types of digital technologies can have a negative impact on individuals' wellbeing. Existing research demonstrates that there is a connection between Instagram use and the mental health of teen girls (Milmo 2021). Girls describe feeling anxiety, depression and poor body image after regularly interacting with Instagram (Gayle 2021). Psychologist Phil Reed and colleagues found that engagement with social media using visual forms has a discernible relationship with the acquisition of narcissistic traits (Reed et. al. 2018, 168). Similarly, media ecologist Sérgio Roclaw Basbaum talks of individuals experiencing the narcissistic pleasure of seeing themselves reflected in a media device when they use WhatsApp, twitter and social networks (Basbaum 2022, 79-80). However, researchers in psychology disagree over whether there has been a growth in the number of individuals who exhibit narcissistic traits since the 1980s (Jarrett 2017).

The connection between the use of some digital technologies and poor mental wellbeing (Milmo 2021) makes it necessary to establish whether the employment of AI as part of the decision-making process in the employment context (Parker 2023) negatively affects the applicants' mental health. Another reason why this inquiry is timely stems from the fact that there is a dearth of literature on how using AI as part of the decision-making process could impact the mental wellbeing of job applicants. Preliminary warning signs hint that there could

be a connection between the employment of AI as part of the decision-making process in employment (Ibid.) and poor mental wellbeing. A legal scholar, Karen Yeung, argues that the use of AI data-driven techniques to personalise delivery of content and services online will foster a culture of narcissism (Yeung 2018, 268). Yeung's research raises the question of whether the employment of AI to select candidates for employment (Parker 2023) creates an environment conducive to individuals experiencing or exhibiting the traits of what are known as personality disorders.

It is at this stage necessary to explain what the article means here by the term personality disorder. Personality disorders are "patterns of thinking, perceiving, reacting and relating that cause significant distress or functional impairment" to the affected person (Zimmerman 2023). The article will use the term personality disorder here to refer to internal processes that are conducive to individuals experiencing a lack of mental wellbeing. This approach to a definition is based on the fact that personality disorders involve behaviours that are damaging to the individual (Adshead and Sarkar 2012, 164). The article treats a personality disorder as a concept that society created and that is shaped by culture (Bjorklund 2006, 11) rather than as a deviation from normal ways of being. First, culture determines what society treats as normal and as a psychological disorder (Kirmayer and Young, 1999, 447). The role of culture in shaping how societies define what is normal results in Pamela Bjorklund drawing a conclusion that a personality disorder is a socially created construct (Bjorklund 2006, 11). Thomas Szasz goes even further when he remarks that mental illness is a myth (Szasz, 1974, 4). He believes this concept "has outlived whatever usefulness it might have had" (Ibid.).

The second reason for treating a personality disorder as a lack of mental wellbeing rather than as a deviation from the norm stems from the fact that individuals can exhibit traits of a personality disorder without having the personality disorder (Heller 2020; American Psychiatric Association 2022, 737). What matters for diagnosis is the intensity and the persistence of the traits, as well as whether these traits cause "significant" distress to the person experiencing these traits (Heller 2020; American Psychiatric Association 2022, 737). All individuals have some characteristics of narcissism, for instance (Heller 2020). In this sense, all individuals are on a personality disorder spectrum (Ibid.). Indeed, Jeremy Coid, Peter Tyrer and Min Yang

found in a study that only 23% of the population had “no evidence of personality disturbance” (Coid, Tyrer and Yang 2010, 194). Furthermore, individuals can develop degrees of severity of what are known as personality disorders (Ibid., 196). Given the role of culture in determining when an individual is manifesting a psychological disorder (Kirmayer and Allan Young, 1999, 447) and that all individuals are on a spectrum (Heller 2020), it is more appropriate to view a personality disorder as a lack of mental wellbeing rather as a deviation from the norm.

A potential counterargument to understanding a personality disorder as a social construct (Bjorklund 2006, 11) capturing a constellation of internal experiences that negatively affect an individual’s mental wellbeing (Zimmerman 2023) is that some scientists claim that there are differences in the brains of individuals who do not have a particular personality disorder and between individuals who do (Chapman 2019, 1151). It is at this stage necessary to note that this claim is contested in the scientific community (Ibid.). A response to this counterargument is that individuals necessarily exercise value judgment in determining whether an individual meets one or several diagnostic criteria for a personality disorder (Adshead and Sarkar 2012, 163). The diagnostic criteria for personality disorders can be interpreted in many different ways because they are disconnected from the context in which an individual is situated (Zandersen and Parnas 2019, 112). The role of the subjective judgment of the clinician in construing whether an individual meets the diagnostic criteria (Adshead and Sarkar 2012, 163) supports Bjorklund’s conception of a personality disorder as a social construct (Bjorklund 2006, 11). Moreover, the fact that individuals can change how they relate to themselves and the world (Luyten et al 2018, 98) supports the view that a personality disorder is a social construct (Bjorklund 2006, 11). Once individuals change the mechanisms of relating to themselves and the world, which are unhelpful to them, they no longer exhibit the traits of the personality disorder (Luyten et al 2018, 98).

Finally, Allan Horwitz’s account of the changing character of what the American Diagnostic and Statistical Manual of Mental Disorders (hereinafter DSM) calls disorders (Horwitz 2021, 149-150) arguably supports viewing a personality disorder as a social construct (Bjorklund 2006, 11). Allan Horwitz calls DSM-5 a social document because culture, politics, economics and how psychiatry developed shaped it (Horwitz 2021, 144-145). He describes intra-professional disputes as influencing the content of the diagnostic criteria (Ibid., 147). For example, drafters removed homosexuality from the DSM following public campaigning (Ibid., 149). They kept

gender dysphoria as a diagnosis because individuals who wished to undergo the gender reassignment procedure wanted to have free access to the gender reassignment procedures (Ibid., 150).

Overall, the range of diagnosis grew over time to encompass behaviours that were previously seen as bad habits (Ibid., 152). The psychiatric profession was embarrassed in the early 2000s when some parents began to seek diagnoses of their children with success in order to manage disruptive or oppositional behaviour with medication (Ibid., 155). This expansion of diagnosis in the DSM led Karl Menninger to remark, “most people have some degree of mental illness at some time, and many of them have a degree of mental illness most of the time” (Ibid., 153; Menninger 1963, 33). Other critics maintain that the labels in the DSM hide a lack of knowledge about the reasons for poor mental wellbeing (Horwitz 2021, 157). Another reason for treating personality disorders as a lack of wellbeing rather than as concrete diagnoses in this article stems from the controversial nature of some diagnoses, such as borderline personality disorder (Perrotta 2020, 47).

#### **IV. Why the Theoretical Framework of Media Ecology is Useful for Understanding the Impact of Employing Artificial Intelligence as Part of the Decision-Making Process on the Job Applicants' Mental Wellbeing**

Christine Nystrom defines media ecology as:

the study of the ways in which our instruments of knowing — our senses and central nervous systems, our technologies of exploration, the physical media they require (like light, sound, electricity), and the conditions in which they are used — construct and reconstruct what we know, and therefore the realities that humans inhabit (Nystrom, Wiebe and Maushart 2021, 108).

Media ecology is arguably a valuable lens for contributing knowledge about some of the impacts of AI to screen applicants for employment (Parker 2023) on people's mental wellbeing. Media ecology enables one to analyse the impact of technology in a comprehensive manner.

This is the case because this approach entails studying the content of the media, the nature of the media and the total cultural environment within which such media function (McLuhan 1969). In particular, media ecology considers how communication media, language and technology comprise an ecology with shifting and interactive elements (Logan 2007, 16). When communication media, language and technology interact, the ecology in which they are situated evolves (Ibid.).

Since media ecology is concerned with exploring how media bring about change (McLuhan and Fiore 1967, 25), media ecology is a suitable theoretical framework for examining how the employment of AI to select individuals for employment (Parker 2023) impacts the psyche of the applicants. Specifically, Jerome Agel, Marshall McLuhan and Quentin Fiore wrote that by changing the environment, media invoke “unique ratios of sense perceptions in people” (Agel, McLuhan and Fiore 1996, 41). By extending a particular sense, technology changes how people perceive, think and act (Ibid.). This scholarship points to the fact that media ecology offers the tools for exploring how the employment of AI as part of the decision-making process in the employment context (Parker 2023) impacts the applicants’ wellbeing.

Additionally, it is put forward that Marshall McLuhan’s approach is beneficial for studying the effects of the use of AI decision-making processes on job applicants’ mental wellbeing because he does not use a fixed theory for making sense of a complex social reality (McLuhan 1969). McLuhan explained, “I’m making explorations. I don’t know where they’re going to take me. My work is designed for the pragmatic purpose of trying to understand our technological environment and its psychic and social consequences” (Ibid.). He saw his scholarship as enabling one to recognise “patterns” and to “map new terrain” rather than to apply fixed categories to make sense of the social reality (Ibid.). Avoiding imposing a superstructure lens to make sense of reality means that researchers begin their inquiry with fewer assumptions. As a result of not trying to make the reality fit a particular framework, the researchers may see new patterns of connection emerge (Ibid.). Since McLuhan’s approach to analysis is designed to make it possible to detect patterns in social reality (Ibid.), this approach to inquiry enables one to learn new information about processes that are already in progress.

Finally, McLuhan’s approach is useful because it allows one to make what can otherwise remain invisible visible (McLuhan 1969). McLuhan said that during the use of new technology,

people are only aware of the environment that preceded the development of this technology (Ibid.). McLuhan claimed, "The present is always invisible because it's environmental and saturates the whole field of attention so overwhelmingly" (Ibid.). He then explained that only a person with an "integral awareness," such as an artist, can reveal the environment that the new technology brings about (Ibid.). McLuhan, therefore, suggested that one needs to be creative in order to uncover the changes in the environment.

By opening a space for creativity (Ibid.), McLuhan allows one to depart from the traditional approaches to reaching logical conclusions (Matie, McLuhan and Toye 1987, 478). Scholars traditionally looked at causes and matched them with effects (Ibid.). In contrast, McLuhan first analysed the effects and then traced the effects to the causes (Ibid.). McLuhan's break from the traditional approach to analysis (Ibid.) is particularly useful for examining the effects that the deployment of AI produces. With complex systems, such as AI, one cannot trace a direct cause-and-effect relationship between the AI's performance and the societal impacts (Krupiy 2021a, 15; Birhane and Sumpter 2022, 5). McLuhan's approach allows one to establish connections between the changes taking place in the environment and the employment of a technology (Agel, McLuhan and Fiore 1996, 41) without the need to demonstrate a linear or direct cause-and-effect relationship. This aspect makes McLuhan's approach particularly valuable for the case study at hand. Additionally, as was already mentioned, McLuhan's approach is particularly well-suited for studying the impact of the use of AI as part of the decision-making process in the employment context (Parker 2023) on the mental wellbeing of the applicants because it is already known that the effects arising from the deployment of AI are not always readily visible (Birhane 2021, 129).

A possible counterargument is that McLuhan's approach to analysis does not yield results with empirical validity (Bates 2011). The response to this argument is that Tony Bates confirmed the validity of some of McLuhan's claims using observation (Ibid.). For example, McLuhan claimed that the "medium is the message" (Matie, McLuhan and Toye 1987, 443). Bates reports that students react to the same knowledge differently depending on whether they learn about it by observing video or reading print (Bates 2011). Since individuals can learn about the world by observing (Ibid.) and since McLuhan's approach relies on exploration (McLuhan

1969), McLuhan's approach to analysis yields valuable insights. Additionally, this article draws conclusions that have weight because it uses literature from social sciences to supplement the media ecology analysis. Finally, the article makes credible conclusions because the author closely examines how AI works and analyses what consequences arise from the manner in which AI operates.

Building on this discussion, studying AI as a medium using the theoretical framework of media ecology exposes how AI creates an environment that, in turn, creates conditions favourable for bringing about transformations in people's mental wellbeing. The application of media ecology as an approach to inquiry makes it possible for human rights defenders to examine what changes the employment of AI as part of the decision-making process in the employment context (Parker 2023) brings about in the environment. The human rights defenders can then scrutinise how these changes in the environment affect the mental wellbeing of the job applicants. In light of this knowledge, one can then put forward how one can use international human rights law norms to address the harms in question.

It will now be explained what concepts from media ecology are relevant for studying how the use of AI to screen candidates for employment (Ibid.) affects the mental wellbeing of the applicants. McLuhan defines media as everything that extends a psychic or a physical faculty (McLuhan and Fiore 1967, 25). The use of AI to screen an applicant for employment (Parker 2023) extends the computational capabilities of individuals. AI enables individuals to compute larger volumes of information faster (Moore 2019). Thus, the AI decision-making tools that organisations use in the employment context (Parker 2023) are media. McLuhan's writings suggest that this acceleration in the processing of information affects the senses of the operators and of the subjects of the decision-making (McLuhan and Fiore 1967, 41). McLuhan saw all media as extending a sense and altering the way in which individuals perceive the world (Jerome, McLuhan, and Quentin 1966, 41). This paper only considers the impact of using AI on applicants due to space constraints.

McLuhan was aware of the co-constitutive nature of humans and the media they design (Coupland and McLuhan 2009, 107). This aspect is captured in the aphorism often attributed to him: "[W]e shape our tools, and they in turn shape us" (Ibid.). McLuhan states that "[P]sychic change [results] from [human]-made or technological environments" (Ibid., 458). McLuhan



elaborates that complex changes occur in the object of study and the environment during all communication (Ibid., 467). Media transform how individuals perceive, think and act by modifying the ratios between the senses through which individuals perceive the environment (McLuhan and Fiore 1967, 41). As McLuhan said, the media “[W]ork us over completely” (Ibid., 25). Since McLuhan drew a link between technology and the psychic changes that the use of technology triggers (Ibid., 41), it is possible to apply McLuhan’s scholarship in order to study some of the changes that the use of AI to select candidates for employment (Parker 2023) creates in the environment. Equally, it is possible to scrutinise how such changes can impact the applicants’ mental wellbeing.

It is suggested that what makes McLuhan’s approach to analysis particularly valuable for exploring the psychic changes that AI employment screening tools produce is that he acknowledges that it can be difficult for individuals to detect the psychic changes that the employment of technology brings about (McLuhan 2013, 34-35). He describes the body as cutting off access to some of the senses in response to the acceleration that technology brings about (Ibid., 34-35). Consequently, the technology can numb human perception (Ibid., 34). In effect, one loses “conscious awareness” of how technology impacts oneself (McLuhan 1969). Having this awareness led McLuhan to develop particular concepts to make it possible for the analyst to detect the changes a technology produces in the environment (Matie, McLuhan and Toye 1987, 415). Since the harmful impacts of AI are not always readily visible (Birhane 2021, 129), the concepts McLuhan developed are particularly suitable for examining such effects.

In order to establish the impact that using AI as part of the decision-making process to screen applicants for employment (Parker 2023) has on how the applicants perceive and act, this article uses McLuhan’s concept known as the figure/ground analysis. McLuhan developed the figure/ground approach to analysis to account for the fact that one can only understand an object of study in the context of the system or environment in which it operates (Logan and Rawady 2021, Preface p 6). The ground of any technology consists of two elements (Matie, McLuhan and Toye 1987, 408). The first element is the situation that gave rise to the technology (Ibid.). The second element is the entire environment that the technology creates as part of producing effects (Ibid.). The ground and the effects of technology are hidden (Ibid.,

477). The figure is the object of the study (Ibid., 194).

McLuhan started the analysis by analysing changes that take place in the ground or the environment as a result of the technology operating (Ibid., 473). He then traced such effects to the figure, namely to the medium/technology (Ibid., 473). McLuhan used probes or hypotheses to study the technology's effects in the environment (Logan and Rawady 2021, Preface p 6). He relied on observation to detect patterns of change in the environment, which the employment of technology brings about (McLuhan 1969). McLuhan created a probe or hypothesis to explain what changes in the environment were taking place (Ibid.). McLuhan would then use different approaches to test the validity of his hypothesis because he was not committed to any particular theory (Ibid.). McLuhan believed that the figure/ground approach to analysis made it possible to evoke a new perception or understanding of the effect of a technology in the environment (Matie, McLuhan and Towe 1987, 415).

Turning to the case study at hand, the employment of the figure/ground mode of inquiry suggests that one needs to begin the analysis by asking what effects the use of AI as part of the decision-making process to screen the applicants for employment (Parker 2023) produces in the environment. Robert Logan studied the impact of digital media on society by applying the concepts of figure and ground (Logan 2020, 10). Since AI is a type of digital technology (IBM 2024), his analysis is relevant to the present discussion. Logan argues that users operate as a figure in the context of digital media (Logan 2020, 10). Meanwhile, digital media perform the role of information ground because they rely on user-generated data to have the capacity to function (ibid.). In addition to being a figure, the users function as the ground for the digital media (ibid.). This is the case because the users provide information for the digital media (Ibid.). The digital media modify their algorithms based on such data (ibid.). Logan concludes that the user and the digital medium operate both as a figure and a ground for one another (ibid.).

Building on the work of Robert Logan, the present author put forward that the subject of the decision-making and the AI decision-making process “operate as both figure and ground in relation to one another” (Krupiy 2021b, 36). The subject of decision-making and the social context serve as the ground for the AI decision-making process (Ibid.). This is the case because the use of the AI decision-making process transforms how the applicants construct

and interpret their identities (Ibid., 37). It achieves this by generating a technologically mediated identity for each applicant and communicating the application decision to the applicant (Ibid., 37). AI generates a technologically mediated identity for each applicant by combining information about the applicant with information about the other applicants (Ibid.). The AI decision-making process uses information about a group of individuals it treats as being similar to the applicant to predict the applicant's future behaviour or performance (Provost and Fawcett 2013, 21; Taylor 2018, 105; Taylor 2017, 14-15). The AI produces an output or a decision based on this process (Provost and Fawcett 2013, 21; Taylor 2018, 105; Taylor 2017, 14-15). Although AI produces a decision about the applicant based on group data (Provost and Fawcett 2013, 21; Taylor 2018, 105; Taylor 2017, 14-15), it purports to inform the applicant about the applicant's characteristics (Krupiy 2021b, 37). The applicants internalise this information because they interpret it as relating to aspects of their personal characteristics or past performance (Ibid.). The AI decision-making process acts as figure because it is producing changes in the environment by shaping how individuals construct their identities (Ibid.).

Simultaneously, the subject of decision-making serves as a figure for the AI decision-making process due to producing effects on the system's model of the external environment (Logan 2020, 10). This is the case because AI modifies its model of the external environment based on the information it receives relating to each job applicant (Ibid.). Thus, the AI decision-making process serves as a ground for the subject of the decision-making (Ibid.). The concept of the figure/ground will be used below to analyse some of the ways in which the employment of AI to screen candidates for employment (Parker 2023) brings about changes in the environment. Additionally, this approach to analysis will be employed below to study what aspects of these changes are likely to trigger psychic transformations in the applicants' psyches.

## **V. Using Media Ecology to Study Some Impacts of Using Artificial Intelligence as Part of the Decision-Making Process on the Mental Wellbeing of the Job Applicants**

McLuhan did not explicitly discuss the relationship between the use of technology and

individuals either acquiring or exhibiting the traits of personality disorders. He did not engage with the topic of technology and mental health in depth. However, McLuhan implicitly addressed the issue of how technology impacts mental health with a degree of detail. He drew a connection between the myth of Narcissus and how the body reacts to being extended by any technology (McLuhan 2013, 34). Chiara Blanco argues that the character of Narcissus in Ovid's myth exhibits the traits of what is now known as narcissistic personality disorder (Blanco 2023). It follows that in discussing what the Narcissus myth can tell people about their relationship with technology (McLuhan 2013, 34), McLuhan made an implicit connection between technology, what is known as narcissistic personality disorder and mental health.

The nature of this connection will now be explored in more detail. McLuhan argues that the Narcissus myth conveys to society that individuals remain "unaware of the psychic and social effects" of a new technology, much like a fish is unaware of the water it swims in (McLuhan 1969). Specifically, when exposed to technology, the body becomes numb (McLuhan 2013, 35). It blocks perception by isolating a sense, a function, or an organ (McLuhan 2013, 35). The purpose of this numbing process is to provide the body with an "immediate relief from strain" that the technology triggers in it (Ibid.). Consequently, McLuhan concludes that technology is an extension of our physical bodies, which induces the body to engage in an act of self-amputation (Ibid., 36). As a result of this process of numbing and self-amputation, any technology changes the ratios among other organs (Ibid., 36). For example, when a technology amplifies the sense of sound perception, it affects the senses of touch, taste and sight (Ibid., 35).

Olivia Harvey interprets McLuhan's discussion of the myth of Narcissus as meaning that technology "mediates the production of subjective identity" of people (Harvey 2006, 336). Harvey's interpretation should be viewed as valid. McLuhan's description of the altering of the ratios between the organs due to exposure to technology (McLuhan 2013, 36) parallels how the DSM-5 views narcissism. The revised DSM-5-TR treats narcissism as being characterised by changes in self-functioning (American Psychiatric Association 2022, 761). For instance, such individuals can view themselves as special (Ibid.). They can feel devastated when other individuals do not give them such recognition (Ibid.). One of the criteria for narcissism in DSM-5 is that an individual makes an excessive reference to the opinions of others to regulate and maintain one's own self-esteem by having a need for admiration (Ibid., 760). Thus, the DSM-5

associates narcissism with an inappropriate balance between the weight that an individual places on one's own reasoning processes and the external environment in order to maintain a sense of self-esteem. While McLuhan focuses on the change in the ratio between the senses (McLuhan 2013, 35-36), the DSM-5 treats narcissism as a change in the ratio of weight an individual places on internally and externally generated opinions. Notwithstanding this difference in approach, the DSM-5 can be described as capturing the change in ratio between the senses. This is the case because one uses aural or visual sense to a greater extent when one relies on the expression of opinions of other people to maintain a sense of self-esteem as opposed to an internal reasoning process. Consequently, what the DSM-5 calls a narcissistic trait refers to greater reliance on using aural or visual sense to construct one's own identity.

It is arguably possible to establish another link between the diagnostic criteria in the DSM-5 and McLuhan's analysis of what the Narcissus myth tells people about how they relate to technology (Ibid., 34). McLuhan alludes to Narcissus to describe how individuals relate to technology by being its "servomechanism" (Ibid., 36). Individuals need to serve the technology in order to use it (Ibid.). Since exploitation of others is one of the traits of narcissistic personality disorder (American Psychiatric Association 2022, 760), McLuhan appears to have drawn a connection between the exploitative dimensions that the use of technology entails (McLuhan 2013, 36) and a narcissistic trait. In particular, he reveals that the extractive aspects of technology (Ibid.) become normalised and invisible to their users. By establishing this connection, McLuhan draws an implicit link between the employment of technology and narcissism.

McLuhan's observation that individuals need to serve the technology to use it (Ibid.) is particularly relevant in the digital age. According to Briny Blackmore and colleagues, users lack awareness that they are in an exploitative relationship with the digital platforms that they use (Blackmore 2023, 446). The users provide data about themselves while using the services, such as social media, email accounts and accessing web pages (Ibid.). The companies use this user-generated data to reap "significant economic value" (Ibid., 446-447). However, the users do not know what value the companies generate by getting access to their data and their contribution to this value (Ibid.).

The employment of AI as part of the decision-making process is similarly extractive due to the use of the applicants' data (Crawford 2021). The employment of AI enables companies to derive "population-level" insights using group data (Viljoen 2021, 1). It is suggested that McLuhan captured the fact that the extractive dimensions of AI technologies can remain invisible to the users (Blackmore 2023, 446). He talks of individuals having "subliminal awareness" about the images of themselves in technologies (McLuhan 2013, 36). Additionally, although McLuhan never expressly addressed the issue, he arguably captured the fact that the possession of narcissistic traits is ubiquitous within the population (Heller 2020). McLuhan describes individuals relating to technology as a "servomechanism" (McLuhan 2013, 36). It is maintained that if the possession of narcissistic traits had not been ubiquitous within the population (Heller 2020), those developing and using AI would not have unleashed a technology that relies on extractive practices to function (Crawford 2021). It follows that although McLuhan did not connect the myth of Narcissus to narcissism, his scholarship on the relationship between technology and human users (McLuhan 1969) is very informative for how technology impacts peoples' mental wellbeing. McLuhan's scholarship arguably continues to have descriptive value for depicting the extractive and invisible dimension of AI in relation to the users (Blackmore 2023, 446).

This discussion raises the question of what media ecology can tell us about the relationship between the use of AI as part of the decision-making process in the employment context (Parker 2023) and changes in the environment. Equally, the issue is what light media ecology can shed on how these changes affect the mental wellbeing of the job applicants. On the application of McLuhan's figure/ground analysis, the employment of AI as part of the decision-making process in hiring (Ibid.) changes the environment. As has already been shown, the AI decision-making process transforms how applicants construct and interpret their identities (Krupiy 2021b, 37). It will now be demonstrated that these changes in the environment are conducive to individuals experiencing or exhibiting the traits of what are known as numerous personality disorders. Using computational decision-making creates new social and cultural conditions affecting peoples' experiences. These experiences are related to suboptimal wellbeing.

## **A. AI And The Traits of The Obsessive-compulsive Personality Disorder**

The use of AI in hiring (Parker 2023) creates an environment that is conducive to individuals exhibiting some of the traits of obsessive-compulsive personality disorder. The characteristics of obsessive-compulsive personality disorder are that the individual is preoccupied with orderliness (American Psychiatric Association 2022, 771). The individual controls one's interpersonal processes (Ibid.). The person pays very close attention to rules and strives for perfectionism (Ibid.). Such individuals exhibit less flexibility and are not open to multiple approaches to completing a task (Ibid., 773). The logic of optimisation underpinning AI decision-making processes (Badar, Umre and Junghare 2014, 39) creates an environment where individuals are "nudged" to strive for perfectionism, view only one approach as right and exhibit a particular behaviour. A nudge involves using regulatory tools to design a decision-making environment in a way that is conducive to the user making particular choices (Teichman and Zamir 2020, 1266). Shoshana Zuboff mentions that big Western companies manufacturing digital technologies, such as driverless cars, design the technologies in a way as to nudge the users to behave in a way which benefits the company's financial interests (Zuboff 2018, 177). The present discussion suggests that Zuboff's argument that some digital technologies nudge users to behave in a particular way (Ibid.) is applicable to the context of the employment of AI decision-making processes to select candidates for employment (Parker 2023).

McLuhan's claim that the "medium is the message" (Matie, McLuhan and Tøye 1987, 443) makes it possible to unpack this argument. McLuhan argued that the message that a medium transmits consists of all of the effects that the medium produces in the environment (Ibid., 448). The audience is the content of the message because the message acquires meaning in the course of the audience interpreting it (Ibid., 443). McLuhan believed that complex changes occur in the figure and the ground during all communication (Ibid., 467). As was already mentioned, the subject of the decision-making process serves both as a ground and as a figure for the AI decision-making process (Krupiy 2021b, 36; Logan 2020, 10). Reasoning by analogy, the message of the AI decision-making process consists of the traits of a personality disorder. In particular, it is put forward that the logic of optimisation underpinning the use of the AI decision-making processes (Badar, Umre and Junghare 2014, 39) creates an environment

that nudges the applicants to exhibit the traits of obsessive-compulsive personality disorder. One such trait is striving for perfectionism (American Psychiatric Association 2022, 771). The AI decision-making process allocates positive decisions to individuals with the highest score (Barocas and Selbst 2016, 678-679). The emphasis on receiving the highest score (Ibid.) nudges the applicants to strive for perfectionism.

Additionally, the operation of the AI decision-making process creates an environment that nudges the job applicants to view a specific approach to performing a task as being the right one and to control their interpersonal processes. This is the case because an AI decision-making process has a fixed definition of what good performance is (Barocas and Selbst 2016, 678-679). This stems from the fact that the AI decision-making process recognises only information that matches the information in its dataset (Buolamwini 2019). The AI decision-making process only rewards those types of performance which maximise the attainment of the programmers' definition of good performance (Barocas and Selbst 2016, 679). As a result, individuals will not be able to get a job unless they adhere closely to the parameters of a good employee as set by AI. According to Mireille Hildebrandt, individuals will exhibit desirable behaviour as defined by AI to increase their chances of getting a positive outcome (Hildebrandt 2017, 7). Given the fact that the employment of the AI decision-making process rewards individuals who seek to match AI logic in their behaviour (Ibid.) and who aim to attain the highest level of performance (Barocas and Selbst 2016, 678-679), the use of AI favours applicants who exhibit the following behaviours. They pay attention to following specific procedures when completing a task. They control their internal processes. They exhibit perfectionism. All of these traits are associated with obsessive-compulsive personality disorder (American Psychiatric Association 2022, 771-772).

Further support for the argument that the use of AI in hiring (Parker 2023) creates an environment that nudges individuals to exhibit particular behaviours can be found in the fact that regulating people through technology is a well-known and established approach (De Cooman 2021, 6). Roger Brownsword maintains that regulators know that they can channel peoples' conduct by using code integrated into the design of a technology (Brownsword 2005, 2). In the hiring context, the use of AI as part of the decision-making process (Parker 2023) has the impact of what Paul de Laat calls "governance by discipline" (de Laat 2019, 319) through operating on the mind (Ibid., 328). De Laat compares using AI to predict how closely



someone adheres to a specific parameter of behaviour and penalising the person for deviating from this parameter to a process of normation (Ibid., 322). He maintains that the employment of AI separates individuals into normal and abnormal categories in the AI's mathematical model of the external environment (Ibid., 324). Whether the AI allocates the individual into the normal or abnormal category depends on how closely that person's data maps in relation to the desired set of parameters in the AI's mathematical model of the external environment (Ibid., 324). De Laat concludes that institutions extend their power over the decision-making subject by using AI to measure how closely an individual adheres to specific parameters (Ibid., 323). Equally, they exercise power over individuals by using AI to predict their behaviour (Ibid., 323).

A possible counterargument is that employers have always exercised a degree of control over how individuals behave. They achieved this by formulating selection criteria and by requiring that the employee demonstrate a "cultural fit" to the company (Epstein 2021). While this is the case, using AI as part of the decision-making process drastically diminishes how much autonomy people can exercise and secure employment. Yeung calls using big data and AI to analyse the users' online behaviour and tailor the content as a "hypernudge" (Yeung 2017, 122). AI channels the users' perceptions and behaviour by shaping what content they see (Ibid., 130). As a result, AI shapes how users understand the world (Ibid., 130). One can arguably extend her argument to the context of the employment of AI as part of the hiring decision-making process (Parker 2023). It is suggested that using the AI as part of the decision-making process in hiring acts as a "hypernudge" (Yeung 2017, 122). This is the case because there is less room for interpretation when determining whether the information the applicant provided matches the decision-making criteria. Since AI detects correlations in the data and separates individuals into different groups (de Laat 2019, 323), individuals need to adhere to the pre-programmed parameters of a good employee to the greatest extent possible to get a positive outcome.

Solange Ghernaoui's scholarship points to the fact that the employment of AI as part of the decision-making process creates an environment that encourages individuals to control their internal processes and to exhibit rigidity in order to shape themselves to embody machine-type

logic. Ghernaouti maintains that deploying AI as part of the decision-making process requires individuals to follow the rules in a way that changes how they live, think and behave (Ghernaouti 2020, 15). It enacts a vision of rationality that imposes uniformity and standardisation on individuals (Ibid.). In the extreme, this logic of economic optimisation rationality can result in “a kind of eugenics of thought and behaviour” (Ibid.). Since many individuals will adjust their behaviour in light of their anticipation of how the AI decision-making processes operate to receive a favourable decision (Hildebrandt 2017, 7), Ghernaouti’s argument has weight. Individuals do not have many options. They can either adhere to the AI-based logic, pursue self-employment, apply to companies not using AI screening tools, or move abroad to countries where organisations do not employ AI decision-making processes. The greater the number of organisations that begin to deploy AI decision-making processes, the narrower the applicants’ choices become.

It should be noted at this stage that the author is not claiming a direct cause-and-effect relationship between the use of AI as part of the decision-making process and the individuals either acquiring or exhibiting the traits of a personality disorder. Instead, using the figure/ground approach to analysis reveals a particular pattern of change (McLuhan 1969). Namely, the analysis suggests that the employment of AI as part of the decision-making process creates an environment that is conducive to individuals having particular psychological experiences and responses. These psychological experiences and responses have parallels with experiencing traits of various personality disorders.

McLuhan talks about technology as an extension of the body, which “demands new ratios or new equilibriums among the other organs” (McLuhan 2013, 36). The medium creates a hidden environment that influences how the individual interprets the communication that the medium transmits (Matie, McLuhan, and Toye 1987, 443). Harvey elaborates that McLuhan envisions technology and a human being as becoming a single system that operates through a feedback mechanism (Harvey 2006, 337). It is claimed here that McLuhan’s approach (McLuhan 2013, 36) and Harvey’s interpretation of McLuhan (Harvey 2006, 337) help to understand that the employment of the AI decision-making process in hiring (Parker 2023) creates a feedback loop between the applicant and the AI. The operation of the AI decision-making process can indirectly affect the applicants’ perception and behaviour by creating an environment that nudges them to behave in a particular way. The applicant and AI are in a reciprocal

relationship, which is conducive to the applicants experiencing particular psychic effects. McLuhan points out that such effects can be concealed (McLuhan 2013, 35).

Consequently, the applicants could have difficulty articulating the precise way in which the operation of the AI decision-making process creates an environment that nudges them to behave in a particular way. This is particularly the case because the applicants will experience these effects both in the short and long term. As Hin-Yan Liu explains, using AI repeatedly produces cumulative and “structural” effects (Liu 2018, 200).

## **B. AI And The Traits of The Dependent Personality Disorder**

Using the AI decision-making process to screen candidates for employment (Parker 2023) creates an environment that nudges individuals to exhibit dependent personality disorder traits. Individuals with a dependent personality disorder exhibit submissive behaviour (American Psychiatric Association 2022, 768). They prefer others to make decisions for them (Ibid.). Additionally, they have difficulty expressing disagreement for fear of losing support (Ibid.). Reliance on the AI decision-making processes creates an environment that nudges individuals to submit to the developers' decisions. This deference to the developers' decisions relates to the areas of the applicants' lives that are linked to attitudes and behaviours associated with increasing the likelihood of receiving a high score from the AI decision-making process. The employment of the AI decision-making process can create a situation where individuals are habituated to machines defining parameters of good behaviour. Since individuals will adjust their behaviour in light of how the AI decision-making operates in order to receive a favourable decision (Hildebrandt 2017, 7), they can become habituated to following machine-based logic.

The deployment of AI as part of the decision-making process not only incentivises individuals to exhibit submissive behaviour. Additionally, it penalises applicants for failing to adhere to the machine-based definition of optimal performance by denying them access to employment. Using AI to monitor the worker's performance during employment as part of algorithmic management (Parent-Rochelleau and Parker 2022, 3) can deepen this effect. Employers want to use AI to set performance goals, monitor whether the employee is adhering to these goals,

rank employee performance, send nudges in relation to particular behaviours and terminate employees (Ibid., 3-4). The employment of AI as part of algorithmic management (Ibid., 3) can create an environment where the employees are expected to anticipate how AI will score them and to adjust their behaviour accordingly (Jarrahi, Newlands, Lee, Wolf, Kinder and Sutherland 2021,4). There is a risk of employees becoming “programmable cogs in machines” (Ibid; Frischmann and Selinger 2017).

Another significant factor is that companies could employ AI-generated scores regarding performance from the former employment place to select candidates for employment. Such usage can create an environment favouring employees who do not express dissenting views. Individuals may start to express disagreement on fewer occasions out of fear of being labelled negatively for the purpose of AI-driven decision-making at a future time. For instance, a manager could negatively evaluate an employee for questioning company practices. This negative evaluation can become a permanent record and a basis for future scoring due to being fed into the AI decision-making process. Additionally, using AI to screen candidates for employment can arguably enable employers to engage in disciplinary practices by analysing social media posts of applicants and by denying employment to candidates who made undesirable statements (Hearst 2023).

This situation is not as farfetched as it may seem. In the United Kingdom racialised and Palestinian academics reported complaints being filed against them with the university and the police after making comments on social media about the Israeli-Palestinian conflict in the fall of 2023 (Ibid.). Amazon illegally fired two employees because they spoke out about the firm's poor working conditions and the disproportionately negative impact which Amazon's environmental practices had on racialised communities (BBC 2021). These developments show that there is a real possibility that companies could use AI to distance themselves from candidates who make undesirable or controversial statements.

### **C. AI And The Traits of The Histrionic Personality Disorder**

Using AI to select candidates for employment (Parker 2023) can also create an environment that nudges the applicants to exhibit the traits associated with the histrionic personality disorder. The traits of this disorder include being easily influenced by the current trends and

the opinions of others (American Psychiatric Association 2022, 757). The employment of AI decision-making process in hiring (Parker 2023) can create an environment that nudges individuals to follow trends.

Shoshana Zuboff talks about the role of social media in causing individuals to feel insecure, pursuing signs of being valued and of increasing peoples' "natural orientation toward the group" (Zuboff 2018, 216). The AI decision-making process achieves the outcome of an environment which nudges individuals to follow trends in a different manner. This aspect stems from the fact that AI uses information about a group of individuals it treats as being similar to predict the applicant's future behaviour (Provost and Fawcett 2013, 21; Taylor 2018, 105; Taylor 2017, 14-15). The use of information about candidates whom the system treats as being similar to the applicant (Ibid.) creates an environment that incentivises individuals to mimic one another. Furthermore, this environment rewards applicants who mimic an influencer they perceive as possessing attributes that the employment of the AI-decision-making process favours. By acting in a similar way to the influencer, applicants maximise the chances that they will receive a positive decision from an AI decision-making process.

The fact that the phenomenon of an influencer is well established supports this argument. Mira Rawady and Robert Logan explain that companies sponsor influencers on social media platforms to promote their brand (Logan and Rawady 2021, ch. 13 p. 7). Individuals emulate influencers, such as celebrities (Ibid., ch. 8 pp. 3-5). They behave this way to get more likes on social media (Ibid.). Jelle Fastenau observes that after choosing an influencer on social media, some individuals copy the behaviour and preferences of that influencer (Fastenau 2018). Psychologists explain that basing one's behaviour on another person's is an aspect of human behaviour in general (Ibid.). While the internet creates an "imagined collective" (boyd 2010, 39), so does social media. The same is the case for the employment of the AI decision-making process. It is suggested that employing AI as part of the decision-making process creates an environment that nudges individuals to emulate an influencer who receives high scores on optimum performance parameters for a good employee (Parker 2023). Through this process, the AI decision-making process reinforces the group dynamic of following an influencer.

#### **D. AI And The Traits of The Borderline Personality Disorder**

The deployment of AI as part of the decision-making process creates an environment that is conducive to individuals exhibiting the traits of borderline personality disorder. People with borderline personality disorder have unstable self-image and sense of self (American Psychiatric Association 2022, 752). They have difficulty identifying a narrative describing their selves and how they came to be this way (Schmidt and Fuchs 2020, 325-326). One individual describes herself as not knowing who she is, her values and her preferences (Zandersen and Parnas 2019, 111). She feels empty (Ibid.). Relationships with others do not help her address this feeling of emptiness (Ibid., 111).

The earlier version of DSM-5 from 2013 included an additional descriptor. The descriptor was that individuals who experience borderline personality disorder have shifting values and goals (American Psychiatric Association 2013, 664). This description no longer appears in the 2022 version of the updated version of DSM-5. It is suggested that this discrepancy between different versions of DSM provides further support for the argument that borderline personality disorder is a social construct (Bjorklund 2006, 11). The present discussion encompasses the earlier criterion of having shifting values (American Psychiatric Association 2013, 664). This choice stems from the fact that the essence of personality disorders is that they cause “distress” for individuals who have them (Ibid., 733). Being in an environment that nudges an individual to have shifting values can be a source of distress for that person.

Helene Detach describes individuals with borderline personality disorder as moulding themselves to their surroundings and mimicking the environment instead of expressing their authentic selves (Detach 2007, 328). They pick up signals from the external world and adapt accordingly (Ibid.). Often, they form their opinions based on the opinion of the majority group to which they want to belong or a person with whom they identify (Ibid., 329). They can quickly change their views simply because their social circle changes rather than through a process of inward reflection (Ibid., 329). Ava Adore describes how she would incorporate aspects of other people's personalities into her own because their traits appealed to her (Adore 2023). Another reason why she did this was to remedy her sense of emptiness (Ibid.). She describes her personality as a composite of the traits of other people, which she incorporates into her own (Ibid.).

The employment of the AI decision-making process fosters an environment where individuals have less possibility to exercise their discretion when deciding how to formulate their narrative about why they are a good candidate for a job. As was already explained, the use of the AI decision-making process creates a technologically mediated identity of the candidate (Krupiy 2021b, 37). It achieves this by amalgamating and processing information relating to multiple applicants to generate a decision about an applicant (Ibid.). The use of AI as part of the decision-making process makes it hard for the applicants to create a narrative about their selves, which connects their conception of self, preferences, experiences and actions to the AI decision outcome coherently.

Consider the study of Elisa Harlan and Oliver Schnuck. Harlan and Schnuck found that the AI employment screening software ranked the same job candidate as less conscientious, agreeable and extroverted when she submitted a video recording with glasses on compared to when she submitted a video with no glasses on (Harlan and Schnuck 2021). The AI changed her personality rankings to be less neurotic, more agreeable, more extrovert and more conscientious when she recorded the same video while wearing a head covering and no glasses (Ibid.). The relationship between personality traits, such as wearing a head scarf and wearing glasses, is arbitrary. For this reason, the AI-generated output about the candidate's personality does not allow the applicant to establish a link between her sense of self and the AI-generated personality description. It is difficult for the applicants to map their values and preferences onto the AI-generated personality ranking. There is a lack of connection between the applicant's values and whether she is wearing contact lenses or glasses on a particular day. The fact that the AI-based decision-making process uses thousands of variables as inputs to determine the candidate's suitability for employment (Magid 2020) exacerbates the difficulty for job candidates to establish a coherent narrative about the relationship between their personality, preferences, characteristics, conduct and the application outcome. Another aggravating factor is that the algorithm's operation does not always disclose to the applicants how the machine calculated the score (Rahman 2021).

An individual can experience a sense of fragmentation of self-identity due to having to integrate the machine-generated conception of self into the personal concept of self. This is

because an applicant can struggle to reconcile and integrate the view of self with the machine-generated conception of self. Having more work or life experience will not necessarily help. Even freelancers with considerable work experience report having difficulty understanding how AI assigned scores to them based on their performance (Rahman 2021). Individuals can experience an inner dissonance due to having to integrate aspects of the personal conception of self with the machine-generated conception of self. As was already explained, the applicants have to carry out this process of integration because the deployment of AI as part of the decision-making process requires individuals to follow the rules and fosters a standardisation of thought (Gheraouti 2020, 15).

A sense of fragmentation can be accompanied by the applicant having difficulty forming a coherent narrative about oneself. This state of affairs can occur due to the applicant perceiving the machine-generated description of self as being different from the inner conception of self. The process of having difficulty understanding the machine's description of self (Rahman 2021) while having to integrate aspects of the personalities of other people into one's own in order to receive a favourable decision from the AI decision-making process on the next attempt can make it hard for the applicants to form a coherent meta-narrative of self. It is maintained that having challenges with formulating a coherent narrative of one's own identity parallels with how an individual experiences a trait of borderline personality disorder (Schmidt and Fuchs 2020, 325-326). Moreover, an individual with the traits of borderline personality disorder can absorb the personalities of other people into oneself as a way of filling a sense of emptiness (Adore 2023). Meanwhile, the job applicant can feel compelled to integrate the machine-generated conception of self to change behaviour in a way that increases the chances of being ranked highly by the AI decision-making process.

McLuhan's writings capture the fact that the employment of the AI-based decision-making processes in hiring (Parker 2023) changes the environment in which individuals construct a conception of self (Krupiy 2021b, 37). McLuhan observes that technology acts on people through their senses and brings about a new equilibrium (McLuhan 2013, 78). Media change how individuals relate to themselves (Ibid., 13). All media provide "artificial perception and arbitrary values" (Ibid., 130). Media trigger irritation in the body and generate physical stress (Ibid., 34). McLuhan's writings capture the sense of fragmentation, dissonance and discomfort that individuals can experience when they integrate the description or score that the AI generates with their conception of self.



Furthermore, the deployment of AI in the decision-making process concerning employment (Parker 2023) arguably creates an environment that nudges individuals to exhibit the trait of borderline personality disorder of having shifting values (American Psychiatric Association 2013, 664). According to Ghernaouti, the developers of the AI decision-making processes can shape how individuals think and act by subjecting them to following the rules (Ghernaouti 2020, 15). Different corporations developing AI can embed varying visions of attributes that correspond to an ideal employment candidate. As a result, applicants will need to behave in different ways in order to be scored highly by different AI systems. This situation can lead to applicants adopting different values and behaviours to increase their chances of getting positive scores from different AI systems. The cumulative use of such systems creates an environment conducive to individuals having different values and exhibiting different behaviours over a short time.

A possible counterargument is that people can shift their values over time for various reasons. For example, people can change their opinions due to exposure to new information. A response to this observation is that typically, individuals change opinions through a process of deliberating (Schmidt and Fuchs 2020, 335). They refer to their sense of self in order to reconcile competing inclinations (Ibid.). On the other hand, it is proffered that the use of the AI decision-making processes in hiring (Parker 2023) confronts individuals with a technologically mediated construction of self (Krupiy 2021b, 37). This machine-generated construction of self does not acknowledge that the subjective, discretionary decisions of the programmers shape how AI depicts the attributes of the applicant in its model of the external environment (Ibid.; Arvidsson and Noll 2023, 60). This opaqueness in the design and operation of AI (Ibid., 90-91; von Eschenbach 2021, 1608) makes it harder for the applicants to identify how the values underpinning the AI decision-making process relate to their own. Consequently, individuals have less possibility to arrive at an informed decision about how their values differ from those programmed in the AI. Not knowing how AI operates (von Eschenbach 2021, 1608) and AI's capacity to mimic behaviour that individuals perceive as being intelligent (UNESCO 2021, 5) can result in the applicants taking the AI-generated outputs at face value.

It is put forward that the applicants can end up chasing positive outcomes without fully realising the role of AI decision-making processes in enacting “governance by discipline” over them (de Laat 2019, 319). Neither will they be aware of how the discretionary judgments of the developers (Arvidsson and Noll 2023, 60) influenced their score or profile. The applicants can experience cognitive dissonance and distress without being able to attribute these experiences to the use of AI as a cause. McLuhan captured this aspect when he noted that the continuous use of technology leads to individuals having “subliminal awareness and numbness” in relation to the images of themselves in technology (McLuhan 2013, 36).

McLuhan’s approach to analysis helps one to understand that different applicants can sense dissonance between the technologically mediated snapshot of self (Krupiy 2021b, 37) and their own conception of self to varying degrees without being fully aware of what specific aspects of the use of AI decision-making process contribute to this experience. Even individuals with intimate knowledge of how AI operates will likely have difficulty establishing how the AI decision-making process generates a specific snapshot of self. This is the case because companies can protect AI as a trade secret under intellectual property law (Foss-Solbrekk 2021, 247). Additionally, individuals will find it hard to test how different factors contribute to decision outcomes. This discussion demonstrates that the employment of AI as part of the employment decision-making process (Parker 2023) fosters an environment where individuals are nudged to experience traits associated with various personality disorders. The findings have weight since this article draws on literature from different disciplines to supplement the media ecology analysis. Moreover, the arguments are not rooted in overgeneralisations due to entailing a detailed analysis of how AI operates.

## **VI. Lessons for the Human Rights Community**

The media ecology analysis can help human rights defenders uncover the effects that the employment of technology, such as AI, produces on individuals. For instance, using the theoretical framework of media ecology can enable human rights defenders to trace an indirect relationship between the employment of AI as part of the hiring decision-making process (Ibid.) and the applicants experiencing a lack of mental wellbeing. The use of media ecology allows human rights defenders to sound an alarm and call for further investigation of the problem before individuals start to experience poor mental wellbeing. By using media ecology to detect

potential harms at an early stage, human rights defenders can prevent a situation where the population serves as a guinea pig. It is too late to wait until individuals begin to experience poor wellbeing in order to empirically test whether the employment of AI as part of the decision-making process results in individuals exhibiting the traits of various personality disorders. Moreover, since any individual can exhibit some traits of a personality disorder without having the disorder in question (Heller 2020; American Psychiatric Association 2022, 737), it may be challenging to employ empirical research methods to establish a causal link between the employment of AI and specific individual outcomes. The value of media ecology is that it helps establish harms that occur indirectly and may not be readily evident.

It is suggested that the findings in this article make it possible for human rights defenders to invoke Articles 2(1) and 12 of ICESCR in relation to the employment of AI to screen candidates for employment (Parker 2023). The two provisions require states to take steps “to the maximum” of available resources in order to achieve “progressively the full realisation of the right” (ICESCR 1966, Art 2(1)) to the enjoyment of the highest attainable standard of mental health (Ibid., Art 12). Human rights defenders can campaign for a ban on using AI as part of the decision-making process to select candidates for employment. For example, it is put forward that human rights defenders can inform states that are party to ICESCR and are part of the European Union that it is insufficient for the Artificial Intelligence Act to merely treat the use of AI in the employment decision-making process as posing a high risk (Artificial Intelligence Act 2024, Annex III Para 4). They can call on these states to revise Annex III to the Artificial Intelligence Act (Ibid.) so that the states ban AI systems intended to be used for recruitment. Since the ban on certain AI applications does not require states to spend additional resources on supplying mental health services to the population, most states cannot argue that they lack sufficient resources to adopt these measures. The budgetary expenditure needed to ban specific AI applications is arguably likely to be much smaller than the resources that states will need to spend on providing mental health services to individuals who are negatively affected by AI's use to screen applicants for employment (Parker 2023). Additionally, human rights defenders can use media ecology to analyse how the use of AI in other contexts is likely to affect individuals in order to formulate their campaigning agendas.

## VII. Conclusion

Applying the media ecology approach to analysis, we learned that using AI to screen applicants for employment (Ibid.) affects how the applicants perceive and act. The use of AI decision-making tools to select applicants for employment (Ibid.) creates an environment that nudges individuals to exhibit the traits of various personality disorders. As a result, individuals experience lower mental wellbeing. Such impacts are not immediately obvious to the eye. The discussion illustrates that human rights advocates can use media ecology to identify how the employment of AI in different contexts impacts individuals through changing society. It is maintained that human rights advocates can use this theoretical framework to identify harms arising from the deployment of AI which may either be not readily visible (Birhane 2021, 129) or which do not occur through a direct pathway of causation (Krupiy 2021, 14-15). As a result of having a wider array of tools to map what types of harms the employment of AI brings about in different contexts, human rights advocates are better positioned to define their campaign agenda.

In turn, having a better understanding of the harms in question allows human rights defenders to determine how they can interpret international human rights law norms so as to apply to govern the emerging social reality. Being equipped with more knowledge better positions human rights defenders to advocate for states to adopt specific laws to govern AI. More broadly, this article shows why interdisciplinary analysis is critical for appropriately identifying social harms arising from using AI. Since legal practitioners cannot apply laws to harms about which they do not know, media ecology enables legal practitioners to better protect the fundamental rights of individuals by enabling them to identify harms that may otherwise remain undetected.

## Bibliography

Adore A. 2023. *What Lack of Identity Feels Like With Borderline Personality Disorder*. <https://themighty.com/topic/borderline-personality-disorder/lack-of-identity-feels-like-borderline-personality-disorder-bpd/>. (Accessed on 11 January 2024).

Adshead G. and Sarkar J. 2012. The Nature of Personality Disorder. *Advances in Psychiatric Treatment*, 18, p. 162–172.

Agarwal, A. (2019). *Introduction to Statistical Methods in AI*. <https://towardsdatascience.com/introduction-to-statistical-methods-in-ai-23f89df72cdb>.

(Accessed on 3 January 2024).

American Psychiatric Association. 2013. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. Arlington: American Psychiatric Association.

American Psychiatric Association. 2002. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5-TR Fifth Edition Text Revision*. Washington DC: American Psychiatric Association.

Australian Human Rights Commission. 2021. *Human Rights and Technology Final Report*. Sydney: Australian Human Rights Commission.

Badar A., Umre B.S., and Junghare, A.S. 2014. Study of Artificial Intelligence Optimisation Techniques Applied to Active Power Loss Minimisation. *IOSR Journal of Electrical and Electronics Engineering*, p. 39-45.

Bahmanteymouri E., Bartlett M., Blackmore B., Burmester B., Chen A. TY., Morreale F., and Thorp M. 2023. Hidden Humans: Exploring Perceptions of User-work and Training Artificial Intelligence in Aotearoa New Zealand, Kōtuitui, *New Zealand Journal of Social Sciences*, 18(4), p. 443-456.

Barocas S. and Selbst A.D. 2016. Big Data's Disparate Impact. *California Law Review*, 104, p. 671-732.

Basbaum S.R. 2022. From Objective Reality to Fake News Galaxy: Narcissus Narcosis Images and Imagination in Contemporary Global Expanded Reality Village, *New Explorations: Studies in Culture and Communications*, 2(3), p. 78-85.

Bates T. 2011. *Marshall McLuhan and His Relevance to Teaching with Technology*. <https://www.tonybates.ca/2011/07/20/marshall-mcluhan-and-his-relevance-to-teaching-with-technology>. (Accessed on 23 February 2024).

BBC News. 2021. *Amazon 'Illegally Retaliated' Against Climate Activists*. <https://www.bbc.co.uk/news/business-56641847>. (Accessed on 3 January 2024).

Bircek N., Osborne L., Reed P., Viganò C., and Truzoli, R. 2018. Visual Social Media Use Moderates the Relationship Between Initial Problematic Internet Use and Later Narcissism, *The Open Psychology Journal*, 11(1), p. 163-170.

Birhane A. 2021. *Automating Ambiguity: Challenges and Pitfalls of Artificial Intelligence* [Unpublished doctoral dissertation]. Dublin: University College Dublin.

Birhane A., and Sumpter D. 2022. The Games We Play: Critical Complexity Improves Machine Learning, *arXiv:2205.08922v1*, p. 1-14.

Bjorklund P. 2006. No Man's Land: Gender Bias and Social Constructivism in the Diagnosis of Borderline Personality Disorder. *Issues in Mental Health Nursing*, 27(1), p. 3-23.

Blanco C. 2023. Unveiling the Myth of Narcissus. Lit & Phil. Newcastle Upon Tyne.

Boyd d. 2010. Social Network Sites as Networked Publics: Affordances, Dynamics and Implications. In Z. Papacharissi (ed) *Networked Self: Identity, Community And Culture on Social Network Sites*. New York: Routledge.

Brown S. 2021. *Machine Learning, Explained*. <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained>. (Accessed 3 January 2024).

Brownsword R. 2005. Code, Control, and Choice: Why East is East and West is West, *Legal Studies*, 25(1), p. 1-21.

Buolamwini J. 2019. *Artificial Intelligence Has a Problem With Gender and Racial Bias. Here's How to Solve It*. <https://time.com/5520558/artificial-intelligence-racial-gender-bias/>. (Accessed on 11 January 2024).

Central Digital and Data Office. 2020. Guidance Data Ethics Framework: Glossary and Methodology. <https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework-glossary-and-methodology>. (Accessed on 5 January 2024).

Chander S. and Jakubowska E. 2021. *Civil Society Calls on the EU to Put Fundamental Rights First in the AI Act*. <https://edri.org/our-work/civil-society-calls-on-the-eu-to-put-fundamental-rights-first-in-the-ai-act/>. (Accessed on 25 August 2022).

Chapman A.L. 2019. Borderline Personality Disorder and Emotion Dysregulation, *Development and Psychopathology*, 31, p. 1143–1156.

Coid J., Tyrer P., and Yang M. 2010. Personality Pathology Recorded by Severity: National Survey, *British Journal of Psychiatry*, 197, p. 193–199.

Coupland D. (2009). *Marshall McLuhan*. Toronto: Penguin.

Crawford K. 2021. *Atlas of AI: Power, Politics and the Planetary Costs of Artificial Intelligence*. New Haven: Yale University Press.

De Cooman J. 2021. From the Regulation of Artificial Intelligence by Society to the Regulation of Society by Artificial Intelligence: All Along the Watchtower. In H. Jacquemin (ed) *Time to Reshape the Digital Society*. Brussels: Larcier.

Department for Science, Innovation and Technology, and Office for Artificial Intelligence. 2023. *AI Regulation: a Pro-innovation Approach*. <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach>. (Accessed on 3 January 2024).

Detach H. 2007. Some Forms of Emotional Disturbance and Their Relationship to Schizophrenia, *The Psychoanalytic Quarterly*, 76(2), p. 325-344.

Donahoe E. and MacDuffee Metzger M. 2019. Artificial Intelligence and Human Rights, *Journal of Democracy*, 30(2), p. 115-126.

Epstein S. 2021. *What Does Being a 'Cultural Fit' Actually Mean?*. <https://www.bbc.co.uk/worklife/article/20210916-why-inexperienced-workers-cant-get-entry-level-jobs>. (Accessed on 25 August 2022).

Eschenbach W von. 2021. Transparency and the Black Box Problem: Why We Do Not Trust AI. *Philosophy & Technology*, 34, p.1607-1622.

European Commission. 2021. *Council and European Parliament Draft Regulation COM (2021) 206 of 21 April 2021 Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>. (Accessed on 25 August 2022).

[The European Parliament and the Council of the European Union. 2024. Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence \(Artificial Intelligence Act\) and Amending Certain Union Legislative Acts 2021/0106 \(COD\)](#). Brussels: The European Parliament and the Council of the European Union.

Fastenau, J. 2018. *Under the Influence: The Power of Social Media Influencers*. Medium. <https://medium.com/crobox/under-the-influence-the-power-of-social-media-influencers-5192571083c3>. (Accessed on 25 August 2022).

Ford B. and Schildhorn B. 2022. *How AI is Deciding Who Gets Hired*. <https://www.bloomberg.com/news/articles/2022-02-03/how-using-artificial-intelligence-in-hiring-might-be-perfecting-bias>. (Accessed on 21 August 2023).

Foss-Solbrekk K. 2021. Three Routes to Protecting AI systems and Their Algorithms Under IP Law: The Good, the Bad and the Ugly, *Journal of Intellectual Property Law & Practice*, 16(3), p. 247-258.

Frischmann B. and Selinger E. 2017., Robots Have Already Taken Over Our Work, But They're Made of Flesh and Bone. *The Guardian*, 25 September 2017. <https://www.theguardian.com/commentisfree/2017/sep/25/robots-taken-over-work-jobs-economy> (Accessed on 22 February 2024).

Fuchs T. and Schmidt P. 2021. The Unbearable Dispersal of Being: Narrativity and Personal Identity in Borderline Personality Disorder, *Phenomenology and the Cognitive Sciences*, 20, p. 321–340.

Gajjar D. 2023. *POSTbrief 57: Artificial Intelligence—An Explainer*. London: Parliamentary Office of Science and Technology.

Gayle D. 2021. Facebook Aware of Instagram's Harmful Effect on Teenage Girls, Leak Reveals. *The Guardian*. <https://www.theguardian.com/technology/2021/sep/14/facebook-aware-instagram-harmful-effect-teenage-girls-leak-reveals>. (Accessed on 21 August 2023).



- Gheraouti S. 2020. Artificial Intelligence and Power Asymmetries: Challenges for Civil Society and Policy Making Processes. In *Promises and Pitfalls of Artificial Intelligence for Democratic Participation Workshop Proceedings*. University of Geneva.
- Harlan E. and Schnuck O. 2021. *Objective or Biased: On the Questionable Use of Artificial Intelligence for Job Applications*. <https://web.br.de/interaktiv/ki-bewerbung/en/>. (Accessed on 21 August 2023).
- Hearst K. 2023. Israel-Palestine War: Social Media Surveillance Creates a 'Culture of Fear' on UK Campuses. *Middle East Eye*, 2023.
- hÉigeartaigh S.O., Flach P., Hernández-Orallo J., Loe B.S., Martínez-Plumed F., and Vold K. 2018. The Facets of Artificial Intelligence: A Framework to Track the Evolution of AI. In *Twenty-Seventh International Joint Conference on Artificial Intelligence*. Stockholm: IJCAI.
- Heller S. 2020. *The Many Faces of Narcissism: Understanding the Spectrum From Healthy to Malignant*. <https://medium.com/invisible-illness/the-many-faces-of-narcissism-a7a1f65a5151>. (Accessed on 21 August 2023).
- Henley J. 2021. Dutch Government Resigns Over Child Benefits Scandal. *The Guardian*. <https://www.theguardian.com/world/2021/jan/15/dutch-government-resigns-over-child-benefits-scandal>. (Accessed on 21 August 2023).
- Hildebrandt M. 2017. Learning as a Machine: Crossovers Between Humans and Machines, *Journal of Learning Analytics*, 4(1), p. 6-23.
- Horwitz A. V. 2021. *DSM: A History of Psychiatry's Bible*. Baltimore: John Hopkins University Press.
- Howard J. 1966. Marshall McLuhan Canada's Talky Social Catalyst: Oracle of the Electric Age. *Life Magazine*, 6(8), p. 91-99.
- Human Rights Council. 2021. The Right to Privacy in the Digital Age: Report of the United Nations High Commissioner for Human Rights UN Doc A/HRC/48/31. Geneva: Human Rights Council.
- IBM. 2024. *What is Artificial Intelligence?*. <https://www.ibm.com/topics/artificial-intelligence>. (Accessed on 21 February 2024).
- Jarrahi MH., Newlands G., Lee MK., Wolf CT., Kinder E. and Sutherland W. 2021. Algorithmic Management in a Work Context, *Big Data & Society*, 8(2), p. 1-14.
- Jarrett C. 2017. *Millennials are Narcissistic? The Evidence is Not So Simple*. <https://www.bbc.com/future/article/20171115-millennials-are-the-most-narcissistic-generation-not-so-fast>. (Accessed on 21 August 2023).
- Jayasuriya K. 1999. Globalization, Law and the Transformation of Sovereignty: The Emergence of Global Regulatory Governance, *Indiana Journal of Global Legal Studies*, 6, p. 425-455.



- Jones E. 2023. *Explainer: What is a Foundation Model?*. <https://www.adalovelaceinstitute.org/resource/foundation-models-explainer/#:~:text=Foundation%20models%20are%20AI%20models%20designed%20to%20produce,as%20a%20%E2%80%98base%E2%80%99%20for%20many%20other%20applications.%205.> (Accessed 3 January 2024).
- Kirmayer L. and Young A. 1999. Culture and Context in the Evolutionary Concept of Mental Disorder, *Journal of Abnormal Psychology*, 108 (3), p. 446-452.
- Krupiy T. 2021a. Meeting the Chimera: How the CEDAW Can Address Digital Discrimination. *International Human Rights Law Review*, 10, p. 1-39.
- Krupiy T. 2021b. Understanding Digital Discrimination: Analysing Marshall McLuhan's Work Through a Human Rights Lens, *New Explorations: Studies in Culture and Communication*, 2(1), p. 1-22.
- Liu HY. 2018. The Power Structure of Artificial Intelligence, *Law, Innovation and Technology*, 10, p. 197-229.
- Logan R. 2007. The Biological Foundation of Media Ecology, *OCAD University Open Research Repository*, p. 1-26.
- Logan R. 2011. McLuhan Misunderstood: Setting the Record Straight. *Razon y Palabra*, 76, p1-32.
- Logan R. 2020. Understanding Human Users: Extensions of Their Digital Media. *New Explorations: Studies in Culture and Communication*, 2020, 1, p. 1-14
- Logan R. and Rawady M. 2021. *Understanding Social Media: Extensions of Their Users (Understanding Media Ecology)*. New York: Peter Lang.
- Magid J.M. 2020. *Does your AI Discriminate?*. <https://theconversation.com/does-your-ai-discriminate-132847>. (Accessed on 3 January 2024).
- Matie M., McLuhan C., and Toye W. 1987. *Letters of Marshall McLuhan*. Toronto: Oxford University Press.
- McLuhan M. 1969. Playboy Magazine Interview. *Playboy Magazine*.
- McLuhan M. and Fiore Q. 1967. *The Medium is the Massage: An Inventory of Effects*. New York: Random House.
- McLuhan, M. 2013. *Understanding Media: Extensions of Man*. New York: Gingko Press.
- Menninger K. 1963. *The Vital Balance*. New York: Viking Press.
- Milmo M. 2021. Frances Haugen Takes on Facebook: the Making of a Modern US Hero. *The*

Guardian. <https://www.theguardian.com/technology/2021/oct/10/frances-haugen-takes-on-facebook-the-making-of-a-modern-us-hero>. (Accessed on 21 August 2023).

Moore M. 2019. *What is AI? Everything You Need to Know About Artificial Intelligence*. <https://www.techradar.com/news/what-is-ai-everything-you-need-to-know>. (Accessed on 21 August 2023).

Maushart S., Nystrom C. And Wiebe, C. 2021. *The Genes of Culture: Towards a Theory of Symbols, Meaning, and Media* (Vol. 1). New York: Peter Lang.

Office of the Secretary-General's Envoy on Technology. 2023. *High-Level Advisory Body on Artificial Intelligence*. United Nations. <https://www.un.org/techenvoy/ai-advisory-body#:~:text=The%20multi-stakeholder%20High-level%20Advisory%20Body%20on%20Artificial%20Intelligence%2C,for%20the%20international%20governance%20of%20artificial%20intelligence%20%28AI%29>. (Accessed on 21 December 2023).

Organisation for Economic Co-operation and Development. 2019. *Artificial Intelligence in Society (Summary)*. Paris: Organisation for Economic Co-operation and Development.

Parent-Rocheleau X. and Parker S.K. 2022. Algorithms as Work Designers: How Algorithmic Management Influences the Design of Jobs, *Human Resource Management Review*, 32(3), p. 1-17.

Parker E. 2023. *Candidate Screening with AI: A Game-Changer in Recruitment Efficiency*. <https://hirebee.ai/blog/automated-candidate-screening-with-hirebee/candidate-screening-with-ai-a-game-changer-in-recruitment-efficiency>. (Accessed on 11 January 2024).

Parnas J. and Zandersen M. 2019. Identity Disturbance, Feelings of Emptiness, and the Boundaries of the Schizophrenia Spectrum, *Schizophrenia Bulletin*, 45(1), p. 106–113.

Perrotta G. 2020. Borderline Personality Disorder: Definition, Differential Diagnosis, Clinical Contexts and Therapeutic Approaches, *Annals of Psychiatry and Treatment*, 4(1), p. 43-56.

Provost F. and Fawcett T. 2013. *Data Science for Business*. Sebastopol: O'Reilly Media Inc.

Rahman H. 2021. *Gig Workers Are Increasingly Rated by Opaque Algorithms. It's Making Them Paranoid*. <https://insight.kellogg.northwestern.edu/article/gig-workers-algorithm> (Accessed on 22 February 2024).

Sapia.ai. 2024. *Save Time and Hire Fast with AI Chat Interviewing*. <https://sapia.ai/products/interview/>. (Accessed on 11 January 2024).

Sapignoli, M. 2021. The Mismeasure of the Human: Big Data and the “AI Turn” in Global Governance, *Anthropology Today*, 37, p. 4-8.

Szacz T. 1974. The Myth of Mental Illness. In R.J. Morris (eds) *Perspectives in Abnormal Behavior: Pergamon General Psychology Series*. Elmsford: Pergamon Press.

Taylor L. 2017. Safety in Numbers? Group Privacy and Big Data Analytics in the Developing

World. In L. Floridi, L. Taylor, and B. Van der Sloot (eds.) *Group Privacy: New Challenges of Data Technologies*. Cham: Springer International Publishing.

Taylor L. 2018. On the Presumption of Innocence in Data-Driven Government. Are We Asking the Right Question?. In I. Baraliuc et al. (eds) *Being Profiled: Cogitas Ergo Sum*. Amsterdam: Amsterdam University Press.

Teichman D. and Zamir E. 2020. Nudge Goes International. *European Journal of International Law*, 30(4), p. 1263-1279.

The Alan Turing Institute. 2024. *Data Science and AI Glossary*.  
<https://www.turing.ac.uk/news/data-science-and-ai-glossary>. (Accessed 3 January 2024).

[UNESCO. 2021. Draft Test of the Recommendation on the Ethics of Artificial Intelligence SHS/IGM-AIETHICS/2021/JUN/3 Rev.2.](https://unesdoc.unesco.org/ark:/48223/pf0000377897)  
<https://unesdoc.unesco.org/ark:/48223/pf0000377897>. (Accessed on 8 November 2023).

United Nations Secretary-General António Guterres. 2018. *Secretary-General's Strategy on New Technologies*.  
<https://www.un.org/en/newtechnologies/images/pdf/SGs-Strategy-on-New-Technologies.pdf>.  
(Accessed on 11 January 2024).

United Nations. 2024. *International Human Rights Law*. <https://www.ohchr.org/en/instruments-and-mechanisms/international-human-rights-law>. (Accessed on 3 January 2024).

Viljoen S. 2021. A Relational Theory of Data Governance, *Yale Law Journal*, 131, p. 573-654.

Yeung, K. 2018. Five Fears About Mass Predictive Personalisation in an Age of Surveillance Capitalism, *International Data Privacy Law*, 8, p. 258-269.

Zimmerman M. 2023. *Overview of Personality Disorders*.  
<https://www.msmanuals.com/professional/psychiatric-disorders/personality-disorders/overview-of-personality-disorders>. (Accessed on 8 November 2023).

Zuboff S. 2018. *The Age of Surveillance Capitalism: the Fight for the Future at the New Frontier of Power*. London: Profile Books.

### International treaties

International Covenant on Economic, Social and Cultural Rights (adopted 16 December 1966, entered into force 3 January 1976) 993 UNTS 3 (ICESCR).