Data Management and Governance Group Report

Analysis of the Data Governance of PimEyes as an Artificial Intelligence Application

Group 8A

Christian Jordan
Daniel Cawley
Jamie Atiyah
Lauren Salkeld

TABLE OF CONTENTS

Analysis of the Data Governance of PimEyes as an Artificial Intelligence Application

Christian Jordan, Daniel Cawley, Jamie Atiyah, and Lauren Salkeld

1. INTRODUCTION

The global demand for artificial intelligence (AI) products is ever-increasing. According to Statista Research Department the market value of AI is estimated to increase to more than 107.5 billion U.S. dollars by 2028 from the estimated 15.84 billion in 2021 (Statista Research Department, 2023). The rapid rise of AI is transforming societies influencing a wide range of domains ranging from finance to employment and healthcare (Busuioc, 2022). According to Brynjolfsson and McAfee we have "entered a second machine age in which machines are not only compliments to humans as in the industrial revolution, but also substitutes as professions and work of all kinds will be affected by AI" (Coeckelbergh, 2020). AI is everywhere, it plays a part in everyday life, and it can be both helpful and pervasive. Since the implementation of machine learning is worldwide, it has widespread implications for people which is why it is paramount that the ethics of AI are widely discussed and kept at the forefront of the minds of those who create and maintain the AI. This report will explore and illustrate the data governance issues relating to the application of PimEyes, a facial recognition application where users upload an image, containing a face, and PimEyes displays the images found across the internet of the uploaded face.

1.1. AIMS & OBJECTIVES

The major aim of this report is to provide an illustration of the social, moral, cultural, environmental, legal, and ethical issues relating to the use of AI and to call attention to these issues within the context of PimEyes. By reviewing and comparing relevant evidence, this report will look to exhibit the current and emerging ethical guidelines, frameworks, principles, and legislation and assess if such ethical practices are being adhered to by PimEyes. The objective is to provide a balanced assessment of the impact PimEyes has on individuals, society, business, and other stakeholders.

2. KEY ETHICAL ISSUES AND DATA GOVERNANCE PRACTICES

To facilitate responsible, considerate design and production of Artificial Intelligence (AI) applications, the Alan Turing Institute developed The FAST (Fairness, Accountability, Sustainability, Transparency) Track Principles (Leslie, 2018). Not only do these principles detail the ethical considerations that must be addressed during AI development, but they also act as the foundation upon which all data-centric products and services should be built and elucidate the importance of executing effective data governance practices.

2.1. FAIRNESS

Fairness is an invaluable concept that must be accounted for at every stage of an application's development, and acts as a key component of Responsible AI; AI that is used effectively by organisations without compromising the rights of the individual (Bateni et al., 2022). Failure to sufficiently address fairness concerns during data collection and pre-processing, application/model design, outcome evaluation, or even application deployment could give rise to irreparable consequences for businesses and end-users alike. That said, even the most diligent attempts to maintain fairness may be undone by human bias ingrained from the outset (Information Commissioner's Office (ICO), 2019). Human bias and underrepresentation in training data is a recurring (and potentially fatal) issue in the field of medicine, where AI is being

used to diagnose and assess patient risk (Ibrahim & Pronovost, 2021). Another study, concerning online advertisement delivery in relation to "racially associated names," identified statistically significant discrimination in the types of ads and websites suggested based on searches of said names (Sweeney, 2013). With the ever-increasing implementation of AI services across an array of industries comes an obligation to reduce bias to the greatest possible extent throughout design and execution. The Turing Institute's FAST Track Principles propose the principle of discriminatory non-harm as a minimum requirement for ensuring fairness in the use of AI systems; a commitment to cause no harm to others resulting from potentially discriminatory application outputs.

2.1.1. FAIRNESS IN DATA ACQUISITION, STORAGE, AND USAGE

Given the rapid, unrelenting amplification of the amount of data available for analysis, machine learning, and AI training, its ethical obtainment and use are of paramount importance. Robust storage solutions to mitigate the risk of breaches, and strict privacy policies detailing how personal data will be used (and by whom) are vital for any business using "big" data (Teich, 2020). Furthermore, the duration of storage and steps taken to protect individuals in instances where sensitive personal data are involved (medical records, for example) must be considered (Smith et al., 2021). The FAST Track Principles discuss the requirement to evaluate the reliability and relevance of sources from which data are collated and ensure that data is not outdated, such that its distribution may impact the fairness of the AI system in development (Leslie, 2018).

2.1.2. FAIRNESS IN DESIGN, REPRESENTATION, AND SYSTEM TRAINING

Ensuring that AI is trained with data that is sufficiently representative of different ethnicities, sexes, genders, ages, social classes, etc. to minimise the potential for an unfair/biased output arising from the underrepresentation of a given population is critical. AI applications are trained using historical data; data that may be heavily influenced by pre-existing social bias (Nadeem et al., 2021), thus eliciting misrepresentation of particular communities, and resultantly leading the system to reach inaccurate and unreliable conclusions when exposed to previously unseen data. Some systems, in which the application is 'improved' in real-time by user input, open the door to training data manipulation, potentially yielding less than favourable results (Vincent, 2016).

Aside from obtaining appropriate training data, feature selection is an aspect of the design process in which human bias has the propensity to subsequently generate AI bias. With the aim of the project unchanged, the use of different feature combinations to fulfil this aim may return drastically different outcomes (Roselli et al., 2019). Fairness depends heavily on the explainability of both the design of a system and the outputs it produces/decisions it makes (Zhou et al., 2022). Unfortunately, the black-box nature of many AI systems makes it extremely difficult to say, with certainty, that the application operates entirely without unintentional bias, and it is often impossible for those who were not directly involved in development to assess potential sources of unfairness (Roselli et al., 2019). For those involved in the development, extensive domain knowledge is indispensable to effectively analyse system outputs and avoid unjustifiable placement of trust in the application.

2.1.3. DATA GOVERNANCE PRACTICES FOR ADDRESSING FAIRNESS

The FAST Track Principles outline several key governance practices that must be implemented into the design and considered throughout development. Firstly, an overarching "Fairness Aware Design" must be incorporated from the project's inception in the interest of addressing discriminatory non-harm. This ensures that all team members are aware of the importance of fairness and are consequently mindful of how elements of their work may introduce bias. With regards to data acquisition, storage, and usage, substantial documentation (in the form of a dataset factsheet) of data sourcing and storage, relevance to the domain, representation of varying communities, and any unexpected shortcomings of the data and system, is imperative. Finally, upon completion of the project, The Turing Institute recommends the preparation of a Fairness Position Statement (FPS) which details the steps taken to ensure fairness in the system, in a fashion that can easily be interpreted by the layperson.

2.2. ACCOUNTABILITY

The principle of accountability in relation to AI systems can often be a complex issue. One school of thought dictates that humans are not to be held accountable for AI outputs/decisions, as they are not in complete control of the applications they design (Matthias, 2004), while others argue that developers must be transparent in their work and responsible for the outcomes they produce (Fox, 2007). Additionally, the aforementioned argument explains that transparency alone does not adequately address accountability, but to some extent, the two do go hand-in-hand.

In a normal working/day-to-day environment, the responsibility of a decision, whether right or wrong, can usually be attributed to one of a small number of individuals. However, in the context of AI, system complexity, and the vast number of parties involved at all stages of the design, development, and deployment process makes assigning blame (or, conversely, plaudits) extremely difficult (Toth et al., 2022).

The Turing Institute's principles explain that, generally, considerations of accountability should pertain to one of two components; answerability and auditability. Respectively, these components concern the idea that the justification of AI-supplemented decision-making is the responsibility of the humans associated with development, and that every element of design and development is well-documented such that they can be reviewed where necessary.

2.2.1. DATA GOVERNANCE PRACTICES FOR ADDRESSING ACCOUNTABILITY

Much of the academic discussion surrounding accountability in data-driven systems concerns the placement of responsibility for the justification of AI output, and the extent to which said responsibility lies with humans. However, acknowledgment of accountability should take place before any system development commences. Completion of a Data Protection Impact Assessment (DPIA) is a legal requirement under the General Data Protection Regulation (GDPR) when the processing and use of data holds the potential to pose a significant risk to human rights (Pandit, 2022). According to the ICO, given the fact that the United Kingdom (UK) GDPR does not explicitly define high-risk scenarios, a DPIA is necessary to identify potential misgivings in data protection and address them accordingly. The DPIA is similar to the dataset factsheet discussed in section 2.1.3 and is essential in the interest of achieving auditability (section 3.0).

Further to DPIAs, the ICO states that the UK GDPR implores public authorities or organisations that 'systematically monitor' individuals, to assign the responsibility of ensuring data protection compliance to an independent Data Protection Officer (DPO). The DPO can be a new, existing, or contracted employee who spearheads endeavours towards ensuring data protection laws are adhered to and will often advise on DPIA conduction.

An additional practice, through which industries or sectors can further evidence their mindfulness of accountability, is the voluntary adoption of a code of conduct (ICO, 2022). These are submitted by trade associations, approved (or disapproved) by the ICO, and intended to aid organisations in following legislation contained within the UK GDPR. Although not compulsory, developing and implementing these codes of conduct acts as a valuable indication of a sector's understanding of the importance of accountability.

2.3. SUSTAINABILITY, AN OVERVIEW

Sustainability in AI refers to creating and continuously maintaining systems, with the understanding that these systems have socio-economic and environmental impacts on the world today, and possibly hundreds of years from now. These long-lasting impacts can affect people and animals in a way that is permanent, which is both exciting and scary for humanity in equal parts. Hence the necessity for ethical guidelines so that society can benefit from its transformative effects.

2.3.1. SUSTAINABILITY OF AI AT THE MICRO AND MACRO LEVEL

AI is a powerful tool, its effects can be felt on the micro levels (individual societies), and the macro level (the world). On a macro level, it is important for countries to communicate effectively about the aims and implications of AI whilst simultaneously being sensitive to multicultural differences. It's important to be aware that 'misunderstandings

between cultures and regions play a more important role in undermining cross-cultural trust' (ÓhÉigeartaigh, et al., 2020: 571) and therefore the sustainability of AI worldwide. Everyone joins the race to produce the most innovative AI as data becomes 'more valuable than gold' (Chanakira, 2022: online).

At the micro level, it impacts everything from jobs and culture to the environment. It has the ability to empower minority groups or vulnerable groups, but it can also have undesirable and disastrous consequences. For example, a report by World Economic Forum predicts that over 80 million jobs worldwide will be replaced by machines globally by 2025 (Verma, 2022), which will mainly impact blue-collar workers. It is evident that the sustainability of AI is hindered by how many people are afraid of the real-life ramifications it can cause. Conversely, some people believe AI has provided positive contributions so far, such as enabling us to 'store and analyze data in multiple industries effectively, to improving our regular routines with virtual and home assistants' (Defined.AI, 2021: online), allowing businesses to progress and strengthen the economy. So, wider acceptance 'might bring economic advantages to those societies who from early on embrace AI technologies' (Sindermann, C. et al.,2022: online) which is thought to play a key role in the success of AI long term.

2.3.2. SUSTAINABILITY AND DEVELOPING THE MODELS

AI sustainability is challenged by society's acceptance, cyber security risks, and the influence of historical data in machine learning models. For example, 'One of the common errors we see is in skewed samples… you see more officers being dispatched to certain neighborhoods and you end up building on and perpetuating historical data' (Dhinakaran, 2022). It is important to remain sensitive to the fact that old data is more likely to have more bias. Historical data is often used to train AI models because training sets require millions of rows of data to validate them. Although it may not be entirely possible to avoid using historical data, it is possible for companies to stay educated in current policies and be able to detect inaccuracies. Factors of bias may include occupation, education, income, wealth and location. So, a human is in the best position to quality check the output for bias if AI is to be successful. This is one of the reasons 'the need for human-powered data labelling only continues to grow: You still need data to fine-tune and validate automatically generated solutions' (Megorskaya, 2022: online). So, humans are likely to play a role in the sustainability of AI in the future, showing that AI may not threaten blue-collar jobs as predicted, but change the way the job market operates.

2.3.3. WHERE IS SUSTAINABILITY NOW?

As AI is advancing and impacting society, it is important to look at where ethics and sustainability of AI currently stand. Most guidelines aren't explicit about the timescales of sustainability or what the boundaries are for ethics, so it is imperative to value everything equally regardless of time or date. Achieving this requires non-anthropocentric (non-human-centric) ethics (Owe, 2021). This is because AI will continue to affect animals, humans, and the environment. So, if humans use AI to speed up tasks that take from the environment without thinking of the consequences, then human existence may become threatened. To quote Sir Stephen Hawkings: 'Artificial Intelligence is either the best or the worst thing to happen to humanity' (Hawking, unknown, cited in Hern, 2016: online). For AI to be sustainable for humans, humans also need to sustain AI and their natural environment.

2.4. TRANSPARENCY, AN OVERVIEW

Transparency is important when creating an artificial intelligence (AI) application, whether it is being used in an institution or being deployed for public use. From conception through to the finished product, it is essential to build an application with the understanding that it will be open to scrutiny, from a quality control perspective and to ensure it is non-discriminatory. Creators and stakeholders must ensure good explainability and be able to justify why a model was created and how it works.

However, some organisations are concerned that 'transparency brings the risk of disclosing intellectual property that they want to protect' (Dhinakaran,2021: online). This raises questions, such as "How much information is too much information?" and, "What is the trade-off between AI progression and explainability?".

A balance between how much detail organisations share about methodologies may be required as trade secrets can give them a competitive edge which enables AI to grow and improve. However, providing no information can seem untrustworthy and may hinder the advancement of AI as society's acceptance is key to its success. Even if organisations were to publish their methodologies in full, it would require an expert to translate them. These experts cannot always see how an algorithm came to make a decision because sometimes models create links between variables that humans cannot see or understand. Even the 'most capable technologies—namely, deep neural networks—are notoriously opaque' (Bleicher, 2017:online). These are commonly referred to as black box methods which work faster with a larger amount of data. Some people believe that when a black-box method is used, like deep learning, it 'improves the performance at the cost of explainability' (Guo, 2020: 39). Since AI applications are often created in fast-paced, high-pressure working environments, it is easy to see how black-box methods may be chosen over white-box methods as efficiency can often be traded off for full transparency and explainability.

There are implications of both black box and white box methods being used. If a serious error occurs in an application where a black box method is used, then the traceability of the fault and its cause may prove more difficult. This could be a serious safety issue, especially if it is within the public domain, hence the requirement for more transparent AI. There are methods that solve the black box problem, either referred to as white box or glass box methods. These methods include decision trees where the logic of each decision made is clear. The limitation is that 'such models aren't always helpful [in] achieving complex tasks or objectives' (Shin, 2021: online). This is why it is important for businesses to consider which method to use for any given task as it is important to consider public safety in the progression of AI.

Some people worry the bar set for automated decision-making is set too high due to 'an unrealistically high estimate of the degree of transparency' (Zerilli, J. et al., 2018: online). Understandably, those in the industry want to make a huge impact with AI and want to support its fast development, preferring to use the fail-fast method. However, boundaries and restrictions around AI are necessary to safeguard those who are vulnerable in society, even if it means slower technological progress for AI because the largest societal progress is made when AI works alongside empathy. Since AI applications often mirror the creator's, users', or society's own bias, the future of ethical problems associated with AI are likely to continue as "algorithms are becoming a ubiquitous part of human lives' (Shin, 2021: 661). and therefore are likely to encounter and reflect all the socio-economic problems existing in the world today. But, as society becomes more aware AI and its applications, often through the media, society can collectively work with AI to mitigate these issues.

Transparency isn't only applicable to a machine learning (ML) algorithm or business-to-business exchanges, but also the group of people within an organisation. A company that operates using an open-loop system (open to failure) 'does lead to progress because the feedback is rationally acted upon' (Syed, 2015). Compared to a company that uses closed-loop systems (hiding failure). Company culture impacts honesty which directly impacts transparency, these factors influence decisions made in every phase of an AI application up until deployment. So, the ethics of an AI application created by a company is reflective of the culture of that company. And, the individuals at each stage of the creation of an application all need to be accountable for transparency.

3. CURRENT AND EMERGING ETHICAL GUIDELINES, FRAMEWORKS, PRINCIPLES, AND LEGISLATION

This section will delve into the past, present, and future ethical guidelines, frameworks, principles, and legislation. Whilst much attention is accredited to the General Data Protection Regulation (GDPR), there is a wide selection of guidelines, frameworks, principles, and legislation that can be discussed. A short timeline will be presented to illustrate the evolution of legislation and frameworks from the earliest, simplistic instances through to the ever more complex, emerging cases.

3.1. EXISTING GUIDELINES, FRAMEWORKS, PRINCIPLES, AND LEGISLATION

3.1.1. EUROPEAN CONVENTION ON HUMAN RIGHTS

One of the earliest legislations that outlines data principles is Article 8 within the European Convention on Human Rights (ECHR) in 1953. Article 8 states that "everyone has the right to respect for his private and family life, his home and his correspondence" (European Court of Human Rights, 1950). Calder (Calder, 2018) critiques the usefulness of the Article 8 as a legal protection for personal information, declaring that it is a "unnecessarily open-ended provision". Certainly Article 8's use of "respect" does impart a broadness to the article, opening the possibility for suggestion and leeway.

3.1.2. EUROPEAN DATA PROTECTION DIRECTIVE

After the ECHR was published, various data protection laws then emerged, leading to the establishment of many modern-day data protections principles. Alhadeff et. al (Alhadeff, 2012) describe the Organization for Economic Cooperation and Development Guidelines within Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, published in 1980, as the "formal debut" of accountability in the international data protection space. The UN published Guidelines concerning Computerized Personal Data Files on the 14th of December 1990. Kuner (Kuner, 2009) describes the guidelines as "influential" but highlights that the "high-level" data protection guidelines are not "legally binding".

The European Data Protection Directive (DPD) was then introduced in 1995. A review commissioned by the Information Commissioner's Office (ICO) (Robinson, et al., 2009) states that that the DPD "harmonises data protection principles" and lead to an "improved awareness of data protection concerns". One of the main takeaways from the DPD is that it addresses data processing when third countries are involved. Article 25 states: "The member states shall provide that the transfer to a third country of personal data which are undergoing processing … may only take place only if … the third country in question ensures an adequate level of protection" (The European Parliament and the Council of the European Union, 1995).

Article 10 of the DPD attends to the transparency of data collection. The DPD declares that data controllers must give any data subject the "identity of the controller and of his representative, if any" and "the purposes of the processing or which the data are intended" (The European Parliament and the Council of the European Union, 1995). However, Robinson, et. al (Robinson, et al., 2009) highlight how the role of Data Protection Authorities in "accountability and enforcement" are "inconsistent". They conclude the directive as a whole "will not suffice in the long term".

3.1.3. GENERAL DATA PROTECTION REGULATION

Wilson & Jahankhani (Wilson & Jahankhani, 2021) give the rationale for the conception of the GDPR as the "expected explosion of data" in the first half of the 21st century. The GDPR came into effect in May 2018 as a follow up law to the EU projection directive. The GDPR warns against any violations of its privacy and security standards with fines that "max out at €20 million or 4% of global revenue (whichever is higher)" (Wolford, 2022).

Crockett, et al. (Crockett, et al., 2018) highlight Article 22 as "one of the major changes" brought upon by the publication of the GDPR. Article 22 refers to provisions relating to Automated individual decision-making, including profiling. The GDPR states: "the data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her" (GDPR.EU, 2018). Wachter (Wachter, 2018) evaluates Article 22 and voices concern of how the "scope of applicability" is likely to be "very limited" due to language term choices, namely the use of terms "solely automated", "legal or similarly significant effects" and how they "remain undefined in practice".

Significant fines imposed over violations of GDPR standards have been imposed. In July 2021, it was reported that Amazon was fined €746 million with the "tech giant's processing of personal data did not comply with EU law" (BBC News, 2021). In November 2022 it was reported that Meta was fined €265 million by the Irish Data Protection

Commission because of a “data breach that saw the personal details of hundreds of millions of Facebook users published online” (BBC News, 2022). In the Meta data breach case, the article in question was a breach of article 25: Data protection by design and by default.

3.1.4. THE INTERNATIONAL COMMISSIONER’S OFFICE (ICO)

Responsible for many high-profile data usage legal cases resulting in hefty fines for companies that misuse data in the UK, the ICO is perhaps the greatest of these was the £7.5m fine for Clearview AI (International Commissioner’s Office, May 2022), who were found guilty of failing to use collected facial recognition data in a way that is fair and transparent. Essentially, people who’s data was being used were not aware that it was being used in this way.

The ICO is therefore the current regulator for data usage in the UK and serves to uphold public data rights as set out the Data Protection Act, Freedom of Information Act, General Data Protection Regulation and others. It addresses ethical issues surrounding data protection for the most part, and introduces it’s own guidelines (International Commissioner’s Office, n.d), of which data protection impact assessments (DPIAs) are introduced.

3.2. EMERGING GUIDELINES, FRAMEWORKS, PRINCIPLES, AND LEGISLATION

With artificial intelligence growth accelerating at unpredictable speeds throughout the 21st century thus far, governments and commissions in globally have been racing to keep up with the expanding number of ethical issues that come with this.

The main aims of these guidelines, frameworks, principles and legislations are to ensure that any artificial intelligence application cannot be deemed harmful in any way, but also to allow great economic growth in the AI sector of a developing country.

3.2.1. FOUNDATIONS FOR THE FUTURE OF AI GOVERNANCE

The basis for the future of data and AI governance is largely based on the research and proposal from UNESCO. The first globally accepted strategy for introducing regulations and legislation, "Recommendations for Artificial Intelligence", (UNESCO, 2022) was a huge step forward for data governance principles.

Encourages regulation on applications based on its core principles.

Proportionality and Do No Harm	Safety and Security
Fairness and Non-Discrimination	Sustainability
Right to Privacy / Data Protection	Human Oversight and Determination
Transparency and Explainability	Responsibility
Responsibility and Accountability	Awareness and Literacy

Table 1: UNSECO ethical principles.

It is important to remember that a lot of the proposals and recommendations made are simply just that, and official regulations and legislation have not yet been introduced in many countries, leaving many grey areas in data governance.

3.2.2. EUROPEAN AI STRATEGY

"Proposal for Regulatory Framework" set out by the European Commission (April 2021) identifies four separate risk categories for each individual application (Unacceptable, High, Limited, Minimal or No Risk). Applications with

unacceptable risk will be banned from the market, while high risk applications are subject to strict obligations before they can be deployed. Limited risk means certain transparency requirements must be met and minimal risk will not be regulated.

Regulations are set to be released late 2022 to early 2023 in a transitional phase and also provides a guide for developers on the steps it will need to take when this comes into force. The earliest that this could be regulated fully is 2024. Applications must undergo conformity assessments and comply with AI requirements before being registered to a national database. It also states that if significant changes are made it must start the process again.

The European Commission also found that AI liability was proving an issue under current legislation, and addressed this with a "Proposed Directive for AI Liability" (European Commission, Sep 2022). Victims of damage caused by AI applications need to prove a wrongful action or omission by the person who caused the damage, but legislation does not currently assign liability to developers. Changes to these laws are designed to increase public trust in such applications, and if developers know they are liable for their applications, they are likely to ensure they are fair and ethically correct.

3.2.3. UNITED KINGDOM AI STRATEGY

Building on the UNESCO (2022) recommendations and the strategy of the European Commission, the UK government have made their intentions clear with a number of proposals.

HM Government (September 2021) published "National AI Strategy" outlining a greater need for data further restrictions on AI and aims to address the governance issue carefully as not to restrict the growth in the sector. Moving to bring attention to the governance issues faced rather than solving them, and state the governments intentions to collude internationally on building transparency in applications. Interestingly, one of the main points made is that currently, deployers of such applications do not have the guidelines or legislation to follow.

The UK's government followed up on its strategy goals with "Proposals on the Future Regulation of AI" (Collins, D. 2022) which lets developers know what responsibilities they will have when the legislations are confirmed. Contradicting the EU strategy for AI, the UK government “will allow different regulators to take a tailored approach to the use of AI in a range of settings”. A target is also set of 12 months since publication for these new regulations to start being put into place. A key takeaway is that a legal person must be responsible for the application, substantiating the notion that the developers are responsible for the actions and "learning" of the applications.

3.3. CRITICISMS OF CURRENT AND EMERGING GUIDELINES

Are regulations and legislation heading in the right direction? It is clear these proposals for change were welcomed by global communities and government officials.

Daniel Schiff et al. (March 2021) found when encoding 112 published AI ethics documents looking for engagment in 25 different ethics topics, an increased number of these topics were considered in publications from public and non-governmental-organisation sectors as apposed to the private sector. Therefore, when considering new proposals and legislation do we need collaboration not only international but also within our own foundations. This problem could become more apparent with the proposals of the UK government that are open to allowing private regulators determine the legislation amongst the private sector.

Further analysis of ethical guidelines in financial services by J. Huang et al. (March 2021) concluded that these guidelines might have to be industry specific, as proposals of guidelines can be vague and do not assign responsibility for complex interactions between AI-enabled systems. Providing this would happen, the complexity of regulation in the UK alone could increase exponentially with AI constantly evolving and new guidelines required.

4. ANALYSIS OF DATA GOVERNANCE AND ETHICS OF PIMEYES AS AN AI APPLICATION

The following sections on social, moral, cultural, environmental and legal issues consider some of the key questions raised in Harms Modelling (Cassidy, D. *et al.*, 2022). Such questions relate to: “What context will the technology

likely be used in?", "Who will be impacted by the technology?" And, "How can harm be prevented?" Risk, huuman rights and democratic structures are duscussed in relation to these questions.

4.1 AN INTRODUCTION TO PIMEYES

AI applications that utilise facial recognition techniques are an ever-present, ever-controversial, journalistic goldmine (Kaye, 2022) (Valentino, 2022) (Smith, 2018). PimEyes, a site where a user uploads any photo and, almost instantly, is presented with other photos from across the internet containing the uploaded face (Harwell 2021), is one such application that has recently become shrouded in this controversy, and not for the first time. The site, launched in 2017, was the subject of a complaint made by privacy campaigners Big Brother Watch to the UK data and privacy watchdog, who were concerned by its potential facilitation of internet stalking (Vallance, 2022). The free-to-use service scours over 900 million images (Laufer and Meineck, 2020) on the internet and, unlike less sophisticated reverse image search engines of a similar ilk, returns images where the searched face is not necessarily the primary subject of the photo. The search results previously included photographs found on social media sites, however this feature was removed following significant backlash (Landymore, 2022).

According to PimEyes' owner, Giorgi Gobronidze, and its FAQ answers, the service is intended to be a self-search tool designed for users to upload their own pictures, and identify other pictures found on the internet that they perhaps did not know existed. This, supposedly, allows users to better protect themselves against nefarious activities such as catfishing or identity theft.

In theory, this proposed application of the software appears harmless, but the reality of its potential implications is deeply concerning. Although the website states that "PimEyes is designed to find photos of the person who is conducting the search, not of other people" (PimEyes, n.d.), there are currently no measures in place to prevent anyone from using a picture of anyone else. What's more, while the free use of the application is limited to only a handful of searches per day, paid subscriptions provide significantly greater search capabilities, the most expensive of which (the advanced plan, priced at £305.28/month) grants unlimited searches and 500 "PimEyes Alerts". These notify the user when a new photo that matches a previous search is uploaded to the internet. One may wonder why a 'self-search' application includes functionality that actively tracks 500 unique searches.

PimEyes has been likened to ClearView AI (Laufer and Meineck, 2020), a piece of facial recognition software hosting a database of over 20 billion images with customers including government organisations and law enforcement agencies, and which has been the subject of its own fair share of scrutiny and controversy (ICO, 2022) (Meaker, 2022) (Heikklä, 2022). The frightening difference between PimEyes and ClearView is that the former is accessible by anyone, and investigations into the platform, such as one conducted by The Intercept (Hvistendahl, 2022), concluded that it could very easily be applied to abhorrent practices; the stalking and tracking of children as just one example. The PimEyes website may state that you can only "find links to websites that have published pictures of you and URLs to source images", and that "you won't find personal data in our results" (PimEyes, n.d.), but from these pieces, it is unlikely to take bad actors long to finish the puzzle. Thus, like many other facial recognition platforms, a litany of ethical questions present themselves when assessing PimEyes and its integrity.

4.1. SOCIAL ISSUES

When considering the effect that PimEyes and other similar facial recognition applications can have, it is easy to exclusively address the negative, harmful purposes for which this technology could be harnessed. While these are in abundance, it is important to remember that, if used correctly, PimEyes could provide significant value to society.

In 2021, 1.7 million identity theft cases were reported to the Federal Trade Commission in the United States. In the same year, $5.8 billion was lost by Americans because of stolen identities (National Council on Identity Theft Protection, 2022). Image-based sexual abuse, "revenge porn", and non-consensual pornography are all terms that refer to a frightfully common practice in the digital age; uploading private, sensitive material of a sexual nature to the internet without the consent of the individual featured in the photo/video (O'Connell, 2020). Similarly, the emergence of synthetic media and deepfake technology has given rise to "deepfake pornography"; the imposition of a person's face onto pornographic photos or videos (Karasavva & Noorbhai, 2020). These are just a few of many examples of instances where a picture of a person can be utilised by others in an attempt to destroy their life. PimEyes may offer an effective solution to combat

this, by allowing users to quickly identify appearances of themselves across the internet, as well as the provenance of these appearances. Furthermore, for users who subscribe to PimEyes PROtect, PimEyes will draft and send up to 80 monthly Digital Millenium Copyright Act (DMCA) and GDPR takedown notices to companies whose website displays an image of the user without their consent. This functionality can offer solace to those who are fearful of online exploitation that could impact them indefinitely and eases the otherwise distressing, laborious process of having non-consensual digital content removed.

That being said, the current iteration of PimEyes exhibits very little consideration of the key ethical principles that should be at the forefront of any AI/data mining application. This is particularly concerning given that the EU deems facial recognition technology to pose a high risk to basic human rights (Scipione & Lo Monaco, 2022). Yes, the site states that the service is designed to be used as a self-search tool and should not be used to track down photos of other people. But, with no requirement for corroboration that the uploaded image is that of the person uploading it, it is perfectly possible for a user to search for any photo, for any purpose. Add to this the paid-for availability of PimEyes alerts (see section 4.0), and it becomes increasingly difficult to understand how Gobronidze and PimEyes can maintain that the application is solely designed for the purpose described. It is indisputable that, in the words of the owner himself, PimEyes is "tailor-designed for stalkers" (Hvistendahl, 2022). Although searches do not return any personal details pertaining to the individual in the searched image, they can facilitate the first step in the pursuit of obtaining substantial information (possibly including the whereabouts) about the person of interest. Consider a hypothetical scenario in which one person takes a picture of a stranger without their knowledge. PimEyes then allows the photographer to search for other photos of the stranger, perhaps returning an image from their employer's website with their name attached. Although the service no longer searches social media sites, it is easy to see how one could still quickly develop an understanding of who a given person is, and how bad actors could use this information nefariously. Admittedly, this is an argument that could be made against any reverse-image search engine that offers a similar service to PimEyes, a point that was made by the company in a recent official statement in response to a BBC article discussing the complaint made by Big Brother Watch (Vallance, 2022). The difference here is that PimEyes will pick faces out of a crowd, and doesn't just return obvious portrait photos that match the searched image. This could allow interested parties to identify attendees at certain events (Laufer & Meineck, 2022). For example, far-right groups may be able to identify, and subsequently harass, individuals attending demonstrations regarding the rights of marginalised members of society.

Aside from how PimEyes may be applied to cause harm to others, information regarding the underlying technology and data storage is distinctly lacking. Details on how the application works are incredibly vague, with the website offering nothing more than "PimEyes is a facial recognition search engine… displaying a list of websites related to a query… the search is performed based on an uploaded photo" (PimEyes, n.d.). The website also states that any uploaded photos are temporarily stored for 48 hours. When interviewed by The Intercept (Hvistendahl, 2022), Giorgi Gobronidze stated that PimEyes "don't have any" stored photos. However, this same investigation by The Intercept found that, while uploaded photos may be deleted after 48 hours, data from these photos are stored for two years. It also found that images from search results are kept on a PimEyes subdomain which, apparently, owner Gobronidze was unaware of. Another investigation, conducted by reporters at Netzpolitik (Laufer & Meineck, 2022) included a 2018 post from the since-deleted PimEyes Facebook page. This post boasted a "premium database" with over 100 million faces, and stated that this database grows by 1.5 million faces each day. A blog post on the PimEyes website (the date and author of which are not available, as is the case for all other blog posts) states that this database is in fact comprised of photo indexes, or "faceprints" which reference the URL addresses where relevant photos can be found, and contains no actual images (PimEyes, n.d.). While this may be true, it is not clearly explained in an obvious, easily identifiable area of the site, but instead a blog post, alongside other posts which argue against privacy concerns and include images such as the renowned tinfoil hat, often used to discredit and point fun at so-called conspiracy theorists. Perhaps not wise if the company hopes to build public trust and, given that the screenshot of the 2018 Facebook post included in the Netzpolitik article clearly describes a "face database", and not a face index database or similar title, the growing skepticism towards PimEyes' conduct and integrity is understandable.

It's important to remember that much of the literature referenced in this section for raising concerns around the social implications of the application takes the form of news articles. The desire to generate a story, and subsequently manipulate, misquote or misrepresent information is an obvious possibility with this form of media. Regardless, investigation of the service by the current authors finds a multitude of shortcomings in PimEyes' consideration of its own potential societal impact, and consequent shortcomings in action taken to mitigate these negative implications. The

proposed use of the technology may offer a valuable tool for users to protect their own online identity, but this is overshadowed by the nonexistence of measures designed to prevent abusive deployment.

4.2. MORAL ISSUES

Facial recognition is generating a lot of fear currently as it is powerful and moving fast. The Biometrics Commissioner, Paul Wiles explains that technical developments today have moved much faster than the legislation (BBC, 2017). This makes auditing and monitoring AI applications difficult as the rules aren't clear on when to hold a company accountable. However, as facial recognition companies face more scrutiny, regulators begin to set a clearer standard for compliance by issuing fines against those that fail to meet criteria.

How does PimEyes try to deal with public concerns that have been voiced by so many media outlets such as Big Brother Watch (Vallance, 2022)? Worryingly, PimEyes have been negligent with respect to these concerns and have changed very little. The public may expect that developers of such applications have moral obligations to deal with such concerns, however, without current legislation forcing their hand, these obligations carry very little threat. The PimEyes application itself does not have any core morals embedded into it, so it has to be the responsibility of the developer who bears this burden. When PimEyes facial matching algorithm determines its results, it has no human thoughts to consider bias or fairness in these instances. Returning the images of vulnerable persons such as children, elderly or the ill is not built into the development of such AI. Moral thinking in machines is very different than that of human thinking, and we need to consider both the capabilities of the human and of the machine (Boddington, 2021). Machines at present do not have the moral aptitude of humans and we must program them to align with these values.

Other prominent ethical concerns of facial recognition are that it may get used for stalking, implicate innocent people, and worsen pre-existing discrimination. More often than not facial recognition tools inaccurately identify people of colour which has moral implications and should be treated seriously by developers. This is due to skewed data within training sets where age, race, gender, etc. are not equally weighted, which leads to bias. For example, if the data is primarily trained on white men, then programs will struggle to accurately identify BIPOC faces and women (Klosowski, 2020: online). PimEyes so far has not been transparent about the data they use in their training sets. Racial bias is especially damaging in America where these facial recognition technologies are used a lot by professional crime investigators, which has real-life complications. For example, over the last few years, three black men have been falsely accused of crimes because of the bias within facial recognition technology. 'All three cases were eventually dropped, but in Parks' case, that took almost a year, including 10 days in jail' (Johnson, 2022: online). Safeguards must be put in place to combat discrimination, otherwise, countries run the risk of victimising innocent civilians.

This further complicates matters when these tools are used by people who are not trained such as the general public. The owner of PimEyes, Giorgi Gobronidze, states that: 'facial recognition technology would be used to control people if governments and big companies had the only access to it' (Hill, 2022: online). But, what happens when these tools get into the wrong hands? In America, it has been turning people into investigators and PimEyes has been used in the pursuit of Capitol rioters, by an online crowdsourced collective of "sedition hunters" (Harwell, 2021: online). This could have major moral consequences if people act on their suspicions based on the results given by the technology as this could harm a guilty or innocent person as everyone has the right to a fair legal trial.

PimEyes can also be misused by individuals for stalking people and exploiting children. One of the issues highlighted by The Intercept was that 'PimEyes returned images of children labeled as "potentially explicit" while still providing a link to the source website' (Landymore, 2022: online) which has clear moral implications. When Giorgi Gobronidze was confronted with issues such as this, he confirmed that PimEyes was 'working to develop better safeguards for children' (Hvistendahl, 2022: online). However, nothing has been made clear about what these safeguards are or when they are coming into effect. This is a huge transparency issue for the user and the general public, who are unaware such a tool is widely available to anyone with an internet connection. Giorgi Gobronidze then went on to suggest that parents should be more responsible (Hvistendahl, 2022), which shifts the responsibility onto the user. These are the kind of risks that transpire when regulations are slower to take hold in society than AI applications. It results in individuals or organisations assuming the responsibility is not their own, and then it becomes difficult to know whom to hold accountable when something goes wrong. Jeramie Scott, director of the Surveillance Oversight Project at the Electronic Privacy Information Center highlighted this as a safeguarding issue as it is a human right to have some protection in society. He explained

that participating in public, whether online or offline, should not mean subjecting yourself to privacy-invasive services like PimEyes. Additionally, congress needs to act to not only protect children, but everyone from the dangers of facial recognition technology (Landymore, 2022: online). Part of the danger of AI is that the public assumes it is clever, capable, and tested. However, if it doesn't include automatic and regular safeguarding then it's not intelligent and ethical AI. So, if facial recognition technology, like PimEyes, is to be used for the power of good, it needs to be managed because it has the ability to achieve many positive things for society. For example, PimEyes told the BBC that it is working with the police to 'combat human trafficking, crimes against children and terrorism' (Nash, 2022: online). So, it can protect people in masses. Unfortunately, with not enough regulations in place, PimEyes could end up hurting the same people it is claiming it wants to help.

Calling to mind the ethical principles of fairness, accountability, sustainability, and transparency (Leslie, 2019), who is responsible for ensuring that the rights of others are not infringed upon through inappropriate use of the service? According to another of PimEyes' mysterious blog posts, it is the sole responsibility of the user. "Everyone can buy a hammer and everyone can either craft with this tool, or kill" (PimEyes, n.d.) writes the anonymous author. In other words, PimEyes is not to be held accountable for the uses of the service it provides. According to these blog posts, accounts of users that exhibit suspicious activity are monitored by the security team and suspended, sometimes permanently. However, what constitutes suspicious activity is not explained, and operators of suspended accounts may simply create new ones using different details, owing to the absence of any identity verification. PimEyes' justification for not requesting said verification is that it does not wish to store personal or biometric data. Understandable, but why a company that supposedly deletes all uploaded photos could not do the same with personal data used as verification is unclear. At present, PimEyes does not sufficiently adhere to any of the key principles discussed in section 2 of the current writing, holding serious potential ramifications for society. In terms of fairness and sustainability, placing complete responsibility on the user to operate ethically and not infringe on the rights or safety of others is inadequate and, frankly, unrealistic. It could be argued that the website does offer an explanation as to how the service functions. However, these explanations come in the form of blog posts written anonymously in response to criticisms of the application. In the interest of transparency, the website should include a thorough explanation of how it retrieves the information it provides in search results, presented in an obvious, dedicated area of the site; a "How It Works" type page accessible from the homepage.

4.3. CULTURAL ISSUES

The cultural implications of PimEyes relate to bias in the interpretation of the results, bias in the way the technology recognises different cultural groups, and different rates of acceptance amongst different cultures and cultural groups. Calling back to the discussion within section 5.0 of photo indexes, the variables in said photo indexes can be established. PimEyes states that images are not labelled based on "race, gender, age, or any other feature necessary to identify an individual" (PimEyes, n.d.). Whilst this can be applauded as a satisfactory example of ethical practice, an analysis of the substitute causes alarm. PimEyes stores an index of "facial-feature measurements" (Metz, 2021). Whilst little is known about the exact measurements used, PimEyes have confirmed that their facial recognition technology (FRT) "works similarly to other such systems" (Metz, 2021).

A look at a similar FRT engine, Clearview AI may indicate that there may be cultural implications ensuing from the use of PimEyes. Clearview AI was deemed by several US Senators to be a threat to "Black communities, other communities of colour and immigrant communities" (Brewster, 2022). This stemmed from the misclassification, and subsequently wrongful arrests, of individuals within these communities, when FRT was used by the police. In 2020, Robert Julian-Borchak Williams was wrongfully arrested because of the misuse of FRT by the Detroit Police Department (Hill, 2022). Williams' case was symptomatic of a wider problem of bias. Clearview AI strenuously denies any suggestion of demographic bias, citing the National Institute of Standards and Technology's (NIST) study which evaluated Clearview AI's FRT as being 99% accurate for all demographics (Zurah, 2022). Whilst there seems to be no algorithmic bias present, we can theorise that the software is bias by virtue of its interpretation. In the Williams case study, the investigative work in detaining Williams may have been affected by human bias (Yan, 2021). As discussed in section 5.0, PimEyes passes any responsibility for correct interpretation to its users. However, it can be argued that PimEyes should take responsibility for ensuring that its software is implemented and interpreted correctly to minimise any bias and resulting discrimination. It can also be suggested that the use of FRT's such as Clearview AI and PimEyes is inherently biased from the database used. If the databases are comprised mainly of individuals with mugshots, then the

database may hold a disproportionate number of Black individuals due to the varying arrest rates for Black and White individuals (Bass, 2022).

PimEyes may also be susceptible popular face-based gender classifiers (from IBM, China's Megvii, and Microsoft) are shown to be heavily trained on lighter-skinned faces (Marks, 2021). Therefore, because of the training data being skewed towards lighter-skinned faces, the misclassification rates of the classifiers were high for persons not heavily featured in the training data. Marks alludes to this by illustrating the misclassification rate for gender classifiers for darker skinned females is 34% where lighter skinned males had a rate of 0.8%, highlighting that FRT has gender bias alongside its racial bias. Minimalizing the misclassification rate in the above classifiers could be achieved by reducing the data size as large sized data sets can "magnify the bias associated with error resulting from sampling" (Kaplan, et al., 2014) . PimEyes states that the company has the biggest database in Europe (PimEyes, n.d.). As PimEyes "works similarly" (Metz, 2021) to the gender classifiers studied by Marks, we can suggest that the training data is also skewed towards lighter-skinned male faces and that reduction of the dataset may be required to reduce the magnification of bias.

PimEyes defines good features for photos that are being uploaded into the FRT engine (PimEyes, n.d.). Amongst these features are the recommendation that photos are "well-lit" and "full-colour", and "high quality". These recommendations may be in place to mitigate the fact that cameras are "not optimized to capture darker skin tones" (Najibi, 2020). PimEyes also recommends that there are "no elements covering the face" (PimEyes, n.d.). This guidance may have been established due to FRTs facing difficulties in detecting faces with beards, moustaches, hijabs, and sunglasses (Alam, et al., 2021) . Alam, et al. recommend that these "complexities" can be overcome with a model which operates using a combination of principal components analysis and segmentation process and removes such complexities to achieve an increase in accuracy. PimEyes could look at implementing such a model to make their FRT more accessible and less discriminatory.

Perceptions of the use of PimEyes will vary across different countries and cultures. A review of the national perceptions of FRT revealed that 67% of Chinese citizens highly support the use of FRT whilst only 38% of German citizens share this view (Kostka, et al., 2021). The differing acceptance is even more stark when focused on the private use of the technology with acceptance rising to 71% and falling to 33% for China and Germany respectively. These differences in acceptability are likely to be replicated across other countries and will depend on cultural values relating to acceptance of the state, view on human rights, and individual freedoms. Within countries, different cultural groups also perceive FRT differently. The use of FRT by law enforcement is deemed as acceptable by a smaller amount of black and Hispanic American adults than white American adults (Smith, 2019) with 47% of black persons, 55% of Hispanics and 64% of whites. It is possible that similar differences will exist in other countries, with more marginalised cultural groups viewing FRT less favourably.

In conclusion, although there is little published evidence relating to cultural implications for PimEyes technology, evidence from other FRT technologies suggest there are important cultural issues to consider in the use of PimEyes as well. Further research is clearly needed in areas such as the interpretation of PimEyes results, bias in the way the technology recognises different cultural groups, and different rates of acceptance between different countries and cultural groups.

4.4. ENVIRONMENTAL ISSUES

In a time where major climate change discussions are being had, such as the "historic" (Carlin, 2022) COP27, companies are now being called to act and reduce their carbon footprint and emissions. The European Commission revealed that the ICT sector accounts for "approximately 7% of global electricity consumption" and predicts that it will rise to "13% by 2030" (European Commission - European Commission, 2022).

The Direct, first-order effect of PimEyes on the environment can be assessed. A major aspect of PimEyes environmental impact is attributed to the database it uses. Database management systems have been judged as one of the major energy consumers (Roukh, et al., 2017). PimEyes states that the company holds the biggest database in Europe (PimEyes, n.d.). This large database will require servers to store and manage the database. Whilst little information is known regarding the specifics of the algorithm used by PimEyes, it is known to have analysed over 900 million faces (Metz, 2021). Therefore, the computational time and processing power required to run the algorithm on the large amount of data will be high and will require severs to manage the computational load. Heat will be generated by such servers and

will require cooling systems to maintain proper temperature and humidity conditions (Songa, et al., 2015). The energy use of these systems is high, accounting for 40% of the total energy consumption in data centres (Songa, et al., 2015) . The psychical components that make up much of IT equipment contain toxic materials and heavy metals such as cadmium, lead, and mercury (Naim, 2021). These heavy metals are polluting water, soils, and the atmosphere when disposed of (Briffaa, et al., 2020).

4.5. LEGAL ISSUES

Legal issues by their nature suggest a perpetrator and a victim, therefore it is important to be able to distinguish between what is lawful and what is not. In terms of AI, this can be difficult with grey areas in current legislation. As mentioned in Section 3, there is a global need for new legislation to tackle this exact issue. If an AI system only learns from historical data given to it, can the developer be held responsible for its actions or decisions?

PimEyes justifies its usage on its FAQ page for the sole purpose of searching for your own personal biometric matches. Contradicting this fair usage policy is the membership area, which allows for up to unlimited searches and 500 face alerts. This opens a host of questions to be answered surrounding the ethics and morals of the developers. Historically, aims of governmental legislation globally has been to solve data protection and privacy related issues, with legislation such as GDPR (Section 3.1.3). The legal concerns raised by Big Brother Watch relate to the privacy of the public for which their biometrical data has been mined and claim that these can be used for unlawful purposes such as stalking.

Proposals for the future of AI governance from the European Commission (2022) and HM Government UK (2021) seek to address current legislation on accountability. They outline the desire for developers to be held accountable for their applications even though they might not be entirely sure how it comes to certain conclusions. The ICO set a clear precedent when sanctioning a fine of 7.5m to Clearview AI in May 2022, where in a similar fashion, the way in which users' data was used was found to be unlawful. Clearview unlawfully obtained and used the personal biometric data of UK residents. They were found to have breached data protection laws by: failing to use the data in a way that is fair and transparent; failing to have lawful reason for collecting people's information; failing to have a process to stop the data being collected indefinitely and asking for additional personal information such as photographs when asked by public if there data was being held (International Commissioner's Office, May 2022). PimEyes have also been collecting biometric data of the public without fairness and transparency, that is, the users in the database are not aware their photographs are being used.

The argument put forward by PimEyes is that they are not providing anything to users in addition to which is already available on the public domain, like how a search engine works. They also stress that images are retained in the database only as indexes, string like objects, of which the original photograph cannot be reconstructed. GDPR may find this unlawful since public data is being kept and used without their knowledge, however PimEyes also state that no personal information such as names and addresses are kept in the database and cannot be subject to GDPR law. There is also the option on the website to have your personal data removed, however, this requires a photograph to be uploaded as it's the only way of matching the data it stores and cannot guarantee removal as the facial recognition system is not 100% accurate. This argument is something current legislation could find problematic legally, and therefore something to be addressed by future legislation if public concern is growing.

It should be noted that there are large scale AI systems used by governments in the UK and worldwide that utilise biometric data in live streams as a force to prevent crime. This has the potential to divide governments and its citizens (Marcus Smith & Seumas Miller, 2022). Has this been legally challenged in the past? On 11 August 2020, the UK court of appeal overturned the High Court's previous dismissal of a challenge to South Wales Police use of facial recognition technology. They found that it's use was unlawful and in violation of the European Convention of Human Rights (Simmons & Simmons), showing that even the police force and other members of government and public service may not be exempt in such situations. Thinking of ethical issues such as inaccuracy of the application, Big Brother Watch (Big Brother Watch, n.d) has said that facial recognition only had an approximate accuracy of 87% between 2016 and 2022 for the police force and over 3000 people were incorrectly identified. Incorrectly identifying criminals could lead to extreme consequences for the victims, they may have been arrested at work or in front of family and could affect their

livelihoods or mental wellbeing. Misidentification is also a major problem for PimEyes, uploading a single photograph returned results of many different individuals which the algorithm determined to be the same person.

PimEyes market themselves as a search engine for similar photographs, and search engines themselves bring an array of ethical issues. Gawker Stalker, launched in 2006, was a search engine for tracking celebrity sightings, which provided a Google maps layover of latest known whereabouts. Although, like PimEyes, only providing information that was previously available in the public domain, it came under much scrutiny for the ease of access to this data. Publicist Stan Rosenfeld commented on CBS News (CBS News, 2006) that the site is "taking advantage of the fact that there is probably no statute on the books at this time that governs this dissemination of information." Gawker Stalker wasn't shut down for its privacy concerns relating to its maps, however filed for bankruptcy in 2016 due to lawsuits for similar accusations of posting private videos. It is obvious that such an application can invade on the privacy of the public and PimEyes may be able to achieve similar stalkerish tendencies. Even considering some of the largest search engines available, such as Google or Yahoo, there is becoming a larger call for privacy and fairness. As data becomes available in faster and in bigger quantities, it is easier to find personal information at the click of a button. It is unlikely that Google will face any legal consequences, but they will need to react to the ever-changing legislations that are being set worldwide. Working in a similar way to google, PimEyes mines the photographic data from billions of websites, and therefore in the same way, they are not liable for misuse under current legislation.

In the terms and conditions of PimEyes (n.d), it has been acknowledged that the application can be used for harm, and an attempt to address this in writing has been made. It is unclear to what degree they are managing the said risk, and there is no discussion on how the monitoring and evaluation system named works. The transparency of the system could be another primary concern for the public, and is what the AI strategies of the future look to make mandatory.

"To ensure the mentioned objectives PimEyes has elaborated monitoring and evaluation system which includes weekly monitoring of guard and alert images to exclude unauthorized searches for children, women, and girls, as far as the mentioned groups are the most vulnerable to online stalking. Besides of the mentioned, set of indicators of suspicious activities have been elaborated and every user's activity is closely monitored notwithstanding the fact that PimEyes never checks for the information obtained by the users. Accounts with suspicious behaviour are banned temporarily, or permanently."

PimEyes could address these issues further and mitigate a great deal of criticism by ensuring the user is retrieving information of themselves by performing some identification checks. This would greatly reduce the risk of stalking and invasion of privacy, but some would remain with the algorithm not being 100% accurate, the likelihood of the results containing biometric data of others is high. A likely conclusion is that this is unlikely to be in the best financial interest of the stakeholders.

5. INDIVIDUAL ASSESSMENT OF PIMEYES' IMPACT ON STAKEHOLDERS

This section aims to offer a critical evaluation of the impact of PimEyes as an AI application on it's users, society, business and it's stakeholders with respect to the ethical issues offered in section 2 of this report. The key ethical areas focused on are issues including bias, fairness, accountability, transparency, explainability, sustainability and risk.

5.1 INDIVIDUALS (USERS)

Users of PimEyes include not only its registered users giving personal details to be kept in the database, but also those that are anonymously using the application. The privacy policy expresses that personal data is collected and kept by its registered users, including email address, full name, address, photo ID, however not all is compulsary to commit. There is no transparency in the way that the data is held and kept but assurances that steps are taken to ensure this data cannot be accessed by third parties. BBC News (June 2020) reported that 6,000 users were signed up at that time, it is easy to imagine the outrage and number of victims if any of the data was to compromised.

The users of the application itself could find themselves in deep waters providing they do not use the application in the correct manor, at least according the PimEyes itself who sees itself as unaccountable for the actions of individuals

using it unlawfully. Although this may be true under current legislation which has difficulties assigning accountability to the developers, emerging legislations in the UK and Europe at least are looking to rectify this in the coming months/years.

Harnessing PimEyes in its perfect form could be a power for good, that is, finding the indexes of a users personal photographs online so that it can be used to combat issues such as leaked photos, revenge porn, fake identities and other instances. There are many instances of online scams, fraud and so-called "catfishing" which can victimise individuals with terrible consequences. This application could be dedicated to providing an easy way to find where personal photographs are being used in this way and prevent such things from escalating. The notification service can easily be used for this purpose.

5.2 SOCIETY

Societal impacts of the application are the most worrying of all ethical implications as discussed in various sections of this report. The main criticisms of the application in the media are the way in which the application can be used unlawfully upon victims of society with regards to their privacy.

Facial recognition has had multitude of negative press reviews with businesses reversing decisions to use it on larger scales due to its issues with bias, that is, not being able to accruately recognise the faces it is attempting to find. PimEyes finds it no different in this respect; when trialling the application, a large number of faces that weren't of the person were returned by the engine. The existing issues in similar applications has been that they are more accurate at identifying people with lighter skin tones, exaggerating the need for review in terms of bias. Furthermore, specific workings of the algorithm used to match faces is kept very much secret, underlining the concerns of bias and fairness even more through a lack of transparency. If the decisions made by the AI application were to be reviewed publicly, even in a small capacity such as as a step-by-step guide, it could easily improve public trust. The reasons this hasn't been made public are unknown, but could imply the business has concerns of its own with regards to the ethical grey areas.

Privacy concerns include the mass crawling of photographic images from a growing list of available images on the internet, including personal photographic data being utilised by the application unknowingly to the members of society it may affect. PimEyes argue that this hasn't been a problem for years for websites such as Google and Yahoo, the results of these websites are text based web-pages and doesn't allow for the use of facial recognition to provide capabilities for stalking. Pehaps it is the sheer ease of use that provides the most concern amongst society for the PimEyes application, the process of uploading a photo and returning hundreds or thousands of photographic results in a matter of seconds. Sure enough, the user could utilise the AI application of Google and obtain similar results, however the whole process could then take months or years to provide access to the same data.

Current legislation GDPR states that "biometric data' means personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data". PimEyes seems to have found a grey area in such legislation as they do not retain actual biometric data in their database, only a string/integer based index dervied from images. The accountability for developers in such situations remains unclear and is something that emerging legislation is looking to clarify.

5.3 BUSINESS

Businesses have a growing number of frameworks and legislations to consider in today's world when building AI applications and PimEyes may have to make changes to the foundations of its application to comply. Legislation to combat the ethical issues of transparency could force PimEyes to be less secretive about its algorithm, and bias could put an end to facial recognition technology altogher in its current form.

Emerging legislations such as the "Proposal for Future Regulation of AI" (Section 3.2.3) aim to make changes to accountability of applications, rendering the developer responsible for the ethical violations. It is not yet clear exactly how future legislation will be implimented, however applications will be regulated on a case-by-case basis forcing developers to manage their AI owing to what they believe to be morally correct. Guidelines suggest that ethical principles are considered when in development however different organisations may take different stances on such cases. Legal

action normally results in hefty fines or application closures in regions, therefore businesses such as PimEyes will have a lot to consider when expanding internationally.

5.4 OTHER STAKEHOLDERS

Primary stakeholders such as owners, investors and employees have the most to lose when it comes to regulation violations. Ethical violations as covered in Section 4 victimise communities and can have legal consequences. Changes to the application to deal with issues including privacy, transparency and risk could mean identification of users and potentially reduce the customer base considerably, thus being financially damaging for all involved (decreased turnover, loss of jobs). Suppliers are an example of secondary stakeholders that could also feel the effects of such changes in terms of financial damage.

Conversely, promoting PimEyes as an application that can protect communities from harm could have the opposite effect, accessing a new customer base by moving towards a transparent and morally inclusive development.

6. CONCLUSION

Ethical implications of applications are a complex, yet vitally important topic in artificial intelligence. This report has provided in-depth reviews of such ethical issues such as fairness, bias, transparency, accountability and more. Globally, organisations and governments such as UNESCO, ICO, HM Government, European Commission and many more are working together to build regulation in response to these issues.

In the specific instance of PimEyes, finding many similarities in prior cases such as Clearview AI (Information Commissioners Office, 2017), in which the facial recognition software was deemed as a threat to black communities (Brewster, 2022). The largest public concern of usage is that the application is "tailor-designed for stalkers" (Hvistendahl, 2022), creating ethical implications in privacy and is likely to be the eventual downfall of the business provided it is not addressed. Database management systems also represent one of the largest energy consumers (Roukh, et al., 2017), and with new resistance fighting global warming, this is another major cause for concern. PimEyes on the face appear committed to safeguarding users and society, themselves confirming they are "'working to develop better safeguards for children" (Hvistendahl, 2022: online), however it is yet to be seen publicly how exactly this is being developed.

On the contrary, issues such as identity fraud, revenge porn and catfishing can help be addressed by using such an application; and with $5.8 billion lost by Americans due to stolen identities (National Council on Identity Theft Protection, 2022), it is important to recognise the potential PimEyes has to benefit societies.

Overall, PimEyes is a major cause of public concern in its current state, the usage is not monitored with it being easy to submit photographs of a multitude of people from a variety of backgrounds without questions asked by the developers. That said, the potential for positive use is there if ethical issues are considered carefully by safeguarding tactics such as identification of users and single identification photographs only can be used.

7. BIBLIOGRAPHY

Alam, M. et al., 2021. Combined PCA–Segmentation Method: An Efficient Technique for Covered Face Recognition. *Advances in Intelligent Systems and Computing*, 26 October, Volume 1397, pp. 831-839.

Alhadeff, J. A. B. a. D. J., 2012. The accountability principle in data protection regulation: origin, development and future directions. In: *Managing privacy through accountability.* London: Palgrave Macmillan, pp. 49-82.

BBC News, 2021. *Amazon hit with $886m fine for alleged data law breach.* [Online]
Available at: https://www.bbc.co.uk/news/business-58024116
[Accessed 2022 December 6].

BBC News, 2022. *Facebook: Meta fined €265m by Irish Data Protection Commission.* [Online]
Available at: https://www.bibme.org/harvard/website-citation/confirm
[Accessed 6 December 2022].

Bass, A., 2022. Smile! You're on Camera: Police Departments' Use of Facial Recognition Technology & the Fourth Amendment. *Adriana Bass, 'Smile! You're on Camera: Police Departments' Use of Facial Recognition Technology & the Fourth Amendment' (2022) 55(4) Loyola of Los Angeles Law Review*, 55(4), pp. 1053-1084.

Big brother watch complaint against facial recognition search engine (1970) *AWO.agency*. Available at: https://www.awo.agency/blog/bbw-pimeyes-complaint/ (Accessed: January 12, 2023).

Big Brother Watch (n.d.) *Stop facial recognition*, *Stop Facial Recognition - Big Brother Watch*. Available at: https://bigbrotherwatch.org.uk/campaigns/stop-facial-recognition/ (Accessed: January 13, 2023).

Bleicher, A. (2017) *Demystifying the Black Box That Is AI*, 9 August. Available at: https://www.scientificamerican.com/article/demystifying-the-black-box-that-is-ai/ (Accessed: December 10, 2022).

Brewster, T., 2022. *A "Threat To Black Communities": Senators Call On Immigration Cops And FBI To Quit Using Clearview Facial Recognition.* [Online]
Available at: https://www.forbes.com/sites/thomasbrewster/2022/02/09/a-threat-to-black-communities-senators-call-on-immigration-cops-and-fbi-to-quit-using-clearview-facial-recognition/?sh=1b37bc756d06
[Accessed 16 December 2022].

Briffaa, J., Sinagrab, E. & Blundell, R., 2020. Heavy metal pollution in the environment and their toxicological effects on humans. *Heliyon*, September.6(9).

Boddington, P. AI and moral thinking: how can we live well with machines to enhance our moral agency?. *AI Ethics* **1**, 109–111 (2021). https://doi.org/10.1007/s43681-020-00017-0

Busuioc, M. (2022, August 12). AI algorithmic oversight: new frontiers in regulation. *Handbook of Regulatory Authorities*, 470-486.

Calder, A., 2018. *EU GDPR: A Pocket Guide.* 2nd Edition ed. s.l.:IT Governance Publishing.

Carlin, D., 2022. *Cop 27 recap: The good, the bad, and what's next after the Climate Conference.* [Online]
Available at: https://www.forbes.com/sites/davidcarlin/2022/12/16/cop-27-recap-the-good-the-bad-and-whats-next-after-the-climate-conference/?sh=69e96092c99b
[Accessed 19 December 2022].

Chanakira, A. (2022) *AI DATA IS THE NEW GOLD RUSH*, 19 September. Available at: https://economist.com.na/73549/columns/ai-data-is-the-new-gold-rush/ (Accessed: January 4, 2023).

Coeckelbergh, M. (2020). AI ethics. *Mit Press*.

Crockett, K., Goltz, S. & Garratt, M., 2018. GDPR impact on Computational Intelligence Research. *nternational Joint Conference on Neural Networks (IJCNN)*, pp. 1-7.

Collins, D. (2022) *UK sets out proposals for new AI rulebook to Unleash Innovation and Boost Public Trust in the technology*, *GOV.UK*. Edited by Department for Digital, Culture, Media & Sport. Available at: https://www.gov.uk/government/news/uk-sets-out-proposals-for-new-ai-rulebook-to-unleash-innovation-and-boost-public-trust-in-the-technology

Defined, A.I. (2021) *Artificial Intelligence Benefits to society - definedcrowd*, *Defined.ai*. Available at: https://www.defined.ai/blog/the-top-5-reasons-to-be-grateful-for-ai/ (Accessed: January 2, 2023).

Dhinakaran, A. (2021) *Overcoming AI's Transparency Paradox*, 10 September. Available at: https://www.forbes.com/sites/aparnadhinakaran/2021/09/10/overcoming-ais-transparency-paradox/?sh=5a07c14a4b77 (Accessed: December 15, 2022).

Dhinakaran, A. (2022) *Recap: The man, the Machine, and the black box*, *Recap: The Man, The Machine, and The Black Box*. Available at: https://arize.com/blog/recap-machine-learning-conference/ (Accessed: December 15, 2022).

D. Schiff, J. Borenstein, J. Biddle and K. Laas, "AI Ethics in the Public, Private, and NGO Sectors: A Review of a Global Document Collection," in *IEEE Transactions on Technology and Society*, vol. 2, no. 1, pp. 31-42, March 2021, doi: 10.1109/TTS.2021.3052127.

European Commission - European Commission, 2022. *Questions and Answers: EU action plan on digitalising the energy system.* [Online]
Available at: https://ec.europa.eu/commission/presscorner/detail/en/qanda_22_6229
[Accessed 19 December 2022].

European Commission (2022) *Proposal for a directive of the European parliament and of the council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)*, *European Union*. European Commission. Available at: https://ec.europa.eu/info/sites/default/files/1_1_197605_prop_dir_ai_en.pdf

European Commission (April 2021) *Regulatory framework proposal on Artificial Intelligence*, *Shaping Europe's digital future*. European Commission. Available at: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

European Court of Human Rights, 1950. *European Convention on Human Rights.* s.l.:s.n.

GDPR.EU, 2018. *Art. 22 GDPR - automated individual decision-making, including profiling.* [Online]
Available at: https://gdpr.eu/article-22-automated-individual-decision-making/
[Accessed 6 December 2022].

Facial Recognition Database 'Risks Targeting Innocent People' (2017), 14 September. Available at: https://www.bbc.co.uk/news/uk-41262064 (Accessed: December 2, 2022).

Guo, W. (2020) "Explainable artificial intelligence for 6G: Improving trust between human and Machine," *IEEE Communications Magazine*, 58(6), pp. 39–45. Available at: https://doi.org/10.1109/mcom.001.2000050.

Harwell, D. (2021). *This facial recognition website can turn anyone into a cop — or a stalker*. [Online]. The Washington Post. Last Updated: 14 May 2021. Available at: https://www.washingtonpost.com/technology/2021/05/14/pimeyes-facial-recognition-search-secrecy/ [Accessed 07 December 2022].

Hern, A. (2016) *Stephen Hawking: AI will be 'either best or worst thing' for humanity*, 19 September. Available at: https://www.theguardian.com/science/2016/oct/19/stephen-hawking-ai-best-or-worst-thing-for-humanity-cambridge (Accessed: December 14, 2022).

Heikkilä, M. (2021). *Clearview scandal exposes limits of transatlantic AI collaboration*. [Online]. Politico. Last Updated: 8 April 2021. Available at: https://www.politico.eu/article/clearview-scandal-exposes-limits-transatlantic-ai-facial-recognition [Accessed 07 December 2022].

Hill, K. (2022) "A Face Search Engine Anyone Can Use Is Alarmingly Accurate," *The New York Times*, 26 May. Available at: https://www.nytimes.com/2022/05/26/technology/pimeyes-facial-recognition-search.html (Accessed: December 10, 2022).

Hvistendahl, M. (2022). *facial recognition database. Photo illustration: Elise Swain/The Intercept Facial Recognition Search Engine Pulls Up "*. [Online]. The Intercept. Last Updated: 16 July 2022. Available at: https://theintercept.com/2022/07/16/facial-recognition-search-children-photos-privacy-pimeyes/ [Accessed 07 December 2022].

Hvistendahl, M. (2022) *The Intercept*, 16 July. Available at: https://theintercept.com/2022/07/16/facial-recognition-search-children-photos-privacy-pimeyes/ (Accessed: December 12, 2022).

Hill, K., 2022. *Wrongfully Accused by an Algorithm.* [Online]
Available at: https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html
[Accessed 16 December 2022].

Huang, J.Y., Gupta, A. & Youn, M. Survey of EU ethical guidelines for commercial AI: case studies in financial services. *AI Ethics* **1**, 569–577 (2021). https://doi.org/10.1007/s43681-021-00048-1

Ibrahim, S., Pronovost, P. (2021). Diagnostic Errors, Health Disparities, and Artificial Intelligence: A Combination for Health or Harm?. *JAMA Health Forum*. 2(9).

Information Commissioner's Office. (2019). *Human bias and discrimination in AI systems*. [Online]. ICO. Last Updated: 25 June 2019. Available at: https://ico.org.uk/about-the-ico/media-centre/ai-blog-human-bias-and-discrimination-in-ai-systems/ [Accessed 03 December 2022]

HM Government (ed.) (2021) *National AI Strategy*, *GOV UK*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1020402/National_AI_Strategy_-_PDF_version.pdf
Information Commissioner's Office. (2022). *ICO fines facial recognition database company Clearview AI Inc more than £7.5m and orders UK data to be deleted*. [Online]. ICO. Last Updated: 23 May 2022. Available at: https://ico.org.uk/about-the-ico/media-centre/news-and-blogs/2022/05/ico-fines-facial-recognition-database-company-clearview-ai-inc/ [Accessed 07 December 2022].
Information Commissioner's Office. (n.d.). *Codes Of Conduct*. [Online]. ICO. Available at: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/codes-of-conduct/ [Accessed 07 December 2022].
International Commissioner's Office (n.d.) *Guidance on AI and Data Protection*, *ICO*. Available at: https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/guidance-on-ai-and-data-protection/
Information Commissioner's Office. (n.d.). *Data protection impact assessments*. [Online]. ICO. Last Updated: 2022. Available at: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/data-protection-impact-asessments/ [Accessed 07 December 2022].
International Commissioner's Office (n.d.) *What is special category data?*, *ICO*. Available at: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/special-category-data/what-is-special-category-data/ (Accessed: January 13, 2023).
Johnson, K. (2022) "How Wrongful Arrests Based on AI Derailed 3 Men's Lives," *WIRED*, 7 March. Available at: https://www.wired.com/story/wrongful-arrests-ai-derailed-3-mens-lives/ (Accessed: December 10, 2022).
Kaplan, R. M., Chambers, D. A. & G. R. E., 2014. Big Data and large sample size: A cautionary note on the potential for bias. *Clinical and Translational Science*, 15 July, 7(4), pp. 342-346.
Karasavva, V., Noorbhai, A. (2021). The Real Threat of Deepfake Pornography: A Review of Canadian Policy. *Cyberpsychology, Behavior, and Social Networking*. 24(3), pp.203-209.
Kaye, B. (2022). *Australian retailer pauses facial recognition trial over privacy complaint*. [Online]. Reuters. Last Updated: 28 June 2022. Available at: https://www.reuters.com/technology/australian-appliances-giant-pauses-facial-recognition-tech-over-privacy-concerns-2022-06-28/ [Accessed 07 December 2022].
Klosowski, T. (2020) "Facial Recognition Is Everywhere. Here's What We Can Do About It.," *New York Times*, 15 July. Available at: https://www.nytimes.com/wirecutter/blog/how-facial-recognition-works/ (Accessed: December 13, 2022).
Kostka, G., Steinacker, L. & Meckel, M., 2021. Between security and convenience: Facial recognition technology in the eyes of citizens in China, Germany, the United Kingdom, and the United States. *Public Understanding of Science*, 26 March, 30(6), pp. 671-690.
Kuner, C., 2009. An international legal framework for Data Protection: Issues and Prospects. *Computer Law and Security Review*, July, 25(4), pp. 307-317.
Marks, P., 2021. Can the biases in facial recognition be fixed; also, should they?. *Communications of the ACM*, March, 64(3), pp. 20-22.
Landymore, F. (2022). *Experts Horrified by Facial Recognition Site That Digs Up "Potentially Explicit" Photos of Children*. [Online]. Futurism. Last Updated: 19 July 2022. Available at: https://futurism.com/experts-horrified-by-facial-recognition-site-that-digs-up-potentially-explicit- [Accessed 07 December 2022].
Leslie, D. (2019). Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector. The Alan Turing Institute.
Matthias , A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*. 6(3), pp.175-183.
Meaker, M. (2022). *Clearview Stole My Face and the EU Can't Do Anything About It*. [Online]. Wired. Last Updated: 07 November 2022. Available at: https://www.wired.co.uk/article/clearview-face-search-engine-gdpr [Accessed 07 December 2022].
Megorskaya, O. (2022) *Training Data: The Overlooked Problem Of Modern AI*, 27 June. Available at: https://www.forbes.com/sites/forbestechcouncil/2022/06/27/training-data-the-overlooked-problem-of-modern-ai/?sh=712d5e5f218b (Accessed: December 15, 2022).
Metz, R., 2021. *Anyone can use this powerful facial-recognition tool — and that's a problem.* [Online] Available at: https://edition.cnn.com/2021/05/04/tech/pimeyes-facial-recognition/index.html [Accessed 16 December 2022].

Nadeem, A., Marjanovic, O., Abedin, B. (2021). *Responsible AI and Analytics for an Ethical and Inclusive Digitized Society*. London: Springer. pp.259-270.
Naim, A., 2021. Green Information Technologies in Business Operations. *Journal of Modern Philosophy, Social Sciences and Humanities*, December, Volume 1, pp. 36-49.
Najibi, A., 2020. *Racial Discrimination in Face Recognition Technology*. s.l.:s.n.
National Council On Identity Theft Protection. (2022). *2022 Identity Theft Facts and Statistics*. [Online]. Available at: https://identitytheft.org/statistics/ [Accessed 11 December 2022].
Nash, J. (2022) *Complaint filed against PimEyes in UK as facial recognition web search options grow: Biometric Update*, *Biometric Update |*. BiometricUpdate.com. Available at: https://www.biometricupdate.com/202211/complaint-filed-against-pimeyes-in-uk-as-facial-recognition-web-search-options-grow (Accessed: January 6, 2023).
Ocasio-Cortez, A. (2019) *We have started to sound the alarm on the way facial recognition technology is expanding in concerning says.from the FBI to ice to Amazon, the bar for consent and civil liberties protection is repeatedly violated, and on top of it all has a disproportionate racial impact, too. https://t.co/5j9avidtcc*, *Twitter*. Twitter. Available at: https://twitter.com/AOC/status/1150117385870987264 (Accessed: December 1, 2022).
O'Connell, A. (2020). Image rights and image wrongs: image-based sexual abuse and online takedown. *Journal of Intellectual Property Law and Practice*. 15(1), pp.55-65.
ÓhÉigeartaigh, S. *et al.* (2020) "Overcoming barriers to cross-cultural cooperation in AI ethics and governance," *Philosophy & Technology*, 33(4), pp. 571–593. Available at: https://link.springer.com/article/10.1007/s13347-020-00402-x (Accessed: January 4, 2023).
Owe, A. and Baum, S.D. (2021) *The ethics of sustainability for Artificial Intelligence*, *The Ethics of Sustainability for Artificial Intelligence*. MAIEI. Available at: https://gcrinstitute.org/papers/060_sustainability-ai.pdf (Accessed: December 15, 2022).
Pandit, H. (2022). *Towards a Knowledge-Aware AI*. Amsterdam: IOS Press. pp.36-50.
PimEyes, n.d. *PimEyes' Blog: How to improve the facial recognition search results.* [Online]
Available at: https://pimeyes.com/en/blog/how-to-improve-the-facial-recognition-search-results
[Accessed 16 December 2022].
PimEyes, n.d. *About Pimeyes.* [Online]
Available at: https://pimeyes.com/en/about
[Accessed 16 December 2022].
PimEyes, n.d. *More about Pimeyes' database and opt-out service.* [Online]
Available at: More about Pimeyes' database and opt-out service
[Accessed 18 December 2022].
PimEyes, n.d. *Pimeyes' blog: How to improve the facial recognition search results.* [Online]
Available at: https://pimeyes.com/en/blog/how-to-improve-the-facial-recognition-search-results
[Accessed 18 December 2022].
PimEyes (n.d.) *Terms and conditions of use*, *PimEyes.* Available at: https://pimeyes.com/en/premium-terms (Accessed: January 12, 2023).
Robinson, N., Graux, H., Botterman, M. & Valeri, M., 2009. *Review of the European Data Protection Directive,* s.l.: the RAND Corporation.
Roselli, D., Matthews, J., Talagala, N. (2019). Managing Bias in AI. *Proceedings of The 2019 World Wide Web Conference*. 1, pp.539-544. [Online]. Available at: https://dl.acm.org/doi/10.1145/3308560.3317590 [Accessed 03 December 2022].
Roukh, A., Ladjel, B., Bouarar, S. & Boukorca, A., 2017. Eco-Physic: Eco-physical design initiative for very large databases. *Information Systems,* August, Volume 68, pp. 44-63.
Scipione, J.,Lo Moncaco, M. (2022). *Has the Horse Bolted? Dealing with Legal and Practical Challenges of Facial Recognition.* [Online]. SSRN Electronic Journal. Last Updated: 25 March 2022. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4019105 [Accessed 11 December 2022].
Smith, A., 2019. More Than Half of U.S. Adults Trust Law Enforcement to Use Facial Recognition Responsibly. *Pew Research Center*, September.
Shin, D. (2021) "The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI," *International Journal of Human-Computer Studies*, 146. Available at: https://doi.org/10.1016/j.ijhcs.2020.102551.

Sindermann, C. *et al.* (2022) "Acceptance and fear of Artificial Intelligence: Associations with personality in a German and a Chinese sample," *Discover Psychology*, 2(1). Available at: https://doi.org/10.1007/s44202-022-00020-y.
Simmons & Simmons (2022) *UK Court of Appeal finds facial recognition technology unlawful*, *Simmons & Simmons*. Available at: https://www.simmons-simmons.com/en/publications/ckelg1z7p8kjt09008l0bet7c/uk-court-of-appeal-finds-facial-recognition-technology-unlawful
Smith, C. (2018). *Facebook will have to face a massive lawsuit over facial recognition*. [Online]. BGR. Last Updated: 17 April 2018. Available at: https://bgr.com/tech/facebook-facial-recognition-class-action-suit/ [Accessed 07 December 2022].
Smith, Marcus & Miller, Seumas (2022). The ethical application of biometric facial recognition technology. AI and Society 37 (1):167-175.
Songa, Z., Zhangb, X. & Erikssona, C., 2015. Data Center Energy and Cost Saving Evaluation. *The 7th International Conference on Applied Energy – ICAE2015*, p. 1255 – 1260.
Statista Research Department. (2023, January 6). *Global AI in marketing revenue 2028*. Retrieved January 9, 2023, from Statista: https://www.statista.com/statistics/1293758/ai-marketing-revenue-worldwide/
Sweeney, L. (2013). *Discrimination in Online Ad Delivery*. [Online]. SSRN. Last Updated: 29 January 2013. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2208240 [Accessed 03 December 2022].
Teich, D. (2020). *Artificial Intelligence And Data Privacy – Turning A Risk Into A Benefit*. [Online]. Forbes. Last Updated: 10 August 2020. Available at: https://www.forbes.com/sites/davidteich/2020/08/10/artificial-intelligence-and-data-privacy--turning-a-risk-into-a-benefit/ [Accessed 03 Decemeber 2022].
Syed, M. (2015) *Black box thinking*. Penguin Publishing Group.
UK Information Commissioner's Office (ICO) fines Clearview AI £7.5m (2022) *Simmons & Simmons*. Available at: https://www.simmons-simmons.com/en/publications/cl3ldfkeb15do0a897k1ijah5/uk-information-commissioner-s-office-ico-fines-clearview-ai-7-5m (Accessed: December 12, 20232).
The European Parliament and the Council of the European Union, 1995. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection. *Official Journal of the European Communities*, Volume 281/31.
Toth, Z., Caruana, R., Gruber, T., Loebbecke, C. (2022). The Dawn of the AI Robots: Towards a New Framework of AI Robot Accountability. *Journal of Business Ethics*. 178, pp.895-916.
UNESCO (2022) *Recommendation on the Ethics of Artificial Intelligence*, *unesco.org*. UNESCO. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000381137_eng
Valentino, S. (2022). *Facial-recognition technology is appearing in US airports. SFO is likely next.*. [Online]. SF Gate. Last Updated: 6 December 2022. Available at: https://www.sfgate.com/travel/article/facial-recognition-technology-coming-to-sfo-17633303.php [Accessed 07 December 2022].
Vallance, C. (2022). *Stalking fears over PimEyes facial search engine*. [Online]. BBC News. Last Updated: 8 November 2022. Available at: https://www.bbc.co.uk/news/technology-63544169.amp [Accessed 13 December 2022].
Verma, S.S. (2022) *5 reasons why AI is not sustainable in the long term*, *Medium*. CodeX. Available at: https://medium.com/codex/5-reasons-why-ai-is-not-sustainable-in-the-long-term-75495e2bf9c (Accessed: January 1, 2023).
Vincent, J. (2016). *Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day*. [Online]. The Verge. Last Updated: 24 March 2016. Available at: https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist [Accessed 03 December 2022].
Wachter, S., 2018. Normative challenges of identification in the Internet of Things: Privacy, profiling, discrimination, and the GDPR. *Computer Law & Security Review*, Volume 34, p. 436–449.
Wilson, D. & Jahankhani, H., 2021. A proposed OKR-based framework for cyber effective services in the GDPR era. *Strategy, Leadership, and AI in the Cyber Ecosystem*, pp. 113-135.
Wolford, B., 2022. *What is GDPR, the EU's new data protection law?.* [Online]
Available at: https://gdpr.eu/what-is-gdpr/
[Accessed 6 December 2022].
Yan, S., 2021. Algorithms are not bias-free: Four mini-cases. *Human Behavior and Emerging Technologies*, 7 October, 3(5), pp. 1180-1184.
Zerilli, J. *et al.* (2018) "Transparency in algorithmic and human decision-making: Is there a double standard?," *Philosophy & Technology*, 32(4), pp. 661–683. Available at: https://doi.org/10.1007/s13347-018-0330-6.
Zhou, J., Chen, F., Holzinger, A. (2020). *Beyond Explainable AI*. London: Springer. pp.375-386.

Zurah, S., 2022. *The myth of facial recognition bias.* [Online]
Available at: https://www.clearview.ai/post/the-myth-of-facial-recognition-bias
[Accessed 18 December 2022]

APPENDIX A

6G7V0035 Data Management and Governance – 1CW60 Feedback Sheet Data Ethics Report

Name:_____ Grade %

Background Research	*limited / missing*	*sufficient*	*extensive*
Key ethical issues			
Data governance practices			
Ethical Guidelines			
Legislation			

Application Selection and Justification	*limited / missing*	*sufficient*	*extensive*
Application description			
Justification			

Analysis	*limited / missing*	*sufficient*	*extensive*
Ethical Issues			
Relationships to ethical principals			
Social issues			
Legal Issues			
Environmental issues			
Cultural / Gender / Age issues			

- Moral issues

Impact *limited / missing* *sufficient* *extensive*

- Individuals (users)
- Society
- Business
- Other stakeholders

Referencing: *limited / missing* *sufficient* *extensive*

- Appropriate citations in text
- Bibliography / Reference list
- MMU Harvard style

Writing style: *incomprehensible* *coherent* *fluent*

- Communication of ideas
- Professional and academic style

Mark	% of marks available	% of marks Awarded	Element
Group			Report Introduction and Conclusion
Group			Key ethical issues and data governance practices
Group			An overview of current and emerging ethical guidelines, frameworks, principles and legislation.

Group			A brief description of an appropriate and recent real world data mining / Artificial Intelligence / data driven application, that has been featured in the media. The application is selected by the group
Group			A rigorous and multifaceted analysis of the social, moral, cultural, environmental, legal and ethical issues of this application with supporting evidence through literature. This section will include a discussion of relevant key ethical issues and their relationship to ethical principles. It may also include the use of protected characteristics.
Individual			Your individual balanced assessment of the impact of the application on individuals, society, business, and other stakeholders.
Group			Report structure, correct referencing style, professionalism
Individual			Peer review

The marking will be the subject of a QA check ☐ **(moderation) and the mark awarded could be revised.**

Individual mark based on peer assessment weighting is:

Comments:

What you did well

What you can improve upon

APPENDIX B

Peer Review Template

Peer Evaluation

Please use this form to evaluate the contributions of each team member to the group effort. Consider attendance and participation in team meetings, individual contributions to idea generation and research, communication within the group, etc. *These evaluations are completely confidential and will never be shown to your team members. Please respond as honestly as possible.*

1. Please allocate a total of 100 percentage points among your team member, including yourself, with higher percentages going to those members who contributed most. In the case of equal contribution, points should be divided equally among team members.

   Your name: **Christian Jordan**

   Your student number: **14061768**

	Name Christian Jordan	**% Points**
Yourself	**29**	
Daniel Cawley	**29**	
Jamie Atiyah	**29**	
Lauren Salkeld	**13**	
	Total	**100%**

2. Explain any particularly high or low allocations, providing concrete examples to illustrate your reasoning.

   Daniel and Jamie worked well however Lauren continuously failed to meet group deadlines for submission. Lauren's first submission for explaining ethical principles was minimal, only finalising the day before deadline despite requests. This was the same time when the group work on moral issues was submitted, leaving very little time for amendments and formatting.