Universität St.Gallen

THE HUMAN ERROR PROJECT

Research Report

AI Errors and the Profiling of Humans:
Mapping the Debate in European News Media

University of St.Gallen, St.Gallen, 08th December 2022

University of St. Gallen (HSG)
School of Humanities and Social Sciences (SHSS HSG)
Institute for Media and Communications Management (MCM HSG)

Conducted by
Chair of Media and Culture,
School of Humanities and Social Sciences (SHSS)
Institute for Media and Communications Management (MCM)
University of St. Gallen (HSG)

Prof. Dr. Veronica Barassi
Dr. Antje Scharenberg
Ms. Marie Poux-Berthe
Ms. Rahi Patra

In collaboration
with Dr. Philip Di Salvo

St.Gallen, 08 December 2022

University of St. Gallen (HSG)
School of Humanities and Social Sciences (SHSS HSG)
Institute for Media and Communications Management (MCM HSG)

Table of Contents

Preface

The Human Error Project

We are living in a historical time when every little detail of our experience is turned into a data point that is used by AI systems to profile and make automated decisions about our lives. These technologies are increasingly used worldwide. Health and education practitioners use them to 'track risk factors' or to find personalized solutions. Employers, banks, and insurers use them to judge clients or potential candidates. Even governments, the police and immigration officials use these technologies to make decisions about individual lives, from one's right to asylum to one's likelihood to commit a crime. The COVID-19 pandemic has only intensified and exacerbated these practices of technological surveillance, algorithmic profiling and automated decision making.

In different sections of society algorithmic profiling is often understood as holding the key to human nature and behavior; it is used to make the process of decision making more efficient, and to 'avoid the human error'. Paradoxically, however - as recent research has shown - these technologies are filled with systemic 'errors', 'biases' and 'inaccuracies' when it comes to human profiling.

Of course, AI systems can bring many positive outcomes. This is clear if we consider their use in tackling specific issues such as diseases or climate change. Yet, when it comes to **human profiling** these technologies cannot grasp the complexity of human experience and behaviors, and their errors can have a real impact on individual lives and human rights.

In 2020, we launched *The Human Error Project: AI, Human Rights, and the Conflict Over Algorithmic Profiling*, because we believed that - in a world where algorithmic profiling of humans is so widespread - critical attention needs to be paid on how institutions, businesses, and individuals **coexist, negotiate and construct meaning out of AI errors**.

In our research we use the term 'the human error of AI' as an umbrella concept to shed light on different aspects of algorithmic fallacy when it comes to human profiling:

> ***Bias*** – AI systems are human made and will always be shaped by the cultural values and beliefs of the humans and societies that created them.
> ***Inaccuracy*** – AI systems process data. Yet the data processed by algorithms is often the product of everyday human practices, which are messy, contradictory and taken out of context, hence algorithmic predictions are filled with inaccuracies, partial truths, and misrepresentations.
> ***Un-accountability*** – AI systems lead to specific predictions that are often un-explainable, and unaccountable. How can we trust or challenge their decisions if we cannot explain or verify them?

The combination of bias, inaccuracy, and lack of transparency in algorithmic predictions, we believe, implies that AI systems are often (if not always) somehow fallacious in reading humans.

The Human Error Project thus shares much of the same understandings of current research in the field of critical AI and data studies which has shown that AI systems are often shaped by systemic inequalities (Eubanks, 2018; Amoore, 2020; Crawford, 2021), by racial biases (Noble, 2018; Benjamin, 2019; Richardson et al. 2019; Atanasoski and Vora, 2019) and by inaccurate and human reductionist analyses of human practices and intentions (Barassi, 2020; Milan, 2020).

Yet we also want to push the debate further, and ask ***What next?*** We want to question what happens when different actors in society realize that AI systems can be fallacious and biased in reading humans; when they discover that AI systems can be racist, sexist, ageist, ableist and so on? How are different sections of society understanding and shaping the political debate on the Human Error of AI? How are they negotiating and coexisting with the human rights implications of AI? What solutions and AI futures are different actors envisaging?

We launched *The Human Error Project* because we believe that one of the most fundamental questions of our times has become that of mapping, studying, and analyzing the emerging debates and conflicts over AI errors and algorithmic profiling. With this project we position ourselves amongst those scholars that have called for an analysis of the 'political life of technological errors' (Aradau and Blanke, 2021) and for a qualitative approach to the understanding of algorithmic failures (Munk et al., 2022; Rettberg, 2022).

Our aim is to **map the discourses** and **listen to the human stories** of different sections of society, to try and understand **how AI errors** - when it comes to the profiling of humans - are **experienced, understood and negotiated.** To achieve our goals, *The Human Error Project* team is researching three different areas of society where these conflicts over algorithmic profiling are being played out in Europe: the media and journalists; civil society organizations and critical tech entrepreneurs. For all these different sections of society we are gathering data primarily through three main methodologies: critical discourse analysis, organizational mapping, and the collection of 100 in-depth interviews.

Our methodological approach, as we will clarify here below, is based on the understanding that whilst most of current research and influential journalism in the field of critical AI studies comes from the U.S. and focuses on algorithmic injustice with reference mostly to U.S.-centric systems of inequality, European countries (within and outside the E.U.) and their cultural specificities can be crucial in shedding light on how debate on AI errors, algorithmic profiling and human rights is being shaped.

What follows is the first report resulting from *The Human Error Project*. This is the product of **a 2-year-long critical discourse analysis of European media coverage** of AI errors in algorithmic profiling. This for us marks the first step of a broader investigation into the human error of AI.

Introduction

News media are a key field of interest when we want to 'map the debate' on the arrival and adoption of new technologies and their implications. As sociologist Pierre Bourdieu (1991) has shown, news media can be understood as the center of symbolic power in our societies, and they can have a fundamental role when it comes **to construct our sense of reality**. Research has shown that this is particularly true with reference to **emerging technologies**. A study conducted in Germany at the turn of the century, for example, highlighted the role news media played in shaping the cultural belief that the Internet was a positive and emancipatory technology (Roessler, 2001). Another interesting study, among others, is the one of Weaver et al. (2009) on how - between 1999 and 2008 – news media became an important space of debate in identifying the risks and opportunities of nanotechnologies.

If we consider this earlier research on the relationship between news-media and emerging technologies, it is not surprising that nowadays the same is happening with reference to AI[1]. As Sun and Zhai (2022) have shown, over the last years, news media have played a fundamental role in shaping the public debate around AI technologies; media presented AI as a sophisticated, powerful, and value-laden issue, and journalists often acted as **translators of technical knowledge** (Sun and Zhai, 2022). The news media have also been crucial in shaping the moral debates around AI Ethics (Ouchchy et al. 2020). This is clear, for instance, if we consider debates on the implementation of specific technologies such as facial recognition technologies. Shahik and Moran' (2022) fascinating work, shows that - when it came to reporting on the risks and opportunities of facial recognition technologies in the US - the "media frame"[2] was shaped by clear ideological stances between more left-wing and right-wing media (Shahik and Moran, 2022). A 2018 study by the Reuters Institute for the Study of Journalism (RISJ) has instead shown how the debate around these issues is frequently dominated by industry and corporate voices (Brennen, Howard and Nielsen, 2018).

In *The Human Error Project*, we shared these studies' understanding of the centrality of news-media in constructing specific discourses of AI. Yet, our approach was different for two main reasons. Firstly, we did not want to focus – as others have done – on the notion of AI as a general, over-embracing term addressing all emerging technologies. Instead, we decided to focus on one specific societal issue: **AI errors in human profiling**.

Secondly, rather than using a quantitative/content analysis approach to explore or identify dominant media frames, we focused on a **qualitative, critical discourse analysis** methodology (Van Dijk, 1996), which was aimed at highlighting the multiple, contradictory, context-specific and complex narratives that appeared within media discourse. Our approach was influenced by the belief that **media power is a subject of great complexity and controversy** (Curran and Seaton, 1997, p. 270), and that the media rather than imposing top-down forms of meaning

[1] Ideas of AI are of course also being shaped by science fiction and popular culture (see Broussard, 2018) but in our work we decided to focus on news media only.

[2] The notion of media frame comes from Goffman's (1974) frame analysis approach and was applied to communication studies to highlight how often media discourse is shaped by a central organizing idea or story that provides meaning' (Gamson and Modigliani, 1989, p. 143)

construction, often become the space where meanings in society are transmitted, negotiated, and contested (Hall, 1997).

Hence, in our analysis, we did not want to identify and quantify the dominant media frames, we were rather interested in shedding light on the juxtaposed, contradictory narratives that are shaping the debate around AI errors, algorithmic profiling, and human rights.

This report is the product of this effort. It is based on a **longitudinal, cross-cultural discourse analysis** of articles collected in the period between **February 2020 and February 2022** in European news media. For our research we focused on three **core countries**: Germany, France, and the UK. We have selected these three countries, not only because they are the biggest economies in Europe and leaders in the race for AI innovation, but also because our research team covers these respective languages, which allowed us to capture the nuances and cultural specificities of media discourse.

The European context, we believe, is a fascinating field of study for two reasons. Firstly, thanks to its multicultural diversity, it offers us an alternative view to the often US-centric discourse of AI and algorithmic justice; secondly, the debate in Europe matters in particular because of Europe's legal and political particularities, which – notably due to EU legislations like the GDPR, Digital Services Act, the Digital Markets Act and the current work in progress on the EU AI Act – presents a legislative environment that could allow for legal and political possibilities to emerge.

Although our research was primarily concerned with three core countries, as we will explain in the methodology part, we also analyzed key articles from **other** countries, when we believed that they were relevant to defining the wider discourse on AI errors or for analyzing specific case studies (e.g. Google AI-powered dermatology assist tool).

Overall, this report discusses an array of case studies and instances that offer an overview of how different AI systems – between 2020 and 2022 - misunderstood and mismeasured humans and how they were covered in the European media. After discussing our methodological choices, in this report, we have decided to divide these case studies in two different parts.

In the first part, titled **The Errors in the Technologies that Profile us,** we focus on the coverage of errors, inaccuracies, and biases in facial, speech, and emotion recognition technologies. The second part, titled **The Impact of AI Errors in Society**, we instead focus on the articles that covered how AI errors are impacting on different areas of society such as: *a) employment and work; b) crime and policing; c) health and d) social media censorship.*

Our research led us to the conclusion that even if news media are becoming an important space for the discussion of how AI technologies can mis-understand and mis-measure humans, media discourse is often defined by a "techno-solutionist" (Morozov, 2013) and sensationalist understanding of the power of AI, which seems to over-ride in-depth debates about algorithmic errors and the human rights implications of AI.

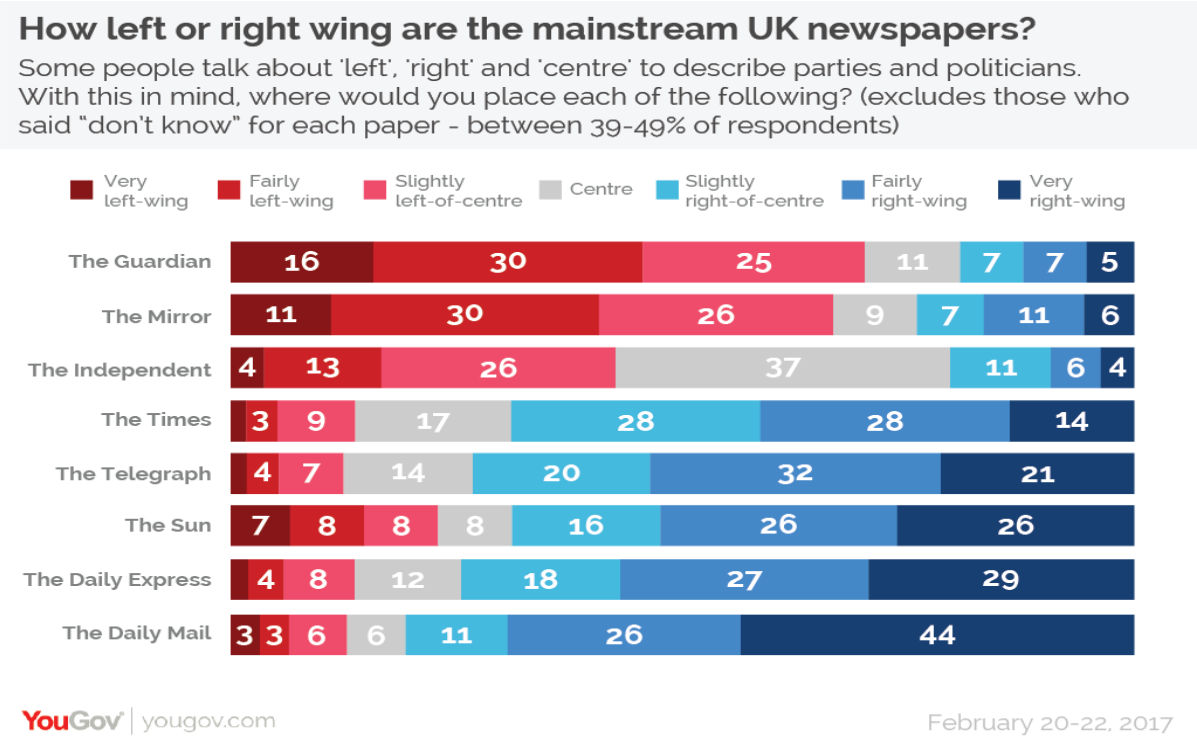
Yet, as this report will show within different media, we are seeing the rise of juxtaposed and contradictory narratives on 'the human error of AI'. This makes them an extraordinarily fascinating field of analysis when we want to shed light on the fact that the race for AI innovation is often shaped by stereotypical and reductionist understandings of human nature, and by new emerging conflicts about what it means to be human.

Methodology

In the period between **February 2020 and February 2022**, The Human Error Project Team carried out a **qualitative, critical discourse analysis** of 520 news media articles that reported on the uses of AI systems in the profiling of humans. For the selection of the articles, we concentrated on three core countries: Germany, France and the UK. As mentioned above, we selected these three countries because they are some of the biggest economies in Europe and thus play a fundamental role in the race for AI innovation, and because our research team covers the respective languages, which allowed us to capture the nuances and cultural specificities of the discourse in each country.

Relevant articles were selected and analyzed following a four steps process. Firstly, we focused on daily or weekly general interest newspapers with a national reach, including those newspapers that are often defined in popular culture as 'tabloid' or 'sensationalist press'.

For each of our key core countries, we selected five key outlets that were balanced across the political spectrum, including left-leaning, center-left/right, center or right-leaning. To map the political spectrum of newspapers - which varied from country to country depending on the media landscape - we relied on external sources such as the 2017 YouGov study in the UK (see figure below), MediaLanscapes.org and the results of a study run by the French statistical institute Ifop (2012).



In the UK, we selected two left-leaning news media (*The Guardian* and *The Independent*), two center-right outlets (*The Times* and *The Telegraph*), and two right-wing publications (*The Sun and The Daily Mail*) that are often considered tabloids. In Germany, we selected one left-wing newspaper (*Die Tageszeitung* - TAZ), two center/ left-leaning ones (*Süddeutsche Zeitung* and

Die Zeit), one conservative-liberal newspaper (*Frankfurter Allgemeine Zeitung - FAZ),* one conservative (*Die Welt*) and one tabloid (*Die Bild*). In France, we selected one left-wing newspaper (*Libération),* one center/left-leaning (*Le Monde*), one rather center/right-leaning (*La Croix*), one rather conservative-liberal newspaper (*Le Figaro*), one right-leaning newspaper (*Les Echos*) and one free daily newspaper (*20 minutes*).

In the period between Feb 2020 and Feb 2022, we looked for articles based on a predefined number of keywords which varied slightly according to the three languages used (German, French and English) by searching directly on the websites or online archives of the news media selected.

Germany	France	UK
KI Ethik	Éthique de l'IA	AI ethics
Algorithmische Verzerrung	Biais algorithmique	Algorithmic bias/ inaccuracy/ error
Algorithmische Profilbildung	Profilage algorithmique	Algorithmic profiling
Algorithmische Fehler	Erreur algorithmique	Algorithmic error
Datenschutz	Droits des données	Data rights/ human rights
Datengerechtigkeit	Justice des données	Data (in)justice/ harms/ (in)equality
Algorithmische Verantwortlichkeit	Responsabilité algorithmique	Algorithmic (un)accountability

Figure 1: Keywords by language

During our research we also used the same keywords on the global news monitoring database Factiva or Google News to identify articles in 'other' media. Thanks to this method - in the countries we studied - we identified articles in publications such as D*eutschlandfunk Kultur* in Germany or *BBC News* in the UK, as well as tech-specific publications like *Wired* in the English-speaking world and alternative media like *Netzpolitik* in Germany. We did this to follow specific research leads (e.g. a case of error in a particular technology) and to explore where the news media discourse fell short of crucial topics. The choice of following specific research leads and keywords also led us to analyze some articles from **other countries** such as the US, Switzerland and other European countries.

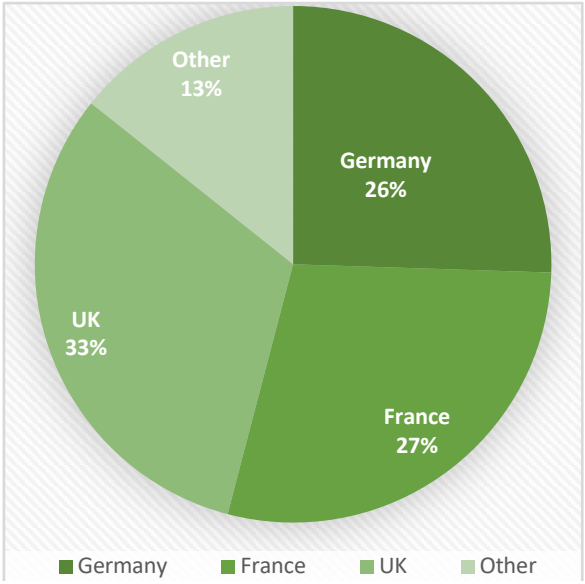


Figure 2: Articles analyzed by country

Our qualitative textual analysis was influenced by the will to focus on the syntagmatic understanding of meaning, namely on the way in which meaning is constructed by a relation of signs into a narrative sequence (Burgelin, 1972; Peterson, 2003). In our analysis, we were also very much influenced by the teachings of critical discourse analysis, and the belief that in the qualitative analysis of narratives, we needed to focus on a problem-oriented rather than a paradigm-oriented approach, one that does not only consider the narrative structure but also the hierarchies and structural inequalities in the framing of discourse (Van Dijk, 1996). In this framework, we were not so much interested in a quantitative analysis of media frames (namely counting the words used or the instances in which a frame appeared). Rather, we were interested in looking at the relationship between dominant narratives and rare narratives in media discourse, paying attention to the "how?" and "when?" of specific narratives.

Our data collection was thus structured as follows:

Key Information	Publication, Language, information on journalist
Content Description	Headline, summary, key quotes
Critical Discourse Analysis	Qualitative textual analysis that was aimed at highlighting the discourses and hierarchies of the text. This included paying attention to: 1. News structure (where is the article positioned?) 2. Thematic structures (what themes are given priority?) 3. Main ideology or message (Is there a main position/value in the article?) 4. Intertextuality (how does the article relate to other texts?)

Figure 3: Data Collection Process

The articles that we studied varied greatly in terms of thematic focus from news pieces directly discussing issues of algorithmic profiling and prediction to others that discussed other themes such as AI ethics, the geopolitical race for AI innovation or new regulations. In our analysis we decided to divide these different thematic focuses in 9 different categories:

Discrimination and bias	Articles focusing on cases of discrimination through algorithms or inbuilt biases, such as racial, gendered, ageist or ableist biases	18.6%
Human machine boundaries	Articles discussing the blurring borders between humans and machines, or human-machine communication, such as face and emotion recognition	16.9%
Profiling and prediction	Articles that explicitly problematize how data is assembled and interpreted with AI or algorithmic technologies to predict human behavior or character	16.6%
Techno-solutionism	Articles presenting AI or automated decision making as a solution to a problem	15.1%
Resistance to ADM	Articles about forms of resistance and how individuals or groups claim agency in the face of algorithmically produced or enhanced inequalities	7.5%
AI Regulation and policy	Articles focusing on policy discussions and regulation, such as DSA, DMA, GDPR, AI Act, Swiss E-ID	9%
Techno-skepticism	Articles primarily expressing a general critique or distrust towards AI	8.5%
AI Ethics	Articles focusing on how AI may be made ethical, or AI impact on human rights and society at large	5.1%
AI innovation race and geopolitics	Articles discussing the geopolitics of AI, particularly Europe's position versus China and US	2.9%

Figure 4: Percentage of Thematic Areas

As a second layer of research for this report, we have decided to divide the articles in two different parts:
The Errors in the Technologies that Profile us: where we explore the coverage of errors, inaccuracies, and biases in facial, speech, and emotion recognition technologies.

The Impact of AI Errors in Society: where we discuss the articles that covered how AI errors are impacting different areas of society such as: *a) employment and work; b) crime and policing; c) health and d) social media censorship.*

PART I:

The Errors in the Technologies that profile Us

1. AI Errors in Facial, Speech and Emotion Recognition

Over the last years, we witnessed key advances in the development and implementation of profiling technologies such as facial recognition, speech recognition or emotion recognition[3]. These technologies are being used in a variety of fields from education to HR, from criminal courts to policing.

The rapid rise in use of technologies that profile our faces, our voices and our emotions is leading to important debates about their inaccuracies and biases and the implications for society. These debates are especially important if we consider the fact that these technologies are used in contexts – such as policing – where their errors can directly endanger human rights.

Of course, debates about the fallacy of profiling technologies are not new in academia. Already in the early 2000s, facial recognition systems were understood as particularly problematic because they were more likely to recognize male and old subjects rather than female and young subjects (Philipps et al. 2003) or different ethnicities (Furl et al., 2002). Yet, over the last years, these debates have been at the top of the research agenda especially amongst those researchers that are concerned with AI ethics (see for instance Buolamwini and Gebru, 2018 on face recognition; Barassi and Scanlon, 2019 on speech recognition; or Crawford, 2021 on emotion recognition).

In the last few years, these debates have also extended well beyond the academic world and have become the center of new social conflicts. A good example comes from 2018, when Amazon workers protested the sale of facial recognition software to police forces, following similar protests at Google and Salesforce against governmental and military contracts (Leskin, 2018).

During our research we were interested in understanding how European news media were shaping the discourses on the fallacy of profiling technologies and we analyzed 54 articles reporting on **facial recognition technologies**, 43 articles reporting on **speech recognition technologies** and 34 reporting on **emotion recognition technologies**. The articles were mostly collected from the news media of our key countries (Germany, France and UK) but through keyword search we also found articles from other countries (US, Ireland, Netherlands and Switzerland).

The media coverage of profiling technologies reveals the growth and spread of these technologies in different areas of society and in everyday life. As shown in the table below (which summarizes some of the headlines of articles taken in 2021) journalists touched upon a plurality of examples of uses and applications of facial, speech or emotion recognition, in a wide variety of contexts from institutional settings (e.g. parliament, schools, prisons) to leisure settings (e.g. gaming, music and entertainment).

[3] The latter ones are the more complex profiling technologies as emotion recognition or affect detection aims to detect and classify emotions by analyzing any face." (Crawford, 2021, p.153) and also use indicators such as heartbeat, breathing rate and voice analysis.

Our analysis not only revealed the plurality of uses and applications of these technologies in everyday life and society, but also the fact that under umbrella terms such as face, emotion and speech recognition, journalists frequently referred to a wider variety of technologies. For instance, the articles that explored speech recognition technologies (abbreviated as SRTs) covered *both* technologies that were designed to recognize and profile humans on the basis of sound data (human voice, noises, coughs etc.) or text data (i.e. language used on social media, on an app, a translation tool etc.).

Figure 5: Headlines on Profiling Technologies in 2021[4]

In our analysis of the coverage of these profiling technologies we came across multiple conflicting and juxtaposed media discourses on AI errors. These discourses, we argue, reflect broader cultural narratives in Western Thought that construct the rise of AI with a mixture of awe and terror. Yet these discourses also shed light on the fact that media coverage is becoming a crucial terrain of reflection on the limits and errors of these technologies, and their implications for our society and human rights.

1.1 Media Coverage on AI: Between Sensationalist Fascination and Fear

In the coverage of profiling technologies perhaps the strongest discourse that we found was a **sensationalist fascination with the power of AI**, where AI technologies were framed not only as if they were human-like (e.g. they are able to read emotions, faces, understand speech) but as if they could **go beyond humans**, and enable us to grasp some of the mysteries of our human mind (e.g. their readings can enable us to understand ourselves, and our human situations better). This fascination with the power of AI, appeared very clearly in many of the articles we have analyzed that adopted a techno-solutionist perspective, which framed these technologies

[4] Please note that for all the in-text citations the titles are translated in English, the original title can be found in the reference list.

as **the potential solutions** to many different problems in our society and thus did not engage directly with the issue of error.

Yet we found these articles particularly interesting because they often engaged with questions about the difference between the power of these technologies vis-à-vis the power of humans. These articles were often shaped by an optimistic and enthusiastic tone which stressed how these technologies were solutions to our human limitations because **they understood human behaviors and contexts better than we do** and could thus be a support to us.

We found a fascinating example of this perspective in France in an article published by *Les Échos*, which talked about a speech recognition technology for nurses working in care homes. The technology was described as an "augmented ear" and was presented in the article as a way to relieve "mental load" of nurses during night shifts. The journalist describes how this technologies placed in the resident's room could detect abnormal noises like someone falling or being sick, and hence help nurses to react more quickly (Meijier, *Les Échos*, "Oso-AI launches a box co-constructed with nursing homes to manage nights", 21.09.2021).

Similarly, in an article that appeared in Germany on the *Süddeutsche Zeitung*, which instead focused on ERTs with dementia or Alzheimer's patients, the journalist claimed that these technologies can help us to understand and engage dementia patients in new ways. In the article, the journalist refers to the story of the daughter of a dementia patient that was surprised to hear that an 83-year-old would enjoy listening to fairy tales, which was something that she would have not known without the AI detecting his emotional state (Gierlinger, *Süddeutsche Zeitung*, "Digital emotion recognition: how artificial intelligence helps dementia patients", 24.07.2020).

During our analysis we found different examples of these type of reporting, defined by a positive fascination with these technologies, and the belief that they can understand us better than what we do. We found one interesting quote that highlights this perspective in *The Times*:

> "Henry James once advised that you can never really know exactly what is going on inside any human heart. To which scientists can now say: yes, you can you just need a radar system powered by artificial intelligence." (Blakely, *The Times*, "AI scan reveals mood in a heartbeat", 04.02.2021, §1)

The sensationalist fascination with the power of profiling technologies was not only prevalent within those news outlets that adopted techno-solutionist or optimistic stances, but also within those articles that revealed **a sense of fear, concern, and anxiety over the power of AI** technologies when it comes to profiling humans and their implications.

During our research we found different articles that discussed how AI solutions for human profiling may be **stripping humans of their sense of agency and autonomy.** For instance, an article that appeared in the *Süddeutsche Zeitung*, titled "The measurement of pupils", raised concerns about what it means for pupils to be under constant surveillance by ERTs in the classroom. The journalist quotes Felicitas Macgilchrist, a Göttingen-based researcher on the implications of educational software, who argues:

> "There is then no more invisibility," she says, "no moment when the pupil is not being observed. With this technology, students are constantly being datafied. But it's important for individuals to have a respite where they don't feel supervised. That's central to a sense of autonomy." (Füller, *Süddeutsche Zeitung*, "The measurement of pupils", 02.05.2021, §6).

We also found articles that were concerned with the fact that ERT and FRT could **modify human emotions and behaviors**. For example, in an article that appeared on the German *Frankfurter Allgemeine Zeitung*, the journalist writes:

> "People who know that they are being permanently observed and their emotions analyzed change their self-image and their behavior […] In the extreme, people could in future find themselves forced to move in public spaces not only inconspicuously, but also always in an artificially good mood... (Lenzen, *Frankfurter Allgemeine Zeitung*, "Emotions don't make it easy for machines", 06.05.2021, §6)

In French news media, the idea of **users being deceived and manipulated by AI** was discussed specifically with reference to speech recognition technologies[5]. A columnist for *Le Monde*, for instance, argued that SRTs are or will soon be able to perform 'humanity' very well and quotes the psychiatrist Serge Tisseron who alerts the readers on the 'sweet' capacity of voice interactive devices to make users believe they have human skills. The psychiatrist also points to the dangers of such power of deception and control to fall in the hands of the private companies, which design them and demands a public debate about the place and role we wish to give to machines before they become so empathetic and nice that we don't want to question them anymore (Tisseron, *Le Monde*, "Connected speakers could turn out to be formidable instruments of manipulation", 29.12.2021).

Journalists' discussion over current regulatory works also tended to revolve around the concern of machines – and the owners of these machines - manipulating humans. A journalist from the French left-leaning media *Libération* reports for instance on the emergence of dead bots which are text-based or voice-based chatbots simulating a conversation with a deceased person. Quoting the report of the National Committee of Digital Ethics in France, she eventually questions how all chatbots in general can "create a dangerous blurring between human and non-human" (Touré, *Libération*, "Deadbots: AI, life after death", 17.11.2021, §2).

During our research we found that the articles that presented a sensationalist fascination with AI (whether it was awe or fear) were particularly interesting in shaping the discourse about profiling technologies, for two main reasons.

On the one hand, this **media discourses reflected some of the dominant cultural narratives on AI** that are present in our society. In fact, this tension between fascination and fear, excitement, and anxiety that we found in media discourse has defined the entire history of artificial intelligence, which has been polarized between positive and negative perceptions (McCorduck, 2004). Academic research on technologies teaches us that this tension is not only

[5] Examples include (Leglu, Sciences et Avenir, "Attention, a chatbot is a robot and not a human being", 10.11.2021; Tisseron, Le Monde, "Connected speakers could turn out to be formidable instruments of manipulation", 29.12.2021; Champenois, Libération, "The connected speaker, services and vices" 30.12.2021).

an aspect of our relationship with AI, but it defines our relationship to new technologies in general. In fact, the Western fascination with the 'newness' of technologies has enabled us to construct mythical and sublime understandings of technologies (Mosco, 2004) and to perceive these technologies as if they have agency and power of a sort. Yet, when we think about the current cultural narratives and myths that shape the discourse around AIs – which we also find in news media discourse - we need to realize that these discourses are not simply emphasizing the advent of 'new technologies' - like the telegraph, satellite technologies or internet technologies - that can transform the organization of society. What they are clearly emphasizing is the advent of a 'new phase of human existence'; one where humans and machines would not only coexist, but also dialogue, empathize or even, in the worst-case scenario, run into conflict (Bory 2019).

On the other hand, we found that this **sensationalist coverage** of AI was particularly interesting for another reason **because this media frame over-rode reflections on AI bias, inaccuracy or failure.** Yet these reflections were starting to emerge in news media**.** In fact, during our research we noticed that journalists reported on a wide variety of errors when it came to profiling technologies.

1.2 Making sense of the Errors in our Profiling Technologies

Our research revealed that journalists were starting to dedicate some attention to the issue of errors in profiling technologies. This was not only particularly true with reference to facial recognition technologies (see section 2.2 on crime and policing) but also with reference to speech recognition technologies which were often presented as **disappointing technologies failing to meet expectations** about their smartness or predictive accuracy. In an article from *Le Monde* published at the occasion of Apple's Siri vocal assistant's tenth anniversary, for instance, the journalist shows how the adoption, functioning and use of smart speakers and voice interactive technologies in general are narrow and limited in France and in the United States. With irony, the journalist mocks the technology and comments on one of the most frequent uses of smart speakers which consists in asking the vocal assistant for jokes: "One wonders whether users always laugh with the assistant, or sometimes at its expense." (Six, *Le Monde*, "Tenth anniversary of Siri: has voice recognition conquered our daily lives?", 13.10.2021, §12).

The reflection on errors was at times very sensationalist in style. The "penny challenge" case was particularly illustrative in this regard. The case refers to how Amazon's vocal assistant Alexa dared an American ten-year-old girl to touch a live plug with a penny. It has been reported as a scandal in France, UK and Switzerland[6].

[6] See for instance (Le Parisien, "Alexa, Amazon's voice assistant, encouraged a 10-year-old child to touch an electrical outlet", 29.12.2021; Champenois*, Libération, "Smart speaker: services and vices", 30.12.2021; BBC News, "Alexa tells 10-year-old girl to touch live plug with penny", 28.12.2021; Barbey, Heidi News, "Alexa challenges a child to touch an electrical outlet with a coin", 29.12.2021)*

Other times reporting on errors was very granular, and journalists reported on very specific examples. Interesting in this regard is the case that was reported in February 2021, in an article on the *The Daily Mail* in the UK that reported on the case of censorship of a YouTube channel dedicated to chess play, when the words "black", "white", "attack" and "threat" used to describe chess strategies were interpreted as racist speech and led to an automatic suspension of the channel (Avery, The *Daily Mail*, "YouTuber blocked for discussing 'black versus white' chess strategy", 19.02.2021).

What we found interesting to notice, however, is that there was a **plurality of ways in which journalists explained why AI errors occurred.** The most common narrative was the understanding that AI made mistakes because **these technologies lack "the human factor"** in the decision-making process. In an article on the use of ERTs in schools that appeared in *Le Monde* in France, for instance, the journalist quotes Corinne Kolinsky, a teacher-researcher at the University of the Littoral Opal Coast in charge of educational innovation, which highlights the difference between a teacher and an ERT: "A teacher will try to understand why a student systematically chooses the wrong answer and try to correct it. A machine is not capable of this type of analysis" (Dautresme, *Le Monde*, "Is "adaptive learning" a revolution in teaching?", 20.05.2021, §13). Another article taken from *Die Zeit* in Germany makes a similar claim:

> "The human factor cannot be replaced by any machine," says Heinz-Peter Meidinger, president of the conservative German Teachers' Association. Learning is "always a social event" that requires contact with the teacher." (Schönert and Spiewak, *Die Zeit*, "Smarter Learning", 08.09.2021, §29)

Other examples of this media frame we found in the coverage of the errors in medicine and health (see section 2.3) and the coverage of errors made by automated vehicles (AVs). For instance, in Germany, news reporting in *Die Tageszeitung* argued that these vehicles: "will not function usefully without a certain amount of human intelligence and emotion" (Ronzheimer , *Die Tageszeitung*, "The human being at the centre", 26.08.2021, §1). An article on *The Times* had a similar perspective and the journalist argued that: "The assumption that every technological innovation improves our lives is surely misguided". The journalist denounces such technological expectations as "madness" and draws on previous cases of accidents by self-driving cars to demonstrate how "rely[ing] on the car's automated lane-keeping system technology with no human input" is problematic. The journalist emphasized the advantage of having "human sight and hearing" for crash avoidance which AVs do not possess bringing the question of human-centricity in the fore (Phillips, *The Times,* "Am I wrong to hate drones and driverless cars?", 11.05.2021, §9).

What we found very interesting during our research, is that media discourse on AI errors in profiling technologies was often structured around **a dichotomy between humans and machines**, which tried to make sense of the difference between human and machine intelligence. This media coverage, we believe, is a crucial example of a broader cultural reflection emerging about **why and when AI matters**. For instance, in an article from *Die Zeit*, on speech recognition technologies used in apps supporting users' mental health, the journalist tries to highlight the distinction between the AI and the therapist:

> "A good therapist is someone who helps when in need. But it is not his or her job to discover those who need help and do not have enough strength to find it. Machines could detect those people. They could make sure that everyone who needs it has access to the health system. And because people in the 21st century have to deal with smartphones all the time, they could possibly make more objective diagnoses than therapists." (Daum, *Die Zeit,* "Can machines heal souls?", 12.07.2021, §33)

The coverage of AI errors therefore was particularly fascinating as it highlighted how media discourse is defined by a process of reflection and negotiation on the relationship between **human and artificial intelligence**, and the ways in which these two **cannot replace one another** but can coexist.

The coverage of AI errors in profiling technologies was also particularly interesting because it highlighted how journalists' reflections on errors in profiling technologies were often trapped in a tension between the understanding that these **errors as technical dysfunctions to be 'fixed'**[7] with good and varied data or with updates to a software design, and the acknowledgement that these **errors as 'unfixable'**[8] because AI systems cannot really grasp the complexity of human experience and often reproduce structural inequalities.

1.3 The Human Error in AI: A Rare but Existing Media Frame

Within those media narratives that understood AI errors as unfixable, we found different examples of media frame that stressed **the impossibility for algorithms to ever understand humans**. In an article that appeared on the German *Die Tageszeitung* on speech recognition technologies, for instance, the journalist quotes Matthias C. Kettemann, the research program leader of the international project "Ethics of Digitization - From Principles to Practices", stating:

> "Humans are communicative beings, we only develop the image of ourselves in the eyes of others. So as long as people are talking, they will be creative – and as long as they are creative, the algorithms have a problem." (Wintermayr, *Die Tageszeitung*, 21.07.2021, "Hate speech online: When machines moderate", 21.07.2021, §17).

A similar invitation to accept that profiling errors in speech recognition will never be solved can be found in a cultural commentary titled "The Chatbot Problem" that appeared in *The New Yorker*, in the words of a Canadian novelist:

> "None of these safeguards will cure us of what needs curing. We are being forced to confront fundamental mysteries of humanity as technical issues: how little we know

[7] Example of these approaches can be found in the following articles: Hidalgo Pérez, *El Pais* and *Le Figaro*, "FluencyBank, the database that studies stuttering and helps Siri understand us better", 18.03.2021; Argiroffo and Butticaz, *RTS* Info, "Removing accents from a voice can make it more understandable", 23.01.2022. *20 minutes*, "Google Translate: A new approach to limit gender biases", 23.04.2020; Ferran, *Le Figaro*, ""Voice 1" or "Voice 2": to be more inclusive, Siri will no longer necessarily be a man or a woman", 04.04.2021)

[8] Examples of these approaches can be found in Hidalgo Pérez, *El Pais and Le Figaro*, "FluencyBank, the database that studies stuttering and helps Siri understand us better", 18.03.2021; Carr, *The Daily Mail*, "University lecturer slams 'sexist' Google Translate", 24.03.2021; Brühl, *Süddeutsche Zeitung*, "Ethical AI: artificial intelligence between good and evil", 03.04.2021; The Independent, "In Africa, rescuing the languages that Western tech ignores", 23.12.2021).

> about the darkness in our hearts, and how faint our control over that darkness is. (Marche, *The New Yorker*, "The chatbot problem", 23.07.2021 §12).

At times journalists expanded on these understandings of AI and human reductionism by referring to the problem of structural inequalities. Yet in our research we found it interesting to notice that reflections on **AI errors in connection to structural inequalities** - which is usually a key discourse in academia - was not so dominant in the media that we analyzed and emerged **mostly within UK left-leaning media** (although the role of The *Daily Mail* must be noted).

One of the most interesting cases that we came across in our analysis, which reflected on the relationship between AI errors and structural injustices was the case of the UK Government's online passport application software which allows British citizens to renew, replace or update their details online. The articles that covered the case explained that this was an application software that screens the details submitted by the applicant and matches these details with the required UK passport rules. This verification software was at the center of media attention because it was found to be racist towards darker skinned people and especially towards women with darker skin color. In one article on the *BBC News* website, the journalist explains: "[p]hotos of women with the darkest skin were four times more likely to be graded poor quality, than women with the lightest skin" (Ahmed, *BBC News*, "UK passport photo checker shows bias against dark-skinned women", 08.10.2020, §14). Citing one of the victims to this system, Ms. Elaine Owusu, a black student, the article indicated the issue of "systemic racism" inbuilt within the software where the lips of black people were misidentified as 'open mouth':

> "If the algorithm can't read my lips, it's a problem with the system, and not with me…I shouldn't have to celebrate overriding a system that wasn't built for me. It should be the norm for these systems to work well for everyone" (Ahmed*, BBC News*, "UK passport photo checker shows bias against dark-skinned women", 08.10.2020, §17-20).

Another case that was strongly reported by the UK news media was the presence of algorithmic error in Uber's face identification software. The identification software provided "real time ID check" by prompting Uber Eats couriers and drivers to verify their identity by submitting their photograph through the software, which was then compared against the photograph stored in the company's database through a "photo comparison" tool. An article from *Wired* reported about the racial bias inbuilt within the software as Uber couriers belonging to the BAME community have been fired as they struggle to prove their identity to Uber's dysfunctional face identification software. To demonstrate the intensity of the error within the software, the article cited David's (an Uber courier) experience with the software:

> "David* had shaved his beard in preparation for a job interview when he logged into the app. When he submitted a selfie, the app said the photo was not of him and asked David to provide information about the person that was replacing him as a courier within 24 hours or risk termination." (Kersley, *Wired*, "Couriers say Uber's 'racist' facial identification tech got them fired", 01.03.2021, §6)

Our analysis revealed that the media discourse on **structural inequalities** of profiling technologies was not limited to facial recognition technologies but extended also to **emotion recognition** and speech recognition. With reference to emotion recognition different articles

appeared in the UK on *The Guardian* and *The Independent* reflecting on the science behind emotion recognition technologies.

> "ERT raises another concern: the science of emotion behind it is controversial. Most ERT is based on the theory of "basic emotions" which holds that emotions are biologically hard-wired and expressed in the same way by people everywhere...Research in anthropology shows that emotions are expressed differently across cultures and societies. In 2019, the Association for Psychological Science conducted a review of the evidence, concluding that there is no scientific support for the common assumption that a person's emotional state can be readily inferred from their facial movements. In short, ERT is built on shaky scientific ground." (Hagerty and Albert, *The Independent*, "Why we should be weary of AI that can read our emotions." 22.04.2021, §6-7)

A central figure quoted in these reflections on ERTs and algorithmic bias is the US scholar Kate Crawford. In her recent book *Atlas of AI: Power, Politics and the Planetary Costs of Artificial Intelligence*, Crawford demonstrates how the flawed idea that there are "universal emotional states which can be read directly from the face" (2021, p.17) originated in the first place, based on the theories of psychologist Paul Ekman, whose idea that there is a set of universal emotions, including happiness, fear, anger, sadness, disgust and surprise, is fundamental to how contemporary ERTs work. Crawford also shows why these ideas still stick, despite having long been discredited by anthropologists such as Margaret Mead and other scholars (Crawford, 2021, p.174). Her work was often cited as crucial to understanding algorithmic bias in ERTs, such as in an article that appeared in *The Guardian*:

> "[Kate] Crawford points to a systematic review of the science in 2019, led by psychologist and neuroscientist Lisa Feldman Barrett, which showed that while under the best recording circumstances, AI can detect expressions like scowls, smiles and frowns, algorithms cannot reliably infer someone's underlying emotional state from them. For example, people scowl in anger only 30% of the time, Barrett says, and they might otherwise scowl for other reasons having nothing to do with anger, such as when they're concentrating or confused, or they hear a bad joke, or they have gas." (Skibba, *The Guardian*, "The computer will see you now: is your therapy session about to be automated?", 24.06.2021, §17)

Reflections on the **structural inequalities** of profiling technologies and the many margins of errors when it came to reading humans were also present in the debate about **speech recognition**. The most fascinating example appeared in a long read in *The Independent*. The journalist details how today's main SRTs make significant errors when they deal with African languages in both spoken and written forms. The chief data scientist interviewed in the article **connects these types of errors to a colonialist past** and explains the link historical inequalities:

> "The history of the African continent and in general in colonized countries, is that when language had to be translated, it was translated in a very narrow way," he said. "You were not allowed to write a general text in any language because the colonizing country might be worried that people communicate and write books about insurrections or

> revolutions. But they would allow religious texts." (*The Independent*, "In Africa, rescuing the languages that Western tech ignores", 23.12.2021, §12-13)

Our research revealed that the media discourse on AI errors with reference to structural inequalities was primarily **driven by UK media and less prominent in our other core countries like Germany and France.** We also found out that the discussion on structural inequalities was much rarer in comparison to other more 'sensationalist' ways of reporting on AI error. Yet, at the same time, we found this media discourse particularly interesting as it related to the debate about the ethics of AI, the human right implications of profiling technologies and role of policy and regulations.

Part II:

The Impact of AI Errors in Society

2. Media Coverage of AI Errors in different contexts

Over the last years, the rise of big data coupled with key developments in artificial intelligence have radically transformed different areas of our society. In fact, today governing institutions, educational bodies, healthcare providers, businesses of all kinds and multiple other agents have not only started to turn every aspect of everyday life into data, a phenomenon usually defined as "datafication" (Mayer-Schonberger and Cuckier, 2013) but are using AI systems for automated decisions making.

Consequently, in understanding the Human Error of AI we cannot only reflect on the errors intrinsic to our technologies, we also need to reflect on **the different ways in which these errors are impacting different areas of society**. This latter point became clear to us during our critical discourse analysis, when we realized that journalists addressed different areas of society that were being affected by AI errors. Hence, we decided that it was important to zoom-into the particular contexts in which these errors were being experienced. In this second part of the report, therefore, we will focus on four different contexts of errors that were reported by news media. We believe that looking at the specific areas can shed light on the context-specificity of the impact of AI errors, and the context specific discourses and negotiations that are consequently emerging. (You can click on the below links if you wish):

2.1 AI Errors in Employment and Work

2.2 AI Errors in Crime Prediction and Policing

2.3 AI Errors in Health

2.4 AI Errors and Social Media Censorship

2.1 AI Errors in Employment and Work

One area of everyday life in which AI and algorithmic technologies become ever more present and which we thus chose to follow up in our analysis of European news media's reporting on algorithmic error, was the uses and errors of AI and algorithmic technologies at work. In our study, we found this theme discussed in different French, Swiss, German and UK media. Our selected news media reported about various case studies of algorithmic technologies employed in **different stages of the professional life cycle** – from recruitment to employment and redundancy.

The first case study of AI being used in work environments, which media reported about, was the use of **automated recruitment tools and the identification of talent** in hiring processes[9].

[9] Examples included, in France, *Les Echos* reported about the cases of companies such as GoJob and StedY, which automatically match temporary job offers with potential candidates (Molga, *Les Echos*, "Gojob brings AI to temporary recruitment", 06.12.2021; Molga, *Les Echos*, "Created by the robotics company Gorgé to hire its temporary workers, StedY breaks free", 04.12.2020). In the UK, *The Guardian* reported about the blind-hiring software like Be Applied, which assesses applicants anonymously so that factors like gender or ethnicity are unknown to recruiters (Pozniak, *The Guardian*, "The bias battle: how software can outsmart recruitment prejudices", 23.12.2020). In Switzerland, the media reported about a "Swiss placement platform" called Job Cloud, which "uses self-learning algorithms to search the Internet for the "most suitable" candidate before a job is even advertised" (Wüstholz and Kaufmann, *Die Wochenzeitung*, "The Software decides who gets the job", 15.04.2021, §6) or the Zürich-based start-up Lionstep, which similarly claims to "use algorithms to tell us whether a person fits a job or not" (Emch, *SRF*, "Computer says no: when the robot decides about your career, 14.03.2021, n.a.).

Although different media focused on the recruitment phase, other media highlighted how AI at work is also used to monitor and make decisions about **employee's career development** (*Les Echos*, "When AI drives talent management", 26.01.2021) **or monitoring levels of productivity** during home office, as reported in the German media (Witzeck, *Frankfurter Allgemeine Zeitung*, "How surveillance has increased in times of home office", 14.12.2021; Seufert, Die Zeit, "What is my boss allowed to do?", 26.01.2022; Wolfangel, *Die Zeit,* "The boss sees everything", 15.09.2021) or **termination of contracts**[10].

Within news media's discussions of AI at work the following two media frames stood out as particularly interesting: (1) **algorithmic bias in AI driven recruitment Vs human bias** (2) **algorithmic misjudgment, surveillance and workers' rights**

One of the most interesting aspects of coverage of AI at work was the open debate on the issue of bias. Many of the articles that we analyzed, whilst they acknowledged that these technologies can speed up the process of sorting through CVs, also dedicated considerable space to the problem of **algorithmic bias** and how this related to human bias. In Germany, journalists reported the fact that algorithmic decisions may prioritize men over women after being trained on historical data where there were more men in particular positions than women, as reported in Germany (Reddig, *Tagesspiegel*, "AI watches along", 19.07.2021; *Süddeutsche Zeitung*, "Do applicants have to adapt now?", 09.07.2021). In the UK, an article cited a study that revealed how Facebook's automated job ads delivery is "skewed by gender" (*The Independent*, "Study: Facebook delivers biased job ads, skewed by gender", 09.04.2021) whilst another reported the findings by the campaign group Global Witness (Tonkin, *The Daily Mail*, "Facebook is accused of discrimination after job adverts for male-dominated roles such as mechanics and pilots are found to be targeted at men", 09.09.2021).

The articles we analyzed also reported on other forms of discrimination. For instance, the Swiss public broadcaster SRF raised in their reporting about the afore mentioned Zürich based start-up Lionstep was the potential problem of **discrimination based on age** (Emch, *SRF*, "Computer says no: when the robot decides about your career, 14.03.2021), while an article in *Süddeutsche Zeitung* reported about the case of an algorithm used by the Austrian employment agency, which discriminated at the intersection of gender and age:

> "Starting in 2021, the algorithm was supposed to determine the labor market chances of unemployed people throughout the country and suggest who might be worth retraining. Points would have been deducted from the calculation of chances for being a woman, for having care responsibilities (but only for women), for being over 30 and especially for being over 50. Likewise for non-EU passports or health problems.

[10] For example, the Swiss free daily 20 minutes reported about the case of the Russian software company Xsolla, which allegedly fired 150 employees after an algorithm had marked them as "unproductive" (Manfredini, *20 minutes*, "The Russian company Xsolla fires 150 employees deemed unproductive by an algorithm", 09.11.2021). In the UK, the *BBC* reported about the Trade Union Congress' demand for further legal protection of workers so that it would not be possible for staff to be fired purely on the basis of an algorithmic decision (*BBC News*, "AI at work: Staff 'hired and fired by algorithm'", 25.03.2021).

> However, the project was paused because personal data may not be processed in this way." (Crysmann, *Süddeutsche Zeitung*, "Dangerous calculation games", 02.12.2020, §11)

The coverage on bias in recruiting technologies was particularly interesting because it often led journalists to explore the fact that AI technologies are not objective and that these systems often **reproduce existing stereotypes and inequalities**. Overall journalists seemed to share the perspective that **"algorithms, which are created by humans, also recreate human bias"** (Crispin, *The Guardian*, "Job-hunting is stressful and humiliating enough. Now robots judge our résumés.", 15.03.2021) and that this can be particularly problematic when it comes to recruiting. For example, an article on the Swiss *RTS Info* quotes a professor from Geneva University to make this point:

> "We are in a technicist myth that tells us: it's a software so it will be objective. Not at all! As there are men and very few women who, in the background, are developing these software, it cannot be objective" reacts Isabelle Collet, professor in Education Sciences at Geneva University" (Tran-Tien and Zufferey, *RTS Info*, "Facebook algorithm is sexist when it comes to job offers", 18.10.2020, §6).

One interesting finding about the debate about algorithmic bias in recruitment tools was that it led several journalists to raise the question of whether unbiased recruiting can really be possible and hence to ponder on who is better at taking less biased decisions: **a human or an algorithm?**

Indeed, some articles took precisely the opposite view to those reporting about algorithmic bias, by arguing that **algorithmic decision making can in fact fix human biases in recruiting**. In *The Guardian* article entitled "The bias battle: how software can outsmart recruitment prejudices", the journalist Helen Pozniak quotes Riham Satti, a clinical neuroscientist turned entrepreneur "who's created technology to make it easier for people to recruit staff more fairly, regardless of background, gender or ethnicity" (Pozniak, *The Guardian*, "The bias battle: how software can outsmart recruitment prejudices", 23.12.2020, §3).

The main understanding was that **algorithmic biases are easier to control than human biases** and that it would be possible to, know where biases lie. This was evident in different articles, one example in France written by Jérôme Ternynck, the founder of a company called SmartRecruiters, and published in France by Les Échos. In this opinion piece, the CEO discussed "Five ways to boost recruitment with AI", ultimately **arguing that AI-driven recruitment "can help eliminate discrimination**" (Ternynck, *Les Échos*, "Opinion | Five ways to boost recruitment with AI", 06.12.2021, §12).

In an article which appeared on *The Guardian* the journalist argued that whilst algorithms can fix human bias in recruitment they can also exacerbate discrimination:

> "The working class; single mothers; people with chronic health issues; people who have spent time in prison or rehab facilities – all are more likely to have gaps in their work

> history. And while there are countless websites that offer tips on how to explain those gaps or overcome a lack of references or credentials during an interview, that explanation doesn't matter if you can't even get your application or résumé in front of a human. And because many of these processes are not transparent, it can be difficult to challenge the algorithm's assessment or even know what part of your application is setting off the rejection." (Crispin, *The Guardian,* "Job-hunting is stressful and humiliating enough. Now robots judge our résumés.", 15.03.2021, §4)

Discrimination was not the only problem reported about AI in recruitment. Another issue highlighted in an article on *Le Monde*, is that of **sorting humans into categories of alleged desirability** and the fact that automated inferences can become a "nightmare for those deemed undesirable because they are deformed or unfit" (Eychenne, *Le Monde*, "With the infernal big data-machine learning couple, we are becoming more readable, more predictable, and dumber too", 01.11.2021, §1). Interestingly, the author traces this tendency of categorizing people through a longer history of how different sciences have always tried to explain and categorize human beings, arguing that AI technologies are but the most recent iteration of this tendency and criticizing, for instance, the assumption that human character may ultimately boil down to only five traits:

> "Our personalities, complex though they are, would boil down to just a few character traits, just as Euclid's geometry boils down to a few axioms. So, we are more or less open-minded, conscientious, extroverted, kind, neurotic. Five-character traits only which would explain a bit of everything: relationships with those close to us, our professional career, our ambitions, and other everyday decisions." (Eychenne, *Le Monde*, "With the infernal big data-machine learning couple, we are becoming more readable, more predictable, and dumber too", 01.11.2021, §3-4)

According to the journalist, part of the nightmarish quality of this development lies in the worry the these machines are also **rising human standards** and thus many would be judged as not fast, competent, smart or fit enough.

An article published in *Die Wochenzeitung* in Switzerland also debunks the promises of AI in recruitment technologies by highlighting how much they are **prone to errors and misjudgements of who a "suitable" candidate** for a job may be. They discuss this by referring to the start-up Retorio from Munich, which uses facial recognition technologies to profile candidates during job interviews:

> "Research by Bayrischer Rundfunk revealed how opaque and confusing Retorio's software makes decisions. For this purpose, the journalists tested the application software with actors who wore different clothes in front of the camera. The result: even minimal changes influence the software's assessments enormously. For example, wearing glasses or headgear led to a completely new result, even if the actress spoke the same text in the same tone of voice. Even irrelevant details changed the automatically calculated ratings – for example, people were rated as "more open" and "more conscientious" if there was a picture hanging in the background or a bookshelf."

> (Wüstholz and Kaufmann, *Die Wochenzeitung*, "The Software decides who gets the job", 15.04.2021, §8)

The issue of AI failure was not only discussed with reference to discrimination and bias but also with reference to AI systems that track and make decisions about **employers' productivity levels**. One case study that was discussed here was a software called Genesys, which amongst other programs, is used by different companies and which "can evaluate employees' performance and rank them" (Wolfangel, *Die Zeit*, "The boss sees everything", 15.09.2021), the article explains:

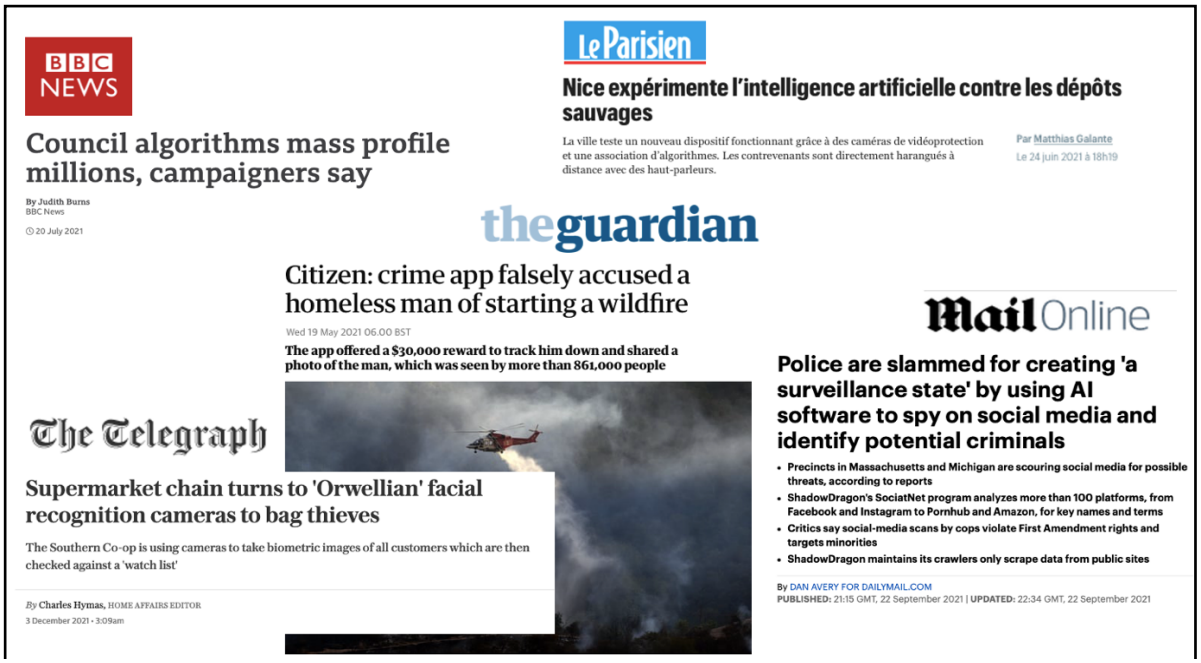
> "Customers can use the software to analyse the behaviour of their employees. The programme can evaluate data from SAP software, which records work activities and is used in many departments. And data from Microsoft programmes and from programmes like Salesforce, which remember who interacted with which customers and when. In themselves, the applications are handy. But they are treacherous when their data is linked together: The software then learns what behaviour is "common" and detects "undesirable" employee behaviour." (Wolfangel, *Die Zeit*, "The boss sees everything", 15.09.2021, §12)

One interesting debate that emerges from the coverage of these surveillance and profiling practices of workers is whether or not companies are **legally allowed** to conduct it. One article that was published by the BBC reported about the Trade Union Congress' claim that further legal protection is required in the UK as well. The article warns of the risk for work to become "dehumanized" and that **workers should have "a legal right to have any "high-risk" decision reviewed by a human"** (*BBC News,* "AI at work: Staff 'hired and fired by algorithm'", 25.03.2021, §3)

Another article that was published on *Die Zeit*, for instance, extensively quotes the Austrian data activist and researcher Wolfie Christl, whose organization Cracked Labs has published a study, which revealed the extent of respective technologies already being used in workplaces today. Referring to the case of the software Celonis, which is used to monitor and analyse employee behavior, he is quoted arguing that respective surveillance may in fact be unlawful:

> "A tool from Celonis can document what employees do on the screen with screenshots. Celonis calls this task mining; Wolfie Christl sees it as espionage. "In any case, the system processes personal data very excessively," he says, "so Celonis may be overstepping not only ethical but also legal boundaries." (Wolfangel, *Die Zeit*, "The boss sees everything", 15.09.2021, §12)

2.2 AI Errors in Crime Prediction and Policing



The use of AI and algorithmic technologies for **the prediction of human risk factors and crime prevention** stood out another key context in the media coverage of algorithmic prediction and error. A great majority of articles in this context concerned the uses of **facial recognition technologies (FRT) and other predictive policing** practices more generally, including, for instance, the use of body-worn, facial recognition cameras and CCTV cameras. We also came across one article reporting about police's screening of social media data.

Much of the reporting on algorithmic crime prediction was based on particular case studies, in which respective technologies were at play and produced errors, including in crimes as diverse as:

1. **domestic violence** (Smee, *The Guardian*, "Queensland police to trial AI tool designed to predict and prevent domestic violence incidents", 13.09.2021).
2. **shoplifting** (Hymas, *The Telegraph*, "Supermarket chain turns to 'Orwellian' facial recognition cameras to bag thieves", 03.12.2021)
3. **littering** (Galante, *Le Parisien*, "Nice experiments with artificial intelligence against illegal trash deposits, 24.06.2021; The Sun, "Bots nab litter louts", 01.03.2021)
4. **fraud** (Peitsmeier, *Frankfurter Allgemeine Zeitung*, "When the algorithm settles the fender bender", 17.01.2022; Zoske, *Frankfurter Allgemeine Zeitung*, "Algorithms should detect dubious investment advisers", 09.06.2021; Watt, *The Telegraph*, "Chinese scientists create 'world's first' AI prosecutor that will decide people's fates", 26.12.2021; Avery, *The Daily Mail*, "More than half of all US states make people filing for unemployment submit to facial-recognition software - and some data could be kept for more than 7 years", 23.07.2021; Booth, *The Guardian*, "Calls for legal review of UK welfare screening system which factors in age", 04.08.2021; Burns, *BBC News*, "Council algorithms mass profile millions, campaigners say", 20.07.2022; Crysmann, *Süddeutsche Zeitung*, "Dangerous calculation games", 02.12.2020)

5. **whistleblowing** (Seufert, *Die Zeit*, "What is my boss allowed to do?", 26.01.22; Wolfangel, *Die Zeit*, "The boss sees everything", 15.09.2021; Witzeck, *Frankfurter Allgemeine Zeitung,* "How surveillance has increased in times of home office", 14.12.2021)
6. **terrorism** (Bertrand and Maxwell, *Le Monde*, "Financing of terrorism: "Making artificial intelligence and data sharing central in anti-terrorism strategies", 19.06.2021; Ayad, *Le Monde*, "To fight terrorism, the fantasy of the preventive algorithm", 11.05.2021; Julien, *L'Express*, "Why AI will not eliminate terrorist threats", 05.11.2020; Bembaron, *Le Figaro*, "Algorithms to fight terrorism: what are we talking about?", 28.04.2021)
7. **child pornography** (Fischermann, *Die Zeit*, "Apple is Watching", 11.08.2021; Krönig and Böttger, *Bild Zeitung*, "The shock card of the child porn raid, 21.05.2021; Krönig and Böttger, *Bild Zeitung*, "Hunt for child porn criminals in Berlin!", 19.05.2021)
8. **knife crime** (Hern, *The Guardian*, "Study finds growing government use of sensitive data to "nudge" behaviour", 08.09.2021)
9. **money laundering** (Bertrand and Maxwell, *Le Monde*, "Financing of terrorism: "Making artificial intelligence and data sharing central in anti-terrorism strategies", 19.06.2021)
10. **cybercrime** (Hern, *The Guardian*, "Study finds growing government use of sensitive data to "nudge" behaviour", 08.09.2021)

The first finding that stood out from our analysis of the news media reporting on crime prediction technologies was that many **journalists were alarmed about the more wide-spread surveillance** that comes with them. For instance, journalists in the UK called respective technologies **"controversial" at best** – as *The Telegraph* called the police's use of facial recognition cameras to identify suspects from an existing databased of criminals' faces (Hymas, *The Telegraph*, "Police use facial recognition cameras for the first time", 10.12.2021) – or **"dystopian" at worst** – as *The Daily Mail* entitled the proposal for changes to the UK's Surveillance Cameras Code of Practice (Gant, *The Daily Mail*, "UK faces 'dystopian' future with facial-recognition AI cameras turning public spaces into 'open air prisons', warn privacy campaigners over new CCTV guidance given to police and councils", 17.08.2021).

A common trope that journalists used to illustrate this concern was a cultural reference to the British writer George Orwell's dystopian novel 1984, in which surveillance technologies have become omnipresent. Indeed, four of the articles we analyzed explicitly referred to Orwell or **used the term "Orwellian".**

Surveillance was not the only concern. **Government manipulation or influence** was also another issue discussed. In one article that appeared on *The Guardian*, for instance, the journalist reflected on the idea of "influence government" or "influence policing" in the UK,

> "Part of the programme also involves the NCA (National Crime Agency) collecting a substantial amount of data about the young people it visits, which can be used to craft profiles of the typical "at-risk" teen. Those profiles can then be used to run an "influence policing" campaign, using targeted advertising aimed at UK teens with an interest in gaming who search for particular cybercrime services on Google." (Hern, *The Guardian*, "Study finds growing government use of sensitive data to "nudge" behaviour", 08.09.2021, §9)

The journalist critically cites Scottish criminal justice researchers' concerns:

> "The rise of "influence government" could cause harm. Not only does it encourage departments to play fast and loose with personal data – using notes from an interview under caution to build a profile of a typical cybercriminal, for instance – it can also focus negative attention on vulnerable and disadvantaged groups in ways that could be counterproductive." (Hern, *The Guardian*, "Study finds growing government use of sensitive data to "nudge" behavior", 08.09.2021, §12)

One of the issue most discussed in the coverage of AI failure in crime was the problem of **racial bias**. One interesting example in the UK is the scoring computer model called Gangs Matrix, which "scores people's "risk" according to how likely a mathematical formula believes they are to commit gang violence based on their previous history of offences, patrol logs, social media activity and friendship networks" and which has been used by the UK's Metropolitan Police. In the article appeared on *The Telegraph* the journalist critically highlights that the tool, "caused controversy when it was revealed 1,000 young, Black men, all under 25, had been flagged in connection to gang violence. However, a review found they posed little to no risk of committing crime" (Meaker, *The Telegraph*, "Minority Report-style tech used to predict crime is 'prejudiced', 20.03.2021, §4-5).

The media coverage did not only engage with the question of racial bias, but also with the fact that predictive policing or crime prevention seemed to be affecting negatively **underprivileged groups**. For instance, in an interview about the current state and potential future use of AI by the police and forensic sciences in France, Jean-Gabriel Ganascia, who is introduced as "CNRS researcher and specialist in artificial intelligence", told *Le Parisien* that stigmatization may affect people from particular neighborhoods:

> "AI could direct law enforcement systematically to the same neighborhoods. The risk is therefore the stigmatization of certain places or certain populations." (Ho, *Le Parisien*, "Jean-Gabriel Ganascia: "Artificial intelligence has become an indispensable tool for the police", 12.10.2021, §4)

Another article that went into detail about the scale of the problem appeared in *Süddeutsche Zeitung*, and which reported about different cases of public authorities using "automated decision making" (ADM) tools, which "sometimes violate data protection and discriminate against certain population groups" in Denmark and The Netherlands, to detect child abuse (Crysmann, *Süddeutsche Zeitung*, "Dangerous calculation games", 02.12.2020, §9).

The article found particularly problematic a case that happened in Denmark, in 2019, when some municipalities had wanted "to use an ADM to assess where children might be at risk at home", but, "the data protection authority intervened". What is particularly striking here is that the algorithm would have selected parents based on categories or keywords attached to parents' files, such as ""divorce, "missed doctor's appointment", "mental health problems" or "job loss"". Another example the journalist cites was the case of an ADM used by the Dutch Ministry of Social Affairs, which "had to be banned by the courts". The journalist described the problem with this ADM as follows:

> "The algorithm had detected potential benefit cheats - but only in places with a particularly high number of low-income households. However, there were no indications that there was above-average cheating there." (Crysmann, *Süddeutsche Zeitung*, "Dangerous calculation games", 02.12.2020, §12).

Perhaps **more than in other fields, the issue of AI error in crime prevention seemed to be taken more seriously** than in the coverage of other areas of society. In reflecting on AI errors in profiling, media discourse did not only focus on issues such as racial bias or the stigmatization of poor people, but some articles also **raised questions about the fairness of technologies that cannot appreciate the complexity and the unpredictability of human experience**.

The aforementioned article that emerged in *The Telegraph*, for instance, raises concerns about **predictive policing's incapacity to fully grasp humans' ability to change** in the course of their lives, arguing not only that AI leads to bias being perpetuated "more quickly and more effectively", but that more fundamental problems are at stake. The journalist quotes a former police officer, named Tony Porter, who shared the same ideas:

> "To try and predict whether that same 17-year-old will go on at some point to commit certain types of crime under totally unknown and unforeseen circumstances, I think involves so many variables that I wonder why we would do it or even presume to think we can do it fairly and accurately." (Meaker, *The Telegraph*, "Minority Report-style tech used to predict crime is 'prejudiced', 20.03.2021, §10)

The media discourse on the errors implicit to crime prevention technologies **highlighted how crime-related decisions required human oversight.** This argument was made, for instance, in the aforementioned *Le Parisien* interview with researcher Jean-Gabriel Ganascia, who argued that, "Human presence then becomes essential because artificial intelligence issues generalities and does not deal with it on a case-by-case basis", but who nevertheless believed that we were heading towards a "lazy" society, "which delegates more and more tasks to machines, which uses artificial intelligence to automate everything, leaving the individual helpless, in disgrace" (Ho, *Le Parisien*, "Jean-Gabriel Ganascia: "Artificial intelligence has become an indispensable tool for the police", 12.10.2021, §4 and §6).

We found a similar approach in Germany in an interview published by the German public broadcaster *Deutschlandfunk Kultur* which cited the psychologist and author Gerd Gigerenzer, who Gigerenzer gives the example of different police departments in Germany, where the police in Hamburg deliberately abstained from algorithmic crime prediction, while Berlin and Bavaria "fell for the advertising hype" (Rahmlow, *Deutschlandfunk Kultur*, "Where the algorithm beats the human - and vice versa - Gerd Gigerenzer in conversation with Axel Rahmlow", 13.09.2021, n.a.).He then goes on to suggest **rules for deciding when decisions should be made by humans and algorithms respectively:** if a situation is predictable and the same rules apply yesterday, today and tomorrow (as in a chess game), algorithms could make better decisions than humans. However, if a situation is unpredictable (such as in the case of crime commitment), humans make the better decisions, and the use of algorithms should be critically interrogated.

2.3 AI Errors in Health

During our analysis we came across 77 articles, which covered issues relating to the use of **AI technologies in health,** and the errors that are emerging when these technologies misread and mismeasure specific health conditions. These reporting were based on the following different examples:

1. **skin condition detection** (Murgia, *Financial Times*, "Google launches AI health tool for skin conditions", 18.05.2021; Chadwick, *The Daily Mail,* "Hey Google, what's this rash?", 19.05.2021; *Stern*, "Google's new AI app aims to detect skin diseases", 01.06.2021; Bleiker, *Deutsche Welle*, "Rash or skin cancer? Google app to detect skin diseases", 23.05.2021; Georges, *Les Échos*, "When artificial intelligence tracks down skin diseases", 19.05.2021; Mertens, *20 minutes*, "I/O 2021: Google will launch an app to analyze cancerous moles", 24.05.2021; Lausson, Numerama, "Are you worried about a mole? Google will be able to analyze it with a picture", 18.05.2021; Davey, *The Guardian*, "Doctors fear Google skin check app will lead to 'tsunami of overdiagnosis'", 20.05.2021; *Le Figaro*, "ANAPIX medical: artificial intelligence at the service of dermatology", 11.06.2021; Randall, *The Daily Mail*, "Dr Google will see you now!", 29.03.2021; Bezmalinovic, *Tagblatt*, "Will Google Doctor soon follow Google Maps?", 22.05.2021; Büchi and Schlauri, *SRF Online*, "Diagnosis via app - choosing digital helpers correctly", 04.06.2021; Dave, *Reuters*, "Google searches for new measure of skin tones to curb bias in products", 19.06.2021; Madhusoodanan, *The Guardian*, "These apps say they can detect cancer. But are they only for white people?", 28.08.2021; *Feathers, Vice*, "Google's New Dermatology App Wasn't Designed for People With Darker Skin", 20.05.2021)

2. **eye condition detection** (Clémens, *Les Échos*, "Artificial intelligence optimizes ophthalmological diagnosis", 18.06.2021; Knight, *Wired US*, "AI Can Help Diagnose Some Illnesses—If Your Country Is Rich", 11.10.2020)
3. **kidney function estimation** (Wüstholz, *Algorithm Watch*, "Health algorithms discriminate against Black patients, also in Switzerland", 03.12.2020)
4. **Covid-19 detection** (Ksoll, *Bild Zeitung*, "How Whatsapp can now detect Corona", 18.02.2022; Knowles, *The Times*, "App can spot Covid cough by its sound", 09.03.2021)
5. **dental health monitoring** (Lentschner, *Le Figaro*, "DentalMonitoring, the new French unicorn", 21.10.2021)
6. **radiology** (Roy, *Le Figaro*, "Why artificial intelligence is far from replacing the radiologist", 13.09.2021; Le Dréau, *La Croix*, "The Therapeutic hopes of Artificial Intelligence", 15.02.2021)
7. **inferences made from X-Ray and medical scans** (Liberatore, *The Daily Mail*, "Artificial Intelligence can guess a person's race with up to 99% accuracy just by looking at their X-rays or other medical scans", 25.08.2021)
8. **AI ethics in health** (Boujemaa, *Le Monde*, "The eye of artificial intelligence less sharp than that of doctors?", 07.09.2021; Lenoir, *Le Figaro*, "Artificial intelligence has to remain a tool serving the doctor without making decisions for her/him", 02.04.2021; Chantadavong, *ZDNet*, "WHO campaigns for more ethics in AI", 03.07.2021)

Of the many different cases collected during our research the case of a recently launched Google dermatology app which was discussed by selected media in France, Germany, Switzerland and the UK has been particularly interesting. The application was launched in Europe in 2021 and asks users to upload images of their skin conditions which are then analyzed by an AI for possibly matching conditions.

Some of the articles that we came across seemed to be very upbeat about the technology. For instance, the *Financial Times* optimistically called the app in its subheading a "[b]reakthrough development [that] will assist users in self-diagnosing issues ranging from acne to melanoma" (Murgia, Financial Times, "Google launches AI health tool for skin conditions", 18.05.2021, emphasis added). In the article, the reporter **focuses primarily on the views, and opinions of the researchers, and developers of the app or experts that focus on the benefits** for early cancer detection and demonstrates a certain level of uncritical awareness and techno-optimism.

A similar approach can be found in the tabloid newspaper *The Daily Mail*, whose article about the app reads more as a Google Press Release with a headline such as "Hey Google, what's this rash? Google unveils AI tool that can identify 288 skin conditions including warts, cysts and lipoma from just THREE photos" (Chadwick, *The Daily Mail*, "Hey Google, what's this rash?", 19.05.2021).

These techno-optimistic approaches were found also in other countries. For instance, both the German media (*Stern*, "Google's new AI app aims to detect skin diseases", 01.06.2021) and French media (Georges, *Les Échos*, "When artificial intelligence tracks down skin diseases", 19.05.2021) presented Google's dermatology app as a solution to the high demand for dermatological advice, and the Swiss media (Bezmalinovic, *Tagblatt*, "Will Google Doctor soon follow Google Maps?", 22.05.2021) introduced it as a new innovation with the headline: "Will

Google Doctor soon follow Google Maps? The internet giant takes the first step in dermatology".

Whilst Google's app was welcomed by different news media as a possible solution to skin cancer diagnosis, different articles focused on the problem of misdiagnosis and errors, and their implications. For instance, in one article published by *The Guardian*, the journalist focuses on how promotional media stories of these apps often can conceal the truth on potential shortcomings. The article cites Dr. Ray Moynihan, Assistant Professor at Bond University's Institute for Evidence-Based Healthcare who explains this:

> "What we have at the moment are puffed up press releases and promotional media stories that make no mention of the potential downsides of these experimental AI tools – and one of the biggest downsides is unnecessary diagnosis, and the harm, anxiety and waste that can cause." (Davey, *The Guardian*, "Doctors fear Google skin check app will lead to 'tsunami of overdiagnosis'", 20.05.2021, §19)

Similarly, to what happened in the case of the reporting on profiling technologies, **the discourse on AI errors in health** was also shaped by the understanding that most of these errors were caused by a **lack of the "human factor".** In the reporting of AI in health this discourse was very strong as most articles emphasized the fact that unlike a real doctor, AI did not have the cultural understanding required to give a diagnosis or the capacity to understand how it is important to consider the whole person, the context and other aspects involved, which is precisely what humans do better than algorithms.

An example of this discourse could be found in an article published by the Swiss German media *SRF Online*, which stated: "One reason for the difficult digital diagnosis: the dermatologist not only examines the mole, but also the other skin changes." (Büchi and Schlauri, *SRF Online*, "Diagnosis via app - choosing digital helpers correctly", 04.06.2021, §8). Similarly, the French newspapers argued: "This is where the limit lies: AI is far from possessing the intuition and experience of the clinician, who will be able to adapt to each patient, to their anatomy, to their antecedents, to rare cases… " (Roy, *Le Figaro*, "Why artificial intelligence is far from replacing the radiologist", 13.09.2021, §4) or medical diagnosis more broadly (Lenoir, *Le Figaro*, "Artificial intelligence has to remain a tool serving the doctor without making decisions for her/him", 02.04.2021).

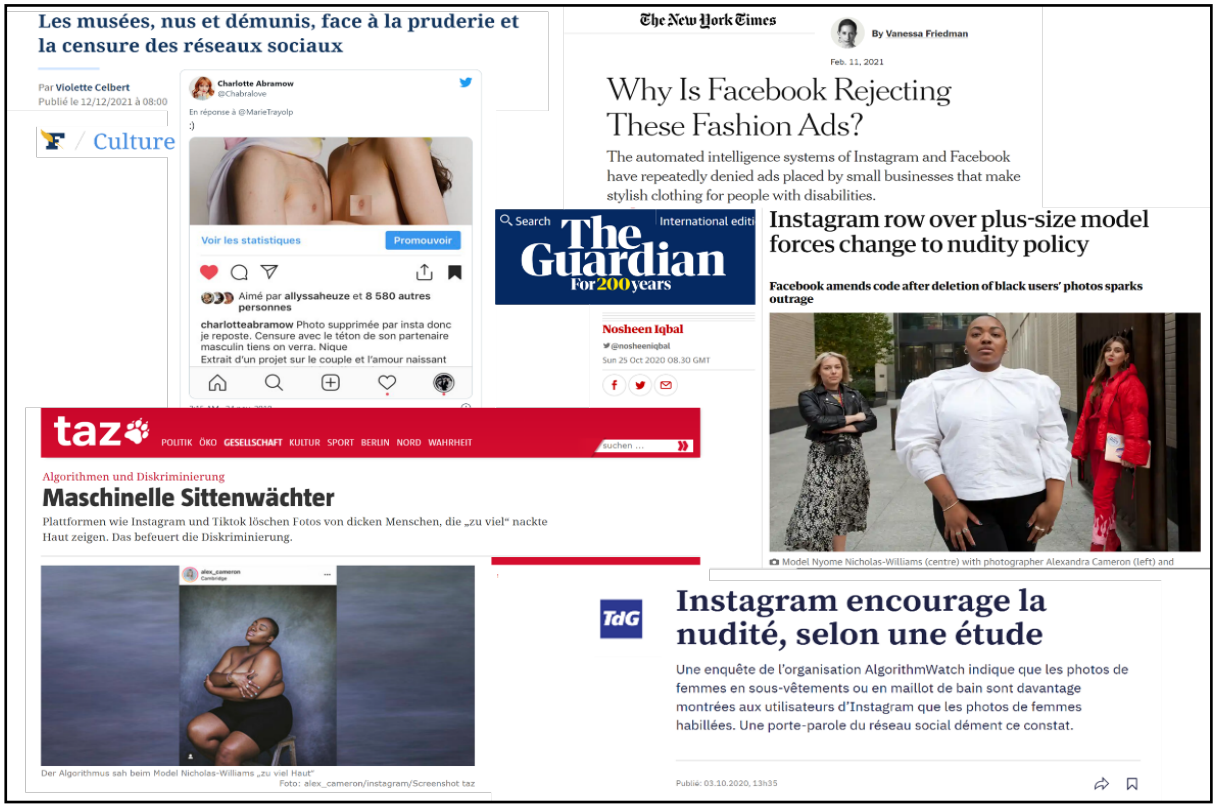
Whilst we found a predominance of discourse on AI errors in health which focused on the understanding that the error of these technologies was linked to a lack of human intelligence and "missing doctor", we also found **few articles that focused on the structural inequalities of data samples**. A *Reuters* news agency article, for instance, explored the issue of "six-color scale known as Fitzpatrick skin type (FST)" which is extensively used by dermatologists and even tech companies to analyze the effectiveness of technologies in detecting different skin colors (Dave, *Reuters*, "Google searches for new measure of skin tones to curb bias in products", 19.06.2021, §2). The issue here, as the article highlights, is that the scale is "a poor and overused measure for care, and often conflated with race and ethnicity" since it does not consider the variety of skin colors outside its six-color range. The Reuters elaborates this problem as:

> "The concern over FST is that its limited scale for darker skin could lead to technology that, for instance, works for golden brown skin but fails for espresso red tones… "Many people would assume I am skin type V, which rarely to never burns, but I burn," said Dr. Susan Taylor, a University of Pennsylvania dermatologist who founded Skin of Color Society in 2004 to promote research on marginalized communities. "To look at my skin hue and say I am type V does me disservice."" (Dave, *Reuters*, "Google searches for new measure of skin tones to curb bias in products", 19.06.2021, §7-14)

During our analysis, we also found different articles mentioning the intersection between race and gender bias in AI health, yet these discourses were not fully developed. As we argue elsewhere (Barassi and Patra, 2022), the prominence of discourses on the 'missing doctor' and on the more sensationalist discussion on the difference between "the human" and "the machine" overshadowed a more critical media discourse on health technologies and scientific bias.

However, what was clear from the media discourse was that these errors could lead not only to profound **individual implications - especially in cases of misdiagnosis -** but also to **clear social implications in cases of over-diagnosis.** For instance, the German *Deutsche Welle* highlighted the tendency of dermatological apps to "make overcautious mistakes" because, "when in doubt, [the app] would rather classify something as a cause for concern than overlook a disease". As the article explains, this can "have a negative knock-on effect due to a possible influx of requests for dermatological examination". Citing Prof. Eoin McKinney from the Center for Artificial Intelligence in Medicine at Cambridge University, the article concludes that this phenomenon could lead to: "panic among patients and high costs for the health system" (Bleiker, *Deutsche Welle*, "Rash or skin cancer? Google app to detect skin diseases", 06.12.2021, §11).

2.4 AI Errors and Social Media Censorship



During our analysis we also collected different news media articles reporting on algorithmic errors when it came to censorship of nudity on social media. The news articles we found were about artistic representations of bodies that museums, artists or activists shared on popular social media platforms such as Instagram, Facebook, Twitch or OnlyFans, and that were wrongly censored. Examples included:

1. **Museums facing algorithmic censorship of art including nudity** (Cherner, *Le Figaro*, "Mu.ZEE in Ostend censored by Facebook because of a photo of sculptures from the last century" 04.08.2021; Celbert, *Le Figaro*, "Museums, naked and helpless, facing the prudery and censorship of social networks", 12.12.2021; Steinitz, *Süddeutsche Zeitung*, "Is this still art or already porn?" , 18.02.2021; Hunt, *The Guardian*, "Vienna museums open adult-only OnlyFans account to display nudes", 16.10.2021; Kahlweit, *Süddeutsche Zeitung*, "30 000 years naked", 20.10.2021; AFP, *La Croix*, "Vienna's museums play provocatively against the censorship of nudity on the internet", 22.10.2021)
2. **Gender bias in nipple censorship** (Thomas, *Libération*, ""Madres Paralelas": on Instagram, the patriarchal nipple war", 12.08.2021 ; Trinks, *Frankfurter Allgemeine* Zeitung, "Mimicry of the nipple", 16.08.2021; Ostermann, *Le Figaro Madame*, "Nipples and breastfeeding women posted in Paris to counter Instagram censorship", 15.10.2021; Tutenges, *Slate*, "Facebook's algorithm blocks image of onions as "too sexy", 10.08.2020; Gayte, *Numerama*, "Twitch's secret rules confirm that women's nipples are banned, but not men's", 15.10.2021)

3. **Racist bias in body censorship** (Iqbal, *The Guardian*, “Instagram row over plus-size model forces change to nudity policy Facebook amends code after deletion of black users’ photos sparks outrage”, 25.10.2020; Lobe, *Die Tageszeitung*, “Machine Guardians of Morals”, 27.07.2021; Rutkowski, *Die Welt*, "Thin, white models post themselves almost naked. Why can't I?", 12.08.2020; Pons, *Le Figaro Madame*, "Nyome Nicholas Williams, the influencer who broke Instagram on nudity", 29.10.2020; Woitier, *Le Figaro*, “Instagram updates its policy on photos showing breasts”, 28.10.2020)
4. **Fatphobic bias in body censorship** (Rutkowski, *Die Welt*, "Thin, white models post themselves almost naked. Why can't I?", 12.08.2020; Woitier, *Le Figaro*, “Instagram updates its policy on photos showing breasts”, 28.10.2020; Metz & Yurieff, *CNN News*, “TikTok empowered these plus-sized women, then took down some of their posts. They still don't know why.”, 03.03.2021)
5. **Ableist bias in body censorship** (Friedman, *The New York Times*, “Why Is Facebook Rejecting These Fashion Ads?”, 11.02.2021; Lobe, *Die Tageszeitung*, “Machine Guardians of Morals”, 27.07.2021)
6. **Racist bias in ‘standard’ face beauty or recognition** (Hern, *The Guardian*, “Student proves Twitter algorithm ‘bias’ toward lighter, slimmer, younger faces”, 10.08.2021; Mahdawi, *The Guardian*, “This AI-powered app will tell you if you're beautiful – and reinforce biases, too” 06.03.2021; Ritschel, *The Independent*, “Female Only App criticized as Transphobic and Racist after AI Software Denies some Women” 27.01.2022; Fiévet, *We Demain*, “When artificial intelligence measures human beauty”, 13.07.2021)

An interesting case that we came across was the censoring by Instagram of the poster of *Madres Paralelas* by the Spanish director Pedro Almodóvar which shows a black and white picture of a nipple leaking a drop of milk framed in a red almond-shaped eye. From a distance, it looks like an eye with a white tear. The case was reported by different media in Europe such as the French Le Figaro Madame (Ostermann, *Le Figaro Madame*, “Nipples and breastfeeding women posted in Paris to counter Instagram censorship”, 15.10.2021) and *Le Figaro* (Woitier, *Le Figaro*, “Instagram updates its policy on photos showing breasts”, 28.10.2020) and the German F*rankfurter Allgemeine Zeitung* (Trinks, Frankfurter Allgemeine Zeitung, "Mimicry of the nipple”, 16.08.2021).

One of the fundamental discourses that were articulated in the news reporting was how **such forms of algorithmic decision making menaced artistic freedom of expression**. For instance, a journalist from the German *Frankfurter Allgemeine Zeitung*, who reported on the Almodovar’s case, engaged in a laudatory description of the movie poster, and discussed several artistic pieces referring to maternity to conclude that Instagram's algorithm represents a real threat to artistic expression.

> “But because the algorithm or the employees who are in focus or both are obviously blind to art, Instagram will continue to be a potential threat in the future: all Maria lactans depictions of art, in which the Mother of God is feeding her son with milk especially in the paintings of the meticulous early Netherlanders in which it can sometimes lead to the formation of drops, which are also a highly revered relic - and

> this is only incidental for the puritanical Instagrammerikans..” (Trinks, *Frankfurter Allgemeine Zeitung*, “Mimicry of the nipple”, 16.08.2021, §5).

The problem about algorithmic censoring and artistic expression was brought to the fore specifically with reference to the case of Austrian museums in Vienna that decided to open an account on OnlyFans - a social media platform known for its erotic content - to bypass social media platforms’ censorship of artistic pieces including nudity. The articles we analyzed reported on how the initiative was not simply a smart marketing move but was meant to question the power of algorithms over artistic freedom. The German media outlet *Süddeutsche Zeitung* for example quotes the head of Vienna Tourist Board Norbert Kettner who declares that “It cannot and should not be the case that algorithms help determine the art of tomorrow.” (Kahlweit, *Süddeutsche Zeitung*, “30000 years naked”, 20.10.2021, §5).

The discourse around the mistaken censorship of bodies did not only focus on the threat of algorithmic logics to artistic expression, but also **with reference to the censorship of ‘other’ bodies and algorithmic discrimination**. For instance, the case of censorship of the Almodovar poster, was described by the French *Libération* as a clear illustration of Instagram incoherent moderation policy as well as a sign of ongoing patriarchal and sexist attitudes towards women. With a sarcastic tone and reference to other censored artistic projects displaying nipples, the journalist Marlène Thomas sheds a different light on the theme of nudity and maternity:

> “If the publication of "female nipples" is strictly prohibited, in its great leniency, Instagram tolerates them "in the context of breastfeeding and in postpartum times". Would women's bodies be tolerable only when they play the mothering role?” (Thomas, *Libération*, ““Madres Parelelas”: on Instagram, the patriarchal nipple war”, 12.08.2021, §3)

Another case that came to the fore about the censoring of bodies and algorithmic discrimination was the one of the plus-size Black model Nyome Nicholas Williams[11]. The British model saw pictures of her posing semi-naked with her arms hiding her breasts being censored on Instagram. She consequently decided to launch the campaign #IwanttoseeNyome to force the platform to change its nudity policy, which eventually succeeded.

As journalists covered the Nyome Nicholas Williams’ case, it was interesting to notice how their focus **varied from one country to another**. *The Guardian* focused on the success of the model and her campaign team to change Instagram’s nudity policy and framed the issue as a feminist and collective victory for all women to be allowed to express their diverse beauty and lifestyles (Iqbal, *The Guardian*, “Instagram row over plus-size model forces change to nudity policy Facebook amends code after deletion of black users’ photos sparks outrage”, 25.10.2020).

[11] Examples include reported in *The Guardian* in the UK (Iqbal, *The Guardian,* “Instagram row over plus-size model forces change to nudity policy Facebook amends code after deletion of black users’ photos sparks outrage”, 25.10.2020), in *Die Tageszeitung* (Lobe, Die Tageszeitung, “Machine Guardians of Morals”, 27.07.2021) and *Die Welt* (Rutkowski, Die Welt, "Thin, white models post themselves almost naked. Why can't I?", 12.08.2020) in Germany and *Le Figaro* (Woitier, Le Figaro, “Instagram updates its policy on photos showing breasts”, 28.10.2020) in France in 2020 and 2021.

While the French media news magazine *Le Figaro Madame* underlined how the case was a sign of "a huge racial imbalance" and quoting the CEO of Instagram of "algorithmic partiality" disfavoring bodies of color (Pons, *Le Figaro Madame*, "Nyome Nicholas Williams, the influencer who broke Instagram on nudity", 29.10.2020, §4), its parent media company Le Figaro focused instead on a technical description of the censorship operated by Instagram and the consequent update implemented (Woitier, *Le Figaro*, "Instagram updates its policy on photos showing breasts", 28.10.2020).

The German *Die Welt* reported on the case focusing on how racist algorithmic biases are linked to the data or lack of data that the algorithms are trained with (Rutkowski, *Die Welt*, "Thin, white models post themselves almost naked. Why can't I?", 12.08.2020), and the *Die Tageszeitung* covered this case together with other cases of discrimination towards large bodies, bodies of color or bodies showing signs of disability to raise awareness on the problematic power of machines and of those who train them to decide which bodies should be seen or not (Lobe*, Die Tageszeitung*, "Machine Guardians of Morals", 27.07.2021).

This was one of the rare articles we found who referred to body algorithmic discrimination, with reference to **ableism**.

> "However, it is not only machines that attack our value system, but also people. For example, moderators of the popular app Tiktok were instructed to hide videos of people with disabilities. As research by Netzpolitik shows, characteristics such as "disfigured face", autism or downs syndrome were mentioned as examples in the guidelines. Queer and fat people were also to disappear from the stage. The range restriction was justified with bullying protection (Lobe, *Die Tageszeitung*, "Machine Guardians of Morals", 27.07.2021, §15)

What we found particularly interesting in our analysis of articles that reported on algorithmic wrongful censorship and discrimination of human bodies on social media platforms, was that the journalists at times reflected on this phenomenon by focusing **not so much on what was being censored but on what was being promoted by algorithmic logics**. For example, the Swiss media *La Tribune de Genève* reported on how incoherently Instagram's algorithm treats pictures of semi-naked or naked bodies. Although they point to the platform's responsibility in promoting "a certain physical norm" (*La Tribune de Genève,* "Instagram encourages nudity, according to a study", 03.10.2020, §11), notably referring to a French case of fatphobia, they do not explore further the roots of the issue nor suggest solutions or alternatives to it. On the contrary, the journalists from *Deutschlandfunk Kultur* who also reported on the study by Algorithm Watch focused their coverage on how sexist Instagram's algorithm can be considered and on what **regulations** exist and are needed to address the issue (Richter, *Deutschlandfunk Kultur*, "More naked skin means more visibility", 12.10.2020).

Conclusions

As Aradau and Blanke have shown (2021), concerns with errors, mistakes, and inaccuracies have shaped political debates about what technologies do and where and how certain technologies can be used and for which purposes. In their groundbreaking article the authors analyze the cultural history of error in scientific thought and the ways in which scientists have always had to coexist with errors especially when developing technologies aimed at measuring and analyzing human beings, such as the biometric technologies.

Aradau and Blanke also argue that when we think about machine learning and error there is much to gain if we develop a "history of the present" approach. Such an approach can enable us to understand how technological error "becomes a problem, raises discussion and debate, incites new reactions, and induces a crisis in the previously silent behavior, habits, practices, and institutions." (Foucault cited in Aradau and Blanke, 2021:2).

During our research we found out that news media are very much part of this "history of the present", as they are becoming the space where important reflections **about failure of AI technologies in algorithmic profiling are linked to broader questions about human reductionism, structural inequalities and human rights.**

The period of our analysis (February 2020 – February 2022), we believe is a particularly important period for the negotiation of AI technologies, because during these years the race for AI innovation and the rapid rise in used of AI systems for profiling humans was accelerated by the Covid-19 pandemic.

As we have shown, news media coverage in this period and in three different countries, was defined by different, contradictory, and juxtaposing discourses. One of the strongest discourses was a **sensationalist fascination with the power of AI** which was simultaneously defined by awe and fear. This sensationalist fascination often **overrode more in-depth and critical discourses** on AI failures and margins of error of these technologies.

However, what we found interesting was that in the period of our analysis and within the articles that we have looked at, these **discourses about the human error of AI were slowly emerging** and created a fruitful terrain for debate and reflection on the difference between human and machine intelligence and whether algorithmic technologies can really grasp the complexity and diversity of human experience.

Our research also reveals a clear cultural diversity in the ways in which news media were reporting on AI errors, with the debate on broader questions of structural inequalities being less advanced in Germany and France compared to their UK counterparts. Yet, in all three countries, we found that the coverage of AI errors opened up broader debates about the **human rights implications of these technologies**. In reflecting on human rights, most media discourse focused on the **right to privacy** and what it means to live in a world where we are constantly surveilled, tracked and evaluated. Yet the debate about AI errors also highlighted the **need for regulation**.

In this regard, many of the articles that we analyzed reflected on the position that is being played by the European Union in the global race for AI innovation. In one article that appeared in Germany for instance the journalists state that: "[t]he EU Commission is now attempting a balancing act: The authority wants to promote the development by setting up a global standard for AI technologies in contrast to the US and China" (Finke, *Süddeutsche Zeitung,* "What the EU's new AI law envisages", 21.04.2021, §1). In one article on emotion recognition technologies that appeared in the UK for instance the journalist quotes Margrethe Vestager, the European Commission's Executive Vicepresident for the digital age saying: "some AI uses should be off-limits completely in "democracies like ours," such as social scoring that can close off someone's privileges in society, and "broad, blanket use of remote biometric identification in public space." (Smith, *The Independent*, "AI systems can have catastrophic effects on human rights, UN says"16.09.2021, §15).

In conclusion, the juxtaposition of different media discourses, we believe demonstrates the fact that the race for AI innovation is triggering **fundamental questions about what it means to be human, about the diversity of human experience and whether technologies can really understand us.** What is also becoming clear is that there is no straightforward answer to these questions and no straightforward solution.

It is for this reason that we need to investigate the political life of AI errors, listen carefully to how different actors in society understand and negotiate with these errors, and what alternatives they are suggesting. This is what we hope to achieve with The Human Error Project. For us this report is just the beginning.

Academic references

Amoore, L. (2020). *Cloud Ethics: Algorithms and the Attributes of Ourselves and Others. Durham,* NC: Duke University Press.

Aradau, C. and Blanke, T. (2021). 'Algorithmic Surveillance and the Political Life of Error'. *Journal for the History of Knowledge,* 2(1). Available at: https://doi.org/10.5334/jhk.42 (Accessed: 5 December 2022).

Atanasoski, N. and Vora, K. (2019). *Surrogate humanity: Race, robots, and the politics of technological futures*. Duke University Press.

Barassi, V. and Scanlon, P. (2019). 'Voice Prints and Children's Rights'. Response to OHCHR Call for Submissions on the General Comment on Children's Rights in Relation to the Digital Environment, May 15th. Available at: https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation (Accessed 5 December 2022).

Barassi, V. (2020) *Child Data Citizen: How Tech Companies are Profiling Us from before Birth*, MIT Press

Barassi, V. and Patra, R. (2022). 'AI Errors in Health?'. *Morals & Machines*, 2(1), pp.34-43.

Benjamin, R. (2019). *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity Press.

Bory, P., 2019. 'Deep New: The shifting narratives of artificial intelligence from Deep Blue to AlphaGo'. *Convergence*, 25(4), pp.627-642.

Bourdieu, P. (1991). *Language and symbolic power*. Harvard University Press.

Buolamwini, J. and Gebru, T. (2018). 'Gender shades: Intersectional accuracy disparities in commercial gender classification'. Proceedings of Machine Learning Research. *Conference on fairness, accountability and transparency*, pp. 77-91. Available at: http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf (Accessed: 5 December 2022).

Brennen, J. S., Howard, P. N. and Nielsen, R. K. (2018). 'An industry-led debate: How UK media cover artificial intelligence'. *The Reuters Institute for the Study of Journalism*. Available at: https://ora.ox.ac.uk/objects/uuid:02126b4c-f4f9-4582-83a0-f8a9d9a65079/download_file?file_format=application%2Fpdf&safe_filename=Brennen%2B-%2BUK%2BMedia%2BCoverage%2Bof%2BAI%2BFINAL.pdf&type_of_work=Report (Accessed: 5 December 2022).

Broussard, M. (2018). *Artificial Unintelligence: How Computers Misunderstand the World*. The MIT Press.

Burgelin, O. (1972). 'Structural analysis and mass communication'. In: McQuail, D. (Ed.). *Sociology of mass communications*. Penguin, pp. 313-333.

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

Curran, J. and Seaton, J. (1997). *Power without Responsibility: Press and Broadcasting in Britain*. Routledge.

Eubanks, V. (2014). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Furl, N., Phillips, P.J. and O'Toole, A.J. (2002). 'Face recognition algorithms and the other-race effect: computational mechanisms for a developmental contact hypothesis'. *Cognitive Science*, 26(6), pp.797-815.

Gamson, W. A., and Modigliani, A (1989). 'Media discourse and public opinion on nuclear power: A constructionist approach. *American Journal of Sociology*, 95 (1), pp. 1-37. Available at: http://doi.org/10.1086/229213 (Accessed: 5 December, 2022).

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience.* Harvard University Press.

Hall, S (1997). 'Representation, meaning, and language'. In: Hall, S., Evans J and Nixon, S. (eds.) *Representation: Cultural representations and signifying practices*. SAGE, pp. 15-64 Mayer-Schönberger, V. and Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.

McCorduck, P. (2004). *Machines who think: A personal inquiry into the history and prospects of artificial intelligence*. CRC Press.

Milan, S. (2020). 'Techno-solutionism and the standard human in the making of the COVID-19 pandemic'. *Big Data & Society*, 7(2). Available at: https://doi.org/10.1177/2053951720966781 (Accessed: 5 December 2022).

Morozov, E. (2013). *To save everything, click here: The folly of technological solutionism*. Public Affairs.

Mosco, V. (2004). *The Digital Sublime. Myth, Power, and Cyberspace*. The MIT Press.

Munk, A.K., Olesen, A.G. and Jacomy, M. (2022). 'The Thick Machine: Anthropological AI between explanation and explication'. *Big Data & Society*, 9(1). Available at: https://doi.org/10.1177/20539517211069891 (Accessed: 5 December 2022).

Noble, S.U. (2018). *Algorithms of Oppression. How Search Engines Reinforce Racism.* New York University Press.

Ouchchy, L., Coin, A. and Dubljević, V. (2020). 'AI in the headlines: the portrayal of the ethical issues of artificial intelligence in the media'. *AI & Society*, 35(4), pp.927-936.

Peterson, M.A. (2003). *Anthropology & Mass Communication: Media and Myth in the New Millennium*. Berghahn Books.

Phillips, P. J., Jiang, F., Narvekar, A., Ayyad, J., and O'Toole, A. J. (2011). 'An other-race effect for face recognition algorithms'. ACM Transactions on *Applied Perception*, 8(2). Available at: http://doi.acm.org/10.1145/1870076.1870082 (Accessed: 5 December 2022).

Rettberg, J.W. (2022). 'Algorithmic failure as a humanities methodology: Machine learning's mispredictions identify rich cases for qualitative analysis'. *Big Data & Society*, 9(2). Available at: https://doi.org/10.1177/20539517221131290 (Accessed: 5 December 2022).

Richardson, R., Schultz, J.M. and Crawford, K. (2019). 'Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice'. *New York University Law Review*, 94, p.16-50.

Roessler, P. (2001). 'Between Online Heaven and Cyberhell The Framing of `The Internet' by Traditional Media Coverage in Germany'. *New Media & Society*, 3(1), pp. 49-66.

Shaikh, S.J. and Moran, R.E. (2022). 'Recognize the bias? News media partisanship shapes the coverage of facial recognition technology in the United States'. *New Media & Society*, Online First. Available at: https://doi.org/10.1177/14614448221090916 (Accessed: 5 December 2022).

Sun, S., Zhai, Y., Shen, B. and Chen, Y. (2020). 'Newspaper coverage of artificial intelligence: A perspective of emerging technologies'. *Telematics and Informatics*, 53. Available at: https://doi.org/10.1016/j.tele.2020.101433 (Accessed: 5 December 2022).

Van Dijk, T. (1996) 'Critical Discourse Analysis' in Tannen D., Hamilton H. and Schiffrin D. (Eds.) *The Handbook of Discourse Analysis* Wiley Blackwell.

Weaver, D.A., Lively, E. and Bimber, B. (2009). 'Searching for a frame: News media tell the story of technological progress, risk, and regulation'. *Science Communication*, 31(2), pp.139-166.

YouGov. (2017). 'How left or right-wing are the UK's newspapers?' YouGov. Available at: https://yougov.co.uk/topics/politics/articles-reports/2017/03/07/how-left-or-right-wing-are-uks-newspapers (Accessed: 5 December 2022).

News media articles references

AFP (2020) 'Instagram encourage la nudité, selon une étude', La Tribune de Genève, 3 October. Available at: https://www.tdg.ch/instagram-encourage-la-nudite-selon-une-etude-683956491065.
AFP (2021) 'Contre la censure de la nudité sur les réseaux, les musées de Vienne jouent la provoc', La Croix, 22 October. Available at: https://www.la-croix.com/Contre-censure-nudite-reseaux-musees-Vienne-jouent-provoc-2021-10-22-1301181923.
Ahmed, M. (2020) 'UK passport photo checker shows bias against dark-skinned women', BBC News, 7 October. Available at: https://www.bbc.com/news/technology-54349538
AI at work: Staff 'hired and fired by algorithm' (2021) BBC News. Available at: https://www.bbc.com/news/technology-56515827
AP (2021) 'Study: Facebook delivers biased job ads, skewed by gender', The Independent, 9 April. Available at: https://www.independent.co.uk/news/study-facebook-delivers-biased-job-ads-skewed-by-gender-facebook-linkedin-study-nvidia-domino-b1829347.html
Argiroffo, F. and Butticaz, E. (2022) Supprimer les accents d'une voix peut la rendre plus compréhensible, rts.ch. Available at: https://www.rts.ch/info/sciences-tech/12796888-supprimer-les-accents-dune-voix-peut-la-rendre-plus-comprehensible.html
Avery, D. (2021a) 'More than half of all US states make people filing for unemployment submit to facial-recognition software - and some data could be kept for more than 7 years', The Daily Mail, 23 July. Available at: https://www.dailymail.co.uk/sciencetech/article-9819499/Millions-Americans-seeking-unemployment-benefits-submit-facial-recognition-software.html
Avery, D. (2021b) 'YouTuber blocked for discussing "black versus white" chess strategy', The Daily Mail, 19 February. Available at: https://www.dailymail.co.uk/sciencetech/article-9279473/YouTube-algorithm-accidentally-blocked-chess-player-discussing-black-versus-white-strategy.html
Ayad, C. (2021) 'Pour lutter contre le terrorisme, le fantasme de l'algorithme préventif', Le Monde, 11 May. Available at: https://www.lemonde.fr/societe/article/2021/05/11/pour-lutter-contre-le-terrorisme-le-fantasme-de-l-algorithme-preventif_6079866_3224.html
Barbey, G. (2021) 'Alexa défie un enfant de toucher une prise électrique avec une pièce', Heidi News, 29 December. Available at: https://www.heidi.news/innovation-solutions/l-assistant-vocal-alexa-defie-un-enfant-de-mettre-ses-doigts-dans-la-prise

BBC News (2021) 'Alexa tells 10-year-old girl to touch live plug with penny', 28 December. Available at: https://www.bbc.com/news/technology-59810383
Bembaron, E. (2021) 'Des algorithmes pour lutter contre le terrorisme : de quoi parle-t-on?', Le Figaro, 28 April. Available at: https://www.lefigaro.fr/secteur/high-tech/des-algorithmes-pour-lutter-contre-le-terrorisme-de-quoi-parle-t-on-20210428
Bertrand, A., Maxwell, W. and Vamparys, X. (2021) 'Financement du terrorisme : « Mettre l'intelligence artificielle et le partage de données au centre du dispositif de lutte »', Le Monde, 19 June. Available at: https://www.lemonde.fr/idees/article/2021/06/19/financement-du-terrorisme-mettre-l-intelligence-artificielle-et-le-partage-de-donnees-au-centre-du-dispositif-de-lutte_6084778_3232.html

Bezmalinovic, T. (2021) 'Kommt nach Google-Maps bald Google-Doktor? Der Internetgigant macht den ersten Schritt in der Dermatologie', Tagblatt, 22 May. Available at: https://www.tagblatt.ch/leben/medizin-kommt-nach-google-maps-bald-google-doktor-der-internetgigant-macht-den-ersten-schritt-in-der-dermatologie-ld.2140055.

Blakely, R. (2021) 'AI scan reveals mood in a heartbeat', The Times, 4 February, p. 18.
Bleiker, C. (2021) Ausschlag oder Hautkrebs? Google App soll Hautkrankheiten erkennen, Deutsche Welle. Available at: https://www.dw.com/de/ausschlag-oder-hautkrebs-google-app-soll-hautkrankheiten-erkennen/a-57625767.
Booth, R. (2021) 'Calls for legal review of UK welfare screening system which factors in age', The Guardian, 18 July. Available at: https://www.theguardian.com/society/2021/jul/18/calls-for-legal-review-of-uk-welfare-screening-system-that-factors-in-age
Boujemaa, N. (2021) 'L'œil de l'intelligence artificielle moins affûté que celui des médecins?', Le Monde, 7 September. Available at: https://www.lemonde.fr/sciences/article/2021/09/07/l-il-de-l-intelligence-artificielle-moins-affute-que-celui-des-medecins_6093767_1650684.html
Brause, C. and Dowideit, A. (2022) 'Ist Instagram das gefährlichste soziale Netzwerk von allen?', Die Welt, 12 January. Available at: https://www.welt.de/politik/deutschland/article236147860/Instagram-Das-gefaehrlichste-soziale-Netzwerk-der-Welt.html?icid=search.product.onsitesearch.
Brühl, J. (2021a) 'Ethical AI: Künstliche Intelligenz zwischen gut und böse', Süddeutsche.de, 3 April. Available at: https://www.sueddeutsche.de/digital/ki-google-facebook-rassismus-sexismus-1.5254308
Brühl, J. (2021b) 'Facebook und Tiktok sind für mich systemgefähr- dende Hochrisiko-Technologie', Süddeutsche Zeitung, 22 April. Available at: https://www.sueddeutsche.de/digital/kuenstliche-intelligenz-eu-1.5272037.

Büchi, S. and Schlauri, E. (2021) Diagnose per App – digitale Helfer richtig auswählen, Schweizer Radio und Fernsehen (SRF). Available at: https://www.srf.ch/wissen/gesundheit/apps-fuer-die-gesundheit-diagnose-per-app-digitale-helfer-richtig-auswaehlen
Burns, J. (2021) Council algorithms mass profile millions, campaigners say, BBC News. Available at: https://www.bbc.com/news/uk-57869647
Carr, J. (2021) 'University lecturer slams "sexist" Google Translate', The Daily Mail, 24 March. Available at: https://www.dailymail.co.uk/news/article-9396937/University-lecturer-slams-sexist-Google-Translate.html
Celbert, V. (2021a) 'Les musées, nus et démunis, face à la pruderie et la censure des réseaux sociaux', Le Figaro, 12 December. Available at: https://www.lefigaro.fr/culture/les-musees-nus-et-demunis-face-a-la-pruderie-et-la-censure-des-reseaux-sociaux-20211212
Celbert, V. (2021b) 'Les musées, nus et démunis, face à la pruderie et la censure des réseaux sociaux', Le Figaro, 12 December. Available at: https://www.lefigaro.fr/culture/les-musees-nus-et-demunis-face-a-la-pruderie-et-la-censure-des-reseaux-sociaux-20211212
Chadwick, J. (2021) 'Hey Google, what's this rash? Google unveils AI tool that can identify 288 skin conditions including warts, cysts and lipoma from just THREE photos', The Daily Mail, 19 May. Available at: https://www.dailymail.co.uk/sciencetech/article-9595525/Google-unveils-AI-tool-identify-288-skin-conditions.html
Champenois, S. (2021) 'L'enceinte connectée, services et vices', Libération, 30 December. Available at: https://www.liberation.fr/lifestyle/hightech/lenceinte-connectee-services-et-

vices-20211230_BZ4ZXUVFMRCNFPP4222XSYVZEE/
Cherner, S. (2021) 'Le Mu.ZEE d'Ostende censuré par Facebook en raison d'une photo de sculptures du siècle dernier', Le Figaro, 4 August. Available at: https://www.lefigaro.fr/culture/le-mu-zee-d-ostende-censure-par-facebook-en-raison-d-une-photo-de-sculptures-du-siecle-dernier-20210804.
Clémens, M. (2021) 'L'intelligence artificielle optimise le diagnostic ophtalmologique', Les Échos, 18 June. Available at: https://www.lesechos.fr/pme-regions/innovateurs/lintelligence-artificielle-optimise-le-diagnostic-ophtalmologique-1324749
Crispin, J. (2021) 'Job-hunting is stressful and humiliating enough. Now robots judge our résumés', The Guardian, 15 March. Available at: https://www.theguardian.com/commentisfree/2021/mar/15/job-hunting-technology-algorithms
Crysmann, T. (2020) 'Gefährliche Rechenspiele', Süddeutsche Zeitung, 2 December. Available at: https://www.sueddeutsche.de/digital/werkstatt-demokratie-algorithmen-verwaltung-diskriminierung-datenschutz-1.5127319
Daum, P. (2021) 'Können Maschinen Seelen heilen?', Die Zeit, 8 July. Available at: https://www.zeit.de/gesundheit/2021-07/ki-psychotherapie-kuenstliche-intelligenz-therapeut-algorithmen-gpt-3/komplettansicht.
Dautresme, I. (2021) 'L'« adaptative learning », une révolution dans l'enseignement ?', Le Monde, 20 May. Available at: https://www.lemonde.fr/education/article/2021/05/20/l-adaptive-learning-une-revolution-dans-l-enseignement_6080819_1473685.html
Dave, P. (2021) 'Google searches for new measure of skin tones to curb bias in products', Reuters, 19 June. Available at: https://www.reuters.com/business/sustainable-business/exclusive-google-searches-new-measure-skin-tones-curb-bias-products-2021-06-18/
Davey, M. (2021) 'Doctors fear Google skin check app will lead to "tsunami of overdiagnosis"', The Guardian, 20 May. Available at: https://www.theguardian.com/society/2021/may/21/doctors-fear-google-skin-check-app-will-lead-to-tsunami-of-overdiagnosis

Deutschlandfunk Kultur (2021) 'Wo der Algorithmus den Menschen schlägt - und umgekehrt' (2021). Available at: https://www.deutschlandfunkkultur.de/sexistische-algorithmen-bei-instagram-mehr-nackte-haut-100.html.

DPA (2021) 'Müssen Bewerber sich jetzt anpassen?', Süddeutsche Zeitung, 7 September. Available at: https://www.sueddeutsche.de/karriere/arbeit-muessen-bewerber-sich-jetzt-anpassen-dpa.urn-newsml-dpa-com-20090101-210708-99-307059.
Emch, S. (no date) 'Computer says no. Wenn der Roboter über die Karriere entscheidet.' (Input). Available at: https://www.srf.ch/audio/input/computer-says-no-wenn-der-roboter-ueber-die-karriere-entscheidet?id=11946700
Eychenne, K. (2021) 'Karl Eychenne : « Avec le couple infernal big data-machine learning, nous devenons plus lisibles, plus prévisibles, plus bêtes aussi »', Le Monde, 1 November. Available at: https://www.lemonde.fr/idees/article/2021/11/01/karl-eychenne-avec-le-couple-infernal-big-data-machine-learning-nous-devenons-plus-lisibles-plus-previsibles-plus-betes-aussi_6100549_3232.html
Feathers, T. (2021) 'Google's New Dermatology App Wasn't Designed for People With Darker Skin', Vice, 20 May. Available at: https://www.vice.com/en/article/m7evmy/googles-new-dermatology-app-wasnt-designed-for-people-with-darker-skin

Ferran, B. (2021) '«Voix 1» ou «Voix 2» : pour être plus inclusif, Siri ne sera plus forcément un homme ou une femme', Le Figaro, 4 April. Available at: https://www.lefigaro.fr/secteur/high-tech/voix-1-ou-voix-2-pour-etre-plus-inclusif-siri-ne-sera-plus-forcement-un-homme-ou-une-femme-20210404
Fiévet, C. (2021) 'Quand l'intelligence artificielle mesure la beauté humaine', We Demain, 26 August. Available at: https://www.wedemain.fr/inventer/intelligence-artificielle-o-mon-ia-suis-je-toujours-le-plus-beau/
Finke, B. (2021) 'Was das neue KI-Gesetz der EU vorsieht', Süddeutsche Zeitung, 21 April. Available at: https://www.sueddeutsche.de/wirtschaft/ki-ai-eu-gesichtserkennung-intelligenz-1.5271653.

Fischermann, T. (2021) 'Apple schaut mit', Die Zeit, 11 August. Available at: https://www.zeit.de/2021/33/apple-iphone-scanning-kinderpornografie-datenschutz-ueberwachung-missbrauch-technologie.
Friedman, V. (2021) 'Why Is Facebook Rejecting These Fashion Ads?', The New York Times, 11 February. Available at: https://www.nytimes.com/2021/02/11/style/disabled-fashion-facebook-discrimination.html
Galante, M. (2021) 'Nice expérimente l'intelligence artificielle contre les dépôts sauvages', Le Parisien, 24 June. Available at: https://www.leparisien.fr/environnement/nice-experimente-lintelligence-artificielle-contre-les-depots-sauvages-24-06-2021-EUZV5KPGCZD5THZII2XK6VRUOM.php
Gant, J. (2021) 'UK faces "dystopian" future with facial-recognition AI cameras', The Daily Mail, 17 August. Available at: https://www.dailymail.co.uk/news/article-9900501/UK-faces-dystopian-future-facial-recognition-AI-cameras-turning-public-spaces-prisons.html
Gayte, A. (2021) Les règles secrètes de Twitch confirment que les tétons des femmes sont interdits, mais pas ceux des hommes, Numerama. Available at: https://www.numerama.com/politique/747577-les-regles-secretes-de-twitch-confirment-que-les-tetons-des-femmes-sont-interdits-mais-pas-ceux-des-hommes.html
Georges, B. (2021) 'Quand l'intelligence artificielle traque les maladies de peau', Les Echos, 19 May. Available at: https://www.lesechos.fr/tech-medias/hightech/quand-lintelligence-artificielle-traque-les-maladies-de-peau-1316371
Gierlinger, M. (2020) 'Emotionserkennung bei Demenzpatienten', Süddeutsche Zeitung, 24 July. Available at: https://www.sueddeutsche.de/digital/demenz-digitalisierung-emotionserkennung-1.4973879

Hagerty, A. and Albert, A. (2021) 'Why we should be weary of AI that can read our emotions', The Independent, 21 April. Available at: https://www.independent.co.uk/news/science/can-ai-read-emotions-b1833668.html
Hern, A. (2021a) 'Student proves Twitter algorithm "bias" toward lighter, slimmer, younger faces', The Guardian, 10 August. Available at: https://www.theguardian.com/technology/2021/aug/10/twitters-image-cropping-algorithm-prefers-younger-slimmer-faces-with-lighter-skin-analysis.
Hern, A. (2021b) 'Study finds growing government use of sensitive data to "nudge" behaviour', The Guardian, 8 September. Available at: https://amp.theguardian.com/technology/2021/sep/08/study-finds-growing-government-use-of-sensitive-data-to-nudge-behaviour?__twitter_impression=true.
Hidalgo Pérez, M. (2021) 'FluencyBank, la base de données qui étudie le bégaiement et aide Siri à mieux nous comprendre', Le Figaro, 18 March. Available at:

https://www.lefigaro.fr/secteur/high-tech/fluencybank-la-base-de-donnees-qui-etudie-le-begaiement-et-aide-siri-a-mieux-nous-comprendre-20210318

Ho, A. (2021) 'Jean-Gabriel Ganascia : «L'intelligence artificielle est devenue un outil indispensable pour la police»', Le Parisien, 21 October. Available at: https://www.leparisien.fr/faits-divers/jean-gabriel-ganascia-lintelligence-artificielle-est-devenue-un-outil-indispensable-pour-la-police-21-10-2021-YAH4JW4PGZAX5NE3ZAFWRGFDL4.php

Hunt, E. (2021) 'Vienna museums open adult-only OnlyFans account to display nudes', The Guardian, 16 October. Available at: https://www.theguardian.com/artanddesign/2021/oct/16/vienna-museums-open-adult-only-onlyfans-account-to-display-nudes

Hymas, C. (2021a) 'Police use facial recognition cameras for the first time', The Telegraph, 10 December. Available at: https://www.telegraph.co.uk/news/2021/12/10/police-use-facial-recognition-cameras-first-time

Hymas, C. (2021b) 'Supermarket chain turns to "Orwellian" facial recognition cameras to bag thieves', The Telegraph, 3 December. Available at: https://www.telegraph.co.uk/news/2021/12/03/supermarket-chain-turns-orwellian-facial-recognition-cameras/.

Iqbal, N. (2020) 'Instagram row over plus-size model forces change to nudity policy', The Guardian, 25 October. Available at: https://www.theguardian.com/technology/2020/oct/25/instagram-row-over-plus-size-model-forces-change-to-nudity-policy

Julian, S. (2020) 'Pourquoi l'intelligence artificielle ne permettra pas d'éliminer la menace terroriste', L'Express, 5 November. Available at: https://www.lexpress.fr/actualite/sciences/pourquoi-l-intelligence-artificielle-ne-permettra-pas-d-eliminer-la-menace-terroriste_2137857.html

Kahlweit, C. (2021) '30 000 Jahre nackt', Süddeutsche Zeitung, 20 October. Available at: https://www.sueddeutsche.de/kultur/pornografie-nacktheit-social-media-1.5443858

Kersley, A. (2021) 'Couriers say Uber's "racist" facial identification tech got them fired', Wired, 1 March. Available at: https://www.wired.co.uk/article/uber-eats-couriers-facial-recognition.

Knight, W. (2020) 'AI Can Help Diagnose Some Illnesses—If Your Country Is Rich', Wired, 11 October. Available at: https://www.wired.com/story/ai-diagnose-illnesses-country-rich/

Knowles, T. (2021) 'App can spot Covid cough by its sound', 9 March. Available at: https://www.thetimes.co.uk/article/app-spot-covid-cough-sound-diagnosis-f83w0kbfk

Kröning, O. and Böttger, D. (2021a) 'Die Schock-Karte der Kinderporno-Razzia', Die Bild Zeitung, 21 May. Available at: https://www.bild.de/bild-plus/regional/berlin/berlin-aktuell/schock-karte-der-kinderporno-razzia-bild-im-berliner-polizei-lagezentrum-76449128.bild.html.

Kröning, O. and Böttger, D. (2021b) 'Jagd auf Kinderporno-Verbrecher in Berlin!', Die Bild Zeitung, 19 May. Available at: https://www.bild.de/regional/berlin/berlin-aktuell/200-beamte-im-einsatz-kinderporno-razzia-in-berlin-76441220.bild.html.

Ksoll, M. (2022) 'Wie WhatsApp Corona am Husten erkennen kann', Bild Zietung, 18 February. Available at: https://www.bild.de/ratgeber/2022/ratgeber/whatsapp-kann-corona-am-geraeusch-erkennen-79196906.bild.html

Lausson, J. (2021) 'Un grain de beauté vous inquiète ? Google pourra l'analyser depuis une photo', Numerama, 18 May. Available at: https://www.numerama.com/sciences/712373-un-

grain-de-beaute-vous-inquiete-google-pourra-lanalyser-depuis-une-photo.html
Le Dréau, A. (2021) 'Les espoirs thérapeutiques de l'intelligence artificielle', La Croix, 15 February. Available at: https://www.la-croix.com/Sciences-et-ethique/espoirs-therapeutiques-lintelligence-artificielle-2021-02-15-1201140827
Le Figaro (2021) 'ANAPIX medical : l'intelligence artificielle au service de la dermatologie', 9 June. Available at: https://sante.lefigaro.fr/article/anapix-medical-l-intelligence-artificielle-au-service-de-la-dermatologie/
Le Parisien (2021) 'Alexa, l'assistant vocal d'Amazon, a encouragé une enfant de 10 ans à toucher une prise électrique', 29 December. Available at: https://www.leparisien.fr/faits-divers/alexa-lassistant-vocal-damazon-a-encourage-une-enfant-de-10-ans-a-toucher-une-prise-electrique-29-12-2021-J7JGTK63SBFQBAO4JBI7TZGLCA.php
Leglu, D. (2021) 'Attention, un chatbot est un robot et pas un être humain', 10 November. Available at: https://www.sciencesetavenir.fr/high-tech/web/les-enjeux-ethiques-des-chatbots_158994
Lenoir, L. (2021) '«L'intelligence artificielle doit rester un outil au service du médecin sans décider à sa place»', Le Figaro, 2 April. Available at: https://www.lefigaro.fr/vox/societe/l-intelligence-artificielle-doit-rester-un-outil-au-service-du-medecin-sans-decider-a-sa-place-20210402
Lentschner, K. (2021) 'DentalMonitoring, nouvelle licorne française', Le Figaro, 21 October. Available at: https://www.lefigaro.fr/societes/dentalmonitoring-nouvelle-licorne-francaise-20211021
Lenzen, M. (2021) 'Gefühle machen es Maschinen auch nicht leicht', Frankfurter Allgemeine Zeitung, 6 May. Available at: https://www.faz.net/aktuell/feuilleton/buecher/rezensionen/sachbuch/catrin-misselhorns-buch-kuenstliche-intelligenz-und-empathie-17325698.html.
L'Express (2021) 'Intelligence artificielle : une aide pour prédire le risque suicidaire chez les étudiants', 15 June. Available at: https://www.lexpress.fr/actualite/societe/sante/intelligence-artificielle-une-aide-pour-predire-le-risque-suicidaire-chez-les-etudiants_2152885.html
Liberatore, S. (2021) 'Artificial Intelligence can guess a person's race with up to 99% accuracy just by looking at their X-rays or other medical scans', Daily Mail, 25 August. Available at: https://www.dailymail.co.uk/sciencetech/article-9925845/AI-guess-persons-race-99-accuracy-just-looking-X-rays.html
Lobe, A. (2021) 'Maschinelle Sittenwächter', Die Tageszeitung: taz, 27 July. Available at: https://taz.de/!5785746/
Madhusoodanan, J. (2021) 'These apps say they can detect cancer. But are they only for white people?', The Guardian, 28 August. Available at: https://www.theguardian.com/us-news/2021/aug/28/ai-apps-skin-cancer-algorithms-darker
Mahdawi, A. (2021) 'This AI-powered app will tell you if you're beautiful – and reinforce biases, too', The Guardian, 6 March. Available at: https://www.theguardian.com/commentisfree/2021/mar/06/ai-powered-app-tell-you-beautiful-reinforce-biases
Manfredini, D. (2021) 'L'entreprise russe Xsolla licencie 150 employés jugés improductifs par un algorithm', 20 minutes, 9 November. Available at: https://www.20minutes.fr/high-tech/3165187-20211104-un-algorithme-licencie-150-employes-juges-pas-assez-productifs
Marche, S. (2021) 'The Chatbot Problem', The New Yorker, 23 July. Available at: https://www.newyorker.com/culture/cultural-comment/the-chatbot-problem

Meaker, M. (2021) 'Minority Report-style tech used to predict crime is "prejudiced"', The Telegraph, 20 March. Available at: https://www.telegraph.co.uk/technology/2021/03/20/minority-report-style-tech-used-predict-crime-prejudiced/

Meijier, I. (2021) 'Oso-AI lance un boitier coconstruit avec les Ehpad pour gérer les nuits', Les Echos, 21 September. Available at: https://business.lesechos.fr/entrepreneurs/idees-de-business/0611700687717-oso-ai-lance-un-boitier-coconstruit-avec-les-ehpad-pour-gerer-les-nuits-345096.php

Mertens, J. (2021) 'Google va lancer une app pour analyser les grains de beauté', 20 minutes, 24 May. Available at: https://www.20minutes.fr/high-tech/3046023-20210520-une-app-pour-analyser-les-grains-de-beaute-et-pathologies-de-la-peau

Metz, R. and Yurieff, K. (2021) 'TikTok empowered these plus-sized women, then took down some of their posts. They still don't know why', CNN, 3 March. Available at: https://www.cnn.com/2021/03/03/tech/tiktok-content-moderation/index.html

Molga, P. (2020) 'Créée par le roboticien Gorgé pour recruter ses intérimaires, StedY s'émancipe', Les Echos, 4 December. Available at: https://www.lesechos.fr/pme-regions/ile-de-france/creee-par-le-roboticien-gorge-pour-recruter-ses-interimaires-stedy-semancipe-1271120

Molga, P. (2021) 'Gojob fait entrer l'IA dans le recrutement intérimaire', Les Echos, 6 December. Available at: https://www.lesechos.fr/pme-regions/innovateurs/gojob-fait-entrer-lia-dans-le-recrutement-interimaire-1369910

Murgia, M. (2021) 'Google launches AI health tool for skin conditions', The Financial Times, 18 May. Available at: https://www.ft.com/content/6d4cd446-2243-43f4-befd-565b4e880811.

Ostermann, O. (2021) 'Des tétons et des femmes allaitantes affichés dans Paris pour contrer une censure d'Instagram', Le Figaro Madame, 15 October. Available at: https://madame.lefigaro.fr/societe/affiches-de-tetons-femmes-en-train-allaiter-collees-dans-paris-pour-contrer-une-censure-instagram-141021-198716.

Peitsmeier, H. (2022) 'Wenn der Algorithmus den Blechschaden reguliert', Frankfurter Allgemeine Zeitung, 17 January. Available at: https://www.faz.net/aktuell/wirtschaft/unternehmen/allianz-software-springt-bei-blechschaeden-ein-17734502.html.

Phillips, M. (2021) 'Am I wrong to hate drones and driverless cars?', The Times, 11 May. Available at: https://www.thetimes.co.uk/article/am-i-wrong-to-hate-drones-and-driverless-cars-qg2vmrjm3

Pons, S. (2020) 'Nyome Nicholas Williams, l'influenceuse qui a fait plier Instagram sur la nudité', Le Figaro Madame, 29 October. Available at: https://madame.lefigaro.fr/style/nyome-nicholas-williams-linfluenceuse-qui-a-fait-plier-instagram-281020-183206

Pozniak, H. (2020) 'The bias battle: how software can outsmart recruitment prejudices', The Guardian, 23 December. Available at: https://www.theguardian.com/careers/2020/dec/23/the-bias-battle-how-software-can-outsmart-recruitment-prejudices

Randall, I. (2021) 'Dr Google will see you now! Looking up symptoms on the Internet IMPROVES your ability to diagnose illness without increasing anxiety, surprising study finds', The Daily Mail, 29 March. Available at: https://www.dailymail.co.uk/sciencetech/article-9414307/Health-Looking-symptoms-online-IMPROVES-ability-diagnose-without-increasing-anxiety.html

Reddig, S. (2021) 'KI guckt mit', Tagesspiegel, 19 July. Available at: https://www.tagesspiegel.de/wirtschaft/ki-guckt-mit-4264693.html.

Richter, A. (2020) Mehr nackte Haut bedeutet mehr Sichtbarkeit, Deutschlandfunk Kultur. Available at: https://www.deutschlandfunkkultur.de/sexistische-algorithmen-bei-instagram-mehr-nackte-haut-100.html.

Ritschel, C. (2022) '"Female only" app called "transphobic" and racist' after AI denies some women', The Independent, 27 January. Available at: https://www.independent.co.uk/life-style/female-only-app-giggle-transphobic-b2002235.html

Ronzheimer, M. (2021) 'Der Mensch im Mittelpunkt', Die Tageszeitung: taz, 26 August. Available at: https://taz.de/Technologiegespraeche-in-Alpbach/!5796392&s=k%C3%BCnstliche+intelligenz/.

Roy, S. (2021) 'Pourquoi l'intelligence artificielle est loin de remplacer le radiologue', Le Figaro, 13 September. Available at: https://www.lefigaro.fr/sciences/pourquoi-l-intelligence-artificielle-est-loin-de-remplacer-le-radiologue-20210913

Rutkowski, M. (2020) '„Dünne, weiße Models posten sich fast nackt. Wieso kann ich es nicht?"', Die Welt, 12 August. Available at: https://www.welt.de/iconist/partnerschaft/article213319610/Iwanttoseenyome-Instagram-loescht-Nacktfotos-von-curvynyome.html.

Schönert, U. and Spiewak, M. (2021) 'Schlauer lernen', Die Zeit, 8 September. Available at: https://www.zeit.de/2021/37/kuenstliche-intelligenz-schule-unterricht-lernen-software-brainix.

Seufert, J. (2022) 'Was darf mein Chef?', Die Zeit, 26 January. Available at: https://www.zeit.de/2022/05/arbeitsrecht-rechte-von-vorgesetzten-arbeit.

Six, N. (2021) 'Dixième anniversaire de Siri : la reconnaissance vocale a-t-elle conquis notre quotidien ?', Le Monde, 7 October. Available at: https://www.lemonde.fr/pixels/article/2021/10/07/dixieme-anniversaire-de-siri-la-reconnaissance-vocale-a-t-elle-conquis-notre-quotidien_6097477_4408996.html

Skibba, R. (2021) 'The computer will see you now: is your therapy session about to be automated?', The Guardian, 4 June. Available at: https://www.theguardian.com/us-news/2021/jun/04/therapy-session-artificial-intelligence-doctors-automated.

Smee, B. (2021) 'Queensland police to trial AI tool designed to predict and prevent domestic violence incidents', The Guardian, 13 September. Available at: https://www.theguardian.com/australia-news/2021/sep/14/queensland-police-to-trial-ai-tool-designed-to-predict-and-prevent-domestic-violence-incidents

Smith, A. (2021) 'AI systems can have "catastrophic effects" on human rights, UN says', The Independent, 16 September. Available at: https://www.independent.co.uk/life-style/gadgets-and-tech/artificial-intelligence-united-nations-face-scan-social-score-b1921311.html

Steinitz, D. (2021) 'Ist das noch Kunst oder schon Porno?', Süddeutsche Zeitung, 18 February. Available at: https://www.sueddeutsche.de/kultur/gerhard-schroeder-linkedin-helmut-newton-1.5210082.

Stern (2021) 'Googles neue KI-App soll Hautkrankheiten erkennen', 6 January. Available at: https://www.stern.de/gesundheit/-derm-assist---google-will-hautkrankheiten-per-app-erkennen-30550698.html

Ternynck, J. (2021) 'Opinion | Cinq pistes pour booster le recrutement par l'IA', Les Echos, 6 December. Available at: https://www.lesechos.fr/idees-debats/cercle/opinion-cinq-pistes-pour-booster-le-recrutement-par-lia-1369898

The Independent (2021) 'In Africa, rescuing the languages that Western tech ignores', 29 December. Available at: https://www.independent.co.uk/news/google-nigeria-south-africa-lagos-english-b1981414.html

The Sun (2021a) 'Bots nab litter louts', 1 March, p. 24.

The Sun (2021b) 'Insta diet blunder', 16 April.

Thomas, M. (2021) '«Madres Paralelas» : sur Instagram, le patriarcat en téton armé', Libération, 12 August. Available at: https://www.liberation.fr/societe/droits-des-femmes/madres-paralelas-sur-instagram-le-patriarcat-en-teton-arme-20210812_6PYJI7YGPNAENE5ZX4Y27I4QIY/

Tisseron (2021) 'Les enceintes connectées pourraient se révéler de redoutables instruments de manipulation', Le Monde, 29 December. Available at: https://www.lemonde.fr/idees/article/2021/12/29/serge-tisseron-les-machines-parlantes-pourraient-se-reveler-de-redoutables-instruments-de-manipulation_6107577_3232.html

Tonkin, S. (2021) 'Facebook is accused of discrimination over male-targeted job adverts', The Daily Mail, 9 September. Available at: https://www.dailymail.co.uk/sciencetech/article-9973037/Facebook-accused-discrimination-male-targeted-job-adverts-mechanics-pilots.html

Top Employers 2021 (2021) 'Quand l'Intelligence artificielle pilote la gestion des talents', Les Echos, 26 January. Available at: https://www.lesechos.fr/partenaires/top-employers-2021/quand-lintelligence-artificielle-pilote-la-gestion-des-talents-1284069

Touré, K.D. (2021) 'Deadbots : IA, une vie après la mort', Libération, 17 November. Available at: https://www.liberation.fr/lifestyle/deadbots-apres-la-mort-ia-lau-dela-20211117_Y6OBCSD4NBBOBJCPJ4GU4DZMLE/

Tran-Tien, C. and Zufferey, D. (2020) L'algorithme de Facebook est sexiste pour les offres d'emploi, rts.ch. Available at: https://www.rts.ch/info/sciences-tech/technologies/11684864-lalgorithme-de-facebook-est-sexiste-pour-les-offres-demploi.html

Tribune de Genève (2020) 'Instagram encourage la nudité, selon une étude', 3 October. Available at: https://www.tdg.ch/instagram-encourage-la-nudite-selon-une-etude-683956491065

Trinks, S. (2021) 'Mimikry der Brustwarze', Frankfurter Allgemeine Zeitung, 16 August. Available at: https://www.faz.net/aktuell/feuilleton/almodovar-filmplakat-auf-instagram-zensiert-17486473.html.

Tutenges, R. (2020) Jugée «trop sexy», une image d'oignons bloquée par l'algorithme de Facebook, Slate. Available at: https://www.slate.fr/story/195851/facebook-algorithme-bloque-image-oignons-jugee-trop-sexy.

Watt, L. (2021) 'Chinese scientists create "world first" AI prosecutor that will decide people's fate', The Telegraph, 26 December. Available at: https://www.telegraph.co.uk/world-news/2021/12/26/chinese-scientists-develop-world-first-ai-prosecutor-can-detect/.

Wintermayr, A. (2021) 'Hate Speech im Netz: Wenn Maschinen moderieren', Die Tageszeitung: taz, 21 July. Available at: https://taz.de/!5781835/

Witzeck, E. (2021) 'Wie im Homeoffice die Überwachung zugenommen hat', Frankfurter Allgemeine Zeitung, 14 December. Available at: https://www.faz.net/aktuell/karriere-hochschule/wie-im-homeoffice-die-ueberwachung-zugenommen-hat-17676498.html

Woitier, C. (2020) 'Instagram met à jour sa politique sur les photos montrant des seins', Le Figaro, 28 October. Available at: https://www.lefigaro.fr/secteur/high-tech/instagram-met-a-jour-sa-politique-sur-la-nudite-apres-la-campagne-d-une-mannequin-grande-taille-20201028

Wolfangel, E. (2021) 'Der Chef sieht alles', Die Zeit, 15 September. Available at: https://www.zeit.de/2021/38/ueberwachung-arbeitsplatz-software-arbeitnehmerrechte-datenschutz/komplettansicht.

Wüstholz, F. (2020) Health algorithms discriminate against Black patients, also in Switzerland, Algorithm Watch. Available at: https://algorithmwatch.ch/en/racial-health-bias/

Wüstholz, F. and Kaufmann, D. (2021) 'Die Software sagt, wer den Job bekommt', Die Wochenzeitung (WOZ), 15 April. Available at: https://www.woz.ch/2115/digitalisierung/die-software-sagt-wer-den-job-bekommt.

Zoske, S. (2021) 'Algorithmen sollen unseriöse Anlageberater aufspüren', Frankfurter Allgemeine Zeitung, 9 June. Available at: https://www.faz.net/aktuell/rhein-main/frankfurt/kuenstliche-intelligenz-soll-unserioese-anlageberater-aufspueren-17377136.html?premium.

20 minutes (2020) 'Google Traduction : Une nouvelle approche pour limiter les a priori liés au genre', 23 April. Available at: https://www.20minutes.fr/high-tech/2766611-20200423-google-traduction-nouvelle-approche-limiter-priori-lies-genre

How to Cite:

Barassi Veronica, Scharenberg Antje, Poux-Berthe Marie, Patra Rahi and Di Salvo Philip (2022), 'AI Errors and the Profiling of Humans: Mapping the Debate in European News Media', Research Report (I) The Human Error Project: AI, Human Rights and the Conflict over Algorithmic Profiling, School of Humanities and Social Sciences and MCM Institute University of St. Gallen, St. Gallen, Switzerland.