ECS7025P – Ethics, Regulation And Law In Advanced Digital Information Processing And Decision Making

INDIVIDUAL ASSIGNMENT (RESUBMISSION)

SCHOOL OF ELECTRONIC ENGINEERING AND COMPUTER SCIENCE (EECS)

MOHAMMAD ASGHAR
STUDENT ID: 160546576
DATE OF SUBMISSION: 08/07/2024

Executive Summary

Artificial intelligence (AI) systems possess immense potential to reshape human interactions and societal progress, but their ethical implementation demands careful consideration of autonomy, the right of explanation, and value alignment. This Ethics report explores these aspects across four detailed sections:

Section A: Literature Review surveys AI's transformative impact in the UK, emphasizing the ethical challenges it presents, such as job displacement, privacy concerns, and biases. It advocates for balanced approaches including regulatory measures and strategies for equitable AI integration.

Section B: Data Ethics examines the British Airways data breach in 2018, discussing GDPR compliance and the ongoing challenge of securing modern IT infrastructures against significant data breaches. It calls for increased investment in cybersecurity.

Section C: AI & Ethical Issues outlines the application of AI at Infosys, focusing on how the company addresses ethical challenges like algorithmic bias and maintains transparency. Infosys's proactive measures include establishing ethical guidelines and ensuring human oversight in AI applications.

Section D: Development and Implementation of an Ethical AI Framework details the creation of a bespoke ethical framework at Infosys, highlighting principles such as fairness, transparency, and accountability. It also describes strategies for framework implementation, including employee training and stakeholder engagement.

The report underscores the significance of continuously addressing ethical concerns in AI development to harness its benefits responsibly and equitably.

Table of Contents

List of Abbreviations

AI:	Artificial Intelligence
AWS:	Amazon Web Services
BA:	British Airways
CDEI:	Centre for Data Ethics and Innovation
CISA:	Cybersecurity and Infrastructure Security Agency
COIN™:	Co-Innovation Network
DPIAs:	Data Protection Impact Assessments
GDPR:	General Data Protection Regulation
GPAI:	Global Partnership on Artificial Intelligence
IBM:	International Business Machines Corporation
ICO:	Information Commissioner's Office
KPIs:	Key Performance Indicators
NCSC:	National Cyber Security Centre
NHS:	National Health Service
NIST:	National Institute of Standards and Technology
OECD:	Organisation for Economic Co-operation and Development
OWASP:	Open Web Application Security Project
R&D:	Research and Development
RUSI:	Royal United Services Institute
TCS:	Tata Consulting Services
US-CERT:	United States Computer Emergency Readiness Team
XAI:	Explainable Artificial Intelligence

Introduction

As AI rapidly advances, its societal, economic, and human impacts necessitate a focus on ethical implications. The motivations for ethical AI are driven by the need to harness AI's potential while mitigating risks and ensuring alignment with human values.

A primary motivation is the protection of human rights and individual freedoms. AI's role in decision-making, from job applications to criminal justice, demands fairness, transparency, and accountability. Addressing potential biases and discrimination is crucial to prevent exacerbating societal inequalities.

Public trust and long-term sustainability of AI technologies are also vital. Without ethical guidelines, AI adoption may face resistance, hindering innovation. As AI systems become more autonomous, questions of moral responsibility and alignment with human values intensify.

To address these concerns, various organizations and governments are developing ethical AI frameworks. These frameworks provide guidelines and best practices, emphasizing principles like beneficence, non-maleficence, accountability, transparency, fairness, and respect for human rights.

Ethical AI considerations include privacy protection, data governance, algorithmic bias, transparency, and socioeconomic impacts of AI-driven automation. By proactively implementing ethical frameworks, we can ensure AI systems advance technological capabilities while positively contributing to human well-being and societal progress.

Section A: Literature Review

The influence of AI in the United Kingdom (UK), along with the ethical challenges it presents, is garnering significant attention from academics, policymakers, and industry experts. For those aspiring to become data scientists or AI professionals, it's essential to grasp the complex effects of AI on both society and the economy. Understanding and addressing these ethical issues is vital to promote the responsible and advantageous use of AI.

AI's Impact in the UK

The UK has established itself as a leader in AI research and development globally, with significant investments and initiatives aimed at harnessing the technology's potential across various sectors. Per the report by Tech Nation (2023), the UK's AI sector brought £6.6 billion in investment in 2022, making it the third-largest AI market globally after the United States and China. Due to this substantial investment, several AI startups have been founded and AI technologies have been incorporated into already-existing enterprises, fostering innovation and economic growth.

The major area where AI is creating immense impact in the UK is healthcare. The National Health Service (NHS) has been actively exploring AI applications to improve patient care, reduce waiting times, and enhance diagnostic accuracy. These tools are being used to analyze medical images, predict patient outcomes, and assist in treatment planning, resulting in a more efficient healthcare system.

Within the finance area, AI is revolutionizing risk assessment, fraud detection, and customer service. The Bank of England (2022) reported that 72% of UK financial institutions are now using AI and machine learning in some capacity, with applications ranging from algorithmic trading to personalized financial advice. This adoption of AI technologies has led to improved operational efficiency and enhanced customer experiences in the banking and finance industry.

Nonetheless, the increased use of AI technologies has also raised concerns about the displacement of humans by AI, underscoring the need for workforce reskilling. A study by PwC suggests that by the mid-2030s, up to 30% of jobs in the UK may be significantly susceptible to automation (PwC UK, 2024). Industries like transportation, manufacturing, and retail are expected to face the highest risks. This highlights the importance of developing comprehensive strategies for workforce adaptation and lifelong learning to assure that the advantages of AI are dispersed fairly throughout society.

Ethical Dilemmas and Issues

The increased use of AI technologies had led to various ethical issues, particularly around bias and discrimination. According to a research by the Alan Turing Institute (2023), AI algorithms risk reinforcing or even heightening societal biases in sectors like employment, finance, and criminal

justice. This issue is especially applicable in the UK, that is increasingly focused on preventing AI systems from deepening existing disparities or introducing new forms of inequality.

To tackle these issues, the government of UK has initiated the Centre for Data Ethics and Innovation (CDEI), which advocates for diverse AI development teams and auditing of AI systems to curb biases. Moreover, the UK's Information Commissioner's Office (ICO) has issued guidelines for AI and data protection, stressing the significance of transparency and accountability in AI-powered decisions (ICO, 2023).

Another significant ethical concern is the influence of AI on privacy and data protection. AI systems require large volumes of data to operate efficiently, raising concerns about how personal data is collected, stored, and utilized. The UK General Data Protection Regulation (GDPR) establishes a data protection framework in the era of AI, yet there are ongoing challenges in harmonizing the demands for data-driven innovation with the protection of individual privacy rights (Department for Science, Innovation and Technology, 2023).

The ethical implications of AI in surveillance and security applications have also come under scrutiny in the UK. The installation of facial recognition technology by law enforcement agencies has ignited discussions over privacy, consent, and potential abuse. According to a report by the Royal United Services Institute (RUSI, 2023), there is a necessity for well-defined governance and public dialogue to ensure that AI surveillance tools are used ethically and are properly regulated.

Autonomous systems and AI-driven decision-making raise questions about accountability and human oversight. In sectors such as healthcare and finance, where AI platforms are widely being utilized to make critical decisions, there is a need to establish clear lines of responsibility and ensure that humans remain in control of key decision-making processes. The UK's AI Council (2023) has emphasized the significance of developing "human-in-the-loop" systems that merge the strengths of AI with human ingenuity.

Handling Ethical Dilemmas

To handle these ethical dilemmas, a multifaceted approach is required which comprises collaboration between government, industry, academia, and civil society. The UK has taken several steps to establish frameworks for responsible AI development and deployment.

The AI Ethics and Safety Framework, developed by the Office for Artificial Intelligence, provides guidance for public sector organizations on the ethical use of AI. This framework emphasizes principles like fairness, accountability, transparency, and explainability, offering practical tools for examining and reducing ethical risks in AI projects (Department for Science, Innovation and Technology, Office for Artificial Intelligence and Centre for Data Ethics and Innovation, 2019).

Education and training have a crucial part in preparing future AI practitioners to navigate ethical dilemmas. Universities across the UK are increasingly incorporating ethics courses into their computer science and data science curricula. For example, the University of Oxford's Institute for Ethics in AI (2023) offers interdisciplinary programs that bring together technical expertise with

philosophical and ethical considerations. With this, Queen Mary University of London's School of Electronic Engineering and Computer Science offers modules on Ethics in Information Processing and Decision Making to teach students how to incorporate ethical and legal standards into AI and Data Science, guided by frameworks like the UK Government's Ethical Framework (Queen Mary University of London, 2024).

Regulatory approaches are evolving to handle the unique challenges presented through AI technologies. The UK's proposed AI regulatory framework, as mentioned in the AI White Paper (Department for Science, Innovation and Technology, 2023), adopts a risk-based strategy that aims to balance innovation with the need for appropriate safeguards. This framework proposes context-specific regulations for high-risk AI applications while encouraging self-regulation and the development of industry standards for lower-risk use cases.

International collaboration is also crucial in addressing global AI ethics challenges. The UK has been actively involved in measures like the Global Partnership on Artificial Intelligence (GPAI) and the OECD AI Principles, working towards developing common ethical standards and governance frameworks for AI on a global scale (Foreign, Commonwealth & Development Office and The Rt Hon James Cleverly, 2023).

As future data scientists and AI practitioners, it is essential to engage with these ethical considerations throughout the AI development lifecycle. This involves conducting thorough impact assessments, implementing robust testing and validation procedures, and actively seeking diverse perspectives to identify and mitigate potential ethical risks.

Furthermore, fostering a culture of ethical awareness and responsibility within organizations is crucial. This can be achieved through the establishment of ethics committees, regular ethics training for AI teams, and the incorporation of ethical considerations into performance evaluations and project planning processes.

Transparency and explainability in AI systems are key in terms of developing public trust and enabling effective oversight. Methods like explainable machine learning and model-agnostic explanation approaches are being developed to make AI decision-making processes more understandable to both experts and non-experts alike (Royal Society, 2023).

The ethical ramifications of AI stretch beyond technical aspects to wider impacts on society (Almasri, 2024). As such, interdisciplinary collaboration between AI practitioners, ethicists, social scientists, and policymakers is essential to develop comprehensive approaches to AI ethics that consider the technology's wider societal implications.

In conclusion, as the UK continues to advance its position as a global leader in AI, handling the ethical challenges associated with the technology will be crucial to ensuring its responsible and beneficial development. Future data scientists and AI practitioners play a vital role in shaping the

ethical landscape of AI, by actively interacting with these issues, implementing best practices, and contributing to the ongoing dialogue on AI ethics and governance.

Section B: Data Ethics

The British Airways Data Breach of 2018

In September 2018, British Airways (BA) experienced a significant data breach that revealed the personal and financial details of around 429,000 customers and employees (ICO, 2020). This incident serves as a notable case study in data protection failures and the enforcement of the GDPR principles.

Nature of the Breach

The breach entailed unauthorized access to BA's systems, leading to the theft of customer data including names, addresses, login information, payment card numbers, expiry dates, and CVV codes (ICO, 2020). The attackers successfully redirected user traffic from the BA website to a fraudulent site, where customer information was harvested (Palmer, 2020).

This type of attack, known as a supply chain attack or digital skimming, exploits the weaknesses in the web application supply chain to insert malicious code into websites, typically targeting payment pages to steal financial data (Palmer, 2020).

Why and How the Breach Occurred?

The breach was primarily attributed to inadequate security measures and poor data protection practices at British Airways. Specifically:

- **Insufficient Network Segmentation**: BA's systems lacked proper network division, enabling the attackers to navigate across the network after getting initial access (ICO, 2020).
- **Weak Access Controls**: The company failed to implement robust access controls, including multi-factor authentication for critical systems (ICO, 2020).
- **Outdated Software**: Some of BA's systems were running outdated software with known vulnerabilities, which the attackers exploited (Palmer, 2020).
- **Inadequate Monitoring**: BA's monitoring systems failed to detect the unauthorized access and data exfiltration for an extended period (ICO, 2020).
- **Third-Party Vulnerabilities**: The attackers initially gained access through a compromised third-party supplier, highlighting the risks associated with supply chain management (Palmer, 2020).
- **Human Error and Insider Threats**: While not explicitly mentioned in the BA case, human error and insider threats often play a significant role in data breaches. A study by IBM Security (2023) found that 19% of data breaches were caused by human error, and 16% involved insider threats. These factors can include employees falling for phishing scams, mishandling sensitive data, or in some cases, deliberately compromising systems.

The breach occurred over a period of about two weeks, from August 21 to September 5, 2018, before being detected and reported to the authorities (ICO, 2020).

GDPR Guidelines and the Specific Breach

The BA data breach violated several key principles of the GDPR:

- **Data Protection by Design and Default (Article 25)**: BA failed to implement appropriate technological and administrative safeguards to ensure data protection. The lack of network segmentation and weak access controls demonstrate a failure to design systems with data protection as a priority (ICO, 2020).
- **Security of Processing (Article 32)**: The company did not apply proper security safeguards in order to prevent personal data theft, including encryption, ongoing confidentiality, integrity, availability, and resilience of processing systems (GDPR, 2013).
- **Data Breach Notification (Articles 33)**: While BA did notify the ICO and affected individuals within the required 72-hour timeframe, the delay in detecting the breach meant that the notification came weeks after the initial compromise (ICO, 2023).
- **Accountability (Article 5(2))**: BA failed to demonstrate compliance with GDPR principles, particularly in terms of applying proper technical and organizational strategies (General Data Protection Regulation, 2018).

The ICO's investigation found that BA had infringed the GDPR by processing personal data without adequate security measures. This included failures to:

- Limit entry to applications, data, and resources to only those needed
- Undertake rigorous testing on the business' systems
- Protect accounts with multi-factor authentication

These failures led to a significant fine of £20 million, reduced from an initial intention to fine £183 million due to the economic impact of COVID-19 and BA's cooperation with the investigation (ICO, 2020; ICO, 2021)

ICO Guidelines on Fines and Penalties

The ICO's approach to fines and penalties is guided by the GDPR and the UK Data Protection Act 2018. The ICO considers several factors when determining fines, including (ICO, 2024):

- Nature, gravity, and duration of the violation
- Intentional or negligent character of the violation
- Actions taken to reduce the damage suffered by data subjects
- Level of accountability held by the data controller or processor

- Record of prior violations
- Extent of collaboration with the regulatory body
- Types of personal data impacted
- Manner in which the violation became known to the regulatory authority
- Compliance with approved codes of conduct or certification processes

This adjustment of £183 to £20 million highlights the complex interplay between regulatory actions and broader economic factors. While the reduction might be seen as necessary given the unprecedented economic challenges posed by COVID-19, it also raises questions about the consistency and effectiveness of penalties in deterring future breaches. The decision to reduce the fine demonstrates the ICO's attempt to balance punitive measures with the need to avoid exacerbating economic hardships during a global crisis (Warwick-Ching, 2020).

Critical Thoughts on Persistent Data Breaches

Despite the implementation of GDPR and increased awareness of data protection issues, significant data breaches continue to occur. Several factors contribute to this persistent problem:

- **Complexity of Modern IT Environments**: Organizations often struggle to secure increasingly complex and distributed IT infrastructures. Cloud computing, IoT devices, and remote work arrangements have expanded the attack surface, making comprehensive security more challenging (Lallie et al., 2021).
- **Evolving Threat Landscape**: Cybercriminals continuously develop new attack techniques and exploit emerging vulnerabilities. The rapid pace of technological change often outstrips organizations' ability to adapt their security measures (Lallie et al., 2021).
- **Human Error and Insider Threats**: As mentioned earlier, human errors and insider actions, whether malicious or unintentional, continue to be significant factors in data breaches. The shift to remote work during and after the COVID-19 pandemic has further complicated this issue, with employees often working on less secure home networks and using personal devices for work purposes (IBM Security, 2023).
- **Inadequate Investment in Cybersecurity**: Despite the potential costs of a breach, many organizations still underinvest in cybersecurity measures, viewing them as a cost center rather than a critical business function (Choi et al., 2022).
- **Supply Chain Vulnerabilities**: As demonstrated in the BA case, vulnerabilities in third-party suppliers and partners can lead to breaches. Managing supply chain risk is complex and often overlooked (Palmer, 2020).
- **Compliance-Focused Approach**: Some organizations focus on meeting minimum regulatory requirements rather than adopting a comprehensive, risk-based approach to security (Choi et al., 2022).
- **Lack of Board-Level Engagement**: Cybersecurity often fails to receive adequate attention at the board level, leading to insufficient resources and strategic focus (Choi et al., 2022).

- **Legacy Systems**: Many organizations struggle with legacy systems that are difficult to secure and update, creating persistent vulnerabilities (Lallie et al., 2021).
- **Inadequate Incident Response Planning:** Poor preparation for potential breaches can lead to delayed detection and ineffective response, exacerbating the impact of attacks (ICO, 2020).

The BA case highlights several of these issues, particularly the challenges of securing complex IT environments, managing supply chain risks, and maintaining adequate monitoring and incident response capabilities.

To address these persistent challenges, organizations need to implement a more holistic & proactive method for data protection and cybersecurity. This includes:

- Implementing a comprehensive risk management framework that goes beyond mere compliance (NIST, 2023)
- Fostering a culture of security awareness throughout the organization (SANS Institute, 2024)
- Investing in advanced security technologies, including AI and machine learning for threat detection (IBM, 2024)
- Routinely upgrading and patching systems to fix known weaknesses (US-CERT, 2023)
- Establishing robust access measures, such as multi-factor authentication (NIST, 2022)
- Performing frequent security audits and penetration tests (OWASP, 2023)
- Developing and regularly testing incident response plans (NIST, 2022)
- Improving supply chain security through rigorous vendor assessment and monitoring (NIST, 2023)
- Adapting security strategies to address the unique challenges of remote and hybrid work environments (CISA, 2024)

Moreover, regulatory bodies and policymakers have a role to play in encouraging better security practices. While GDPR has raised awareness and imposed significant penalties for non-compliance, there may be a need for more prescriptive guidelines and support, particularly for smaller organizations that may lack resources and expertise (EU GDPR, 2023).

The ICO's response to the BA breach, including the substantial fine and detailed recommendations, sends a clear message about the importance of data protection (ICO, 2020). However, it also highlights the need for ongoing vigilance and improvement in organizational security practices.

As data breaches continue to pose significant risks to individuals and organizations alike, it is crucial that data protection remains a top priority. This requires ongoing investment, innovation, and collaboration between organizations, regulators, and cybersecurity professionals to be ahead of emerging threats and effectively safeguard sensitive personal information (NCSC, 2024).

Section C: AI & Ethical Issues

Industries and Companies Using New Technologies and Data

In an era of swiftly transforming digital environment, various companies are using advanced technologies and data to develop cutting-edge capabilities. Infosys, as a global leader in next-generation digital services and consulting, stands at the forefront of this technological revolution alongside other tech giants like IBM, Accenture, and Tata Consulting Services (TCS). Infosys has developed a comprehensive suite of AI and automation solutions that enable businesses to improve their operations, enhance client satisfaction, and foster innovation. For instance, the company's AI-powered platform, Infosys Nia, helps organizations automate complex business processes, enhance decision-making capabilities, and unlock new revenue streams (Infosys, 2023b). This approach is similar to IBM's Watson platform, which also utilizes AI for business process optimization and decision support (IBM, 2023).

In the realm of data analytics, Infosys has been instrumental in helping clients extract valuable insights from large volumes of data. The company's data analytics solutions leverage sophisticated machine learning algorithms and predictive modeling methods to reveal high-risk patterns and trends, enabling businesses to make data-driven decisions (Infosys, 2023c). This capability is comparable to Accenture's Applied Intelligence platform, which also combines AI with data analytics to drive business value (Accenture, 2023).

Furthermore, Infosys has been actively investing in research and development (R&D) to be at the forefront of emerging technologies. The company's innovation hubs and digital studios across the globe serve as collaborative spaces where clients can co-create solutions and experiment with cutting-edge technologies (Infosys, 2023d). This approach is similar to TCS's Co-Innovation Network (COIN™), which also focuses on collaborative innovation with clients and partners (TCS, 2023).

While Infosys has made significant strides in AI and data analytics, it's important to consider that other companies are also making significant contributions. For example, Google's DeepMind has made groundbreaking advancements in AI research, particularly in areas like reinforcement learning and natural language processing (DeepMind, 2023). Similarly, Amazon Web Services (AWS) has developed an extensive range of cloud-based AI and machine learning services that compete with Infosys's offerings (AWS, 2023).

In the context of my current internship at Infosys, I have the opportunity to work with and learn from these cutting-edge technologies. The company's dedication to ongoing innovation and its commitment to remain at the cutting edge of technological progress provides a rich learning environment. However, it's crucial to keep an eye on the broader industry landscape, including the innovations of other tech giants, to maintain a comprehensive understanding of the field and identify potential areas for collaboration or differentiation.

Infosys' Perspective on AI for Organizational Decision-Making

Infosys views AI as a transformative force in organizational decision-making, capable of augmenting human intelligence and driving business value. The company's approach to AI-driven decision-making is centered around three key principles: augmentation, automation, and amplification (Infosys, 2023e).

Augmentation entails leveraging AI to improve human decision-making capabilities through providing real-time insights and recommendations. Automation focuses on streamlining routine decision-making techniques, liberating human resources for more important tasks. Amplification leverages AI to scale decision-making capabilities across the organization, enabling faster and more consistent decision-making at all levels (Infosys, 2023e).

Infosys has developed AI-powered solutions that cater to various aspects of organizational decision-making. For example, the company's AI-driven supply chain management solution helps businesses optimize inventory levels, predict demand, and improve logistics efficiency (Infosys, 2023f). In the financial services sector, Infosys has implemented AI-powered risk management solutions that enable banks and insurance companies to make more informed lending and underwriting decisions (Infosys, 2023g).

Handling AI Ethical Issues at Infosys

Infosys recognizes the significance of addressing ethical concerns in AI development and deployment. The company has established a comprehensive framework for responsible AI that encompasses principles such as fairness, transparency, accountability, and privacy (Infosys, 2023h).

To ensure ethical AI practices, Infosys has implemented several measures:

1. **Ethics Review Board**: An internal committee that reviews AI projects for potential ethical implications and provides guidance on mitigating risks (Infosys, 2023h).
2. **AI Ethics Training**: Mandatory training sessions for employees engaged in AI development to increase awareness of ethical standards and best practices (Infosys, 2023i).
3. **Identifying and Reducing Bias**: Implementation of tools & techniques to identify and address biases in AI algorithms and datasets (Infosys, 2023j).
4. **Explainable AI**: Development of AI models that provide transparent explanations for their decisions, enhancing accountability and trust (Infosys, 2023k).

As an intern at Infosys, my role in addressing AI ethical issues involves participating in ethics training programs, adhering to the company's responsible AI guidelines, and raising awareness about potential ethical concerns within my team.

Harnessing the Power of AI Solutions at Infosys

Infosys has leveraged AI solutions across various domains to drive innovation and create value for its clients. Some notable examples include:

1. **Infosys Nia**: An AI-powered platform that automates complex business processes, enhances decision-making capabilities, and unlocks new revenue streams for clients across industries (Infosys, 2023b).
2. **AI-Driven Customer Experience**: Implementation of chatbots and virtual assistants to improve customer service and engagement for clients in retail, banking, and telecommunications sectors (Infosys, 2023l).
3. **Predictive Maintenance**: Development of AI-powered solutions that predict equipment failures and optimize maintenance schedules for manufacturing and energy clients (Infosys, 2023m).
4. **Fraud Detection**: Implementation of AI algorithms to detect and prevent fraudulent activities in financial transactions for banking and insurance clients (Infosys, 2023n).
5. **Healthcare Analytics**: Deployment of AI-powered solutions to analyze medical imaging data, assist in diagnosis, and improve patient outcomes for healthcare providers (Infosys, 2023o).

Data Privacy, Information Governance, and Legislative Standards

In the UK, data privacy and information governance are controlled by several key regulations and standards, comprising the GDPR, the Data Protection Act 2018, and industry-specific regulations (Information Commissioner's Office, 2023a).

As an intern at Infosys, I recognize the importance of adhering to these regulations and maintaining high standards of data protection and security. Infosys has implemented a comprehensive data privacy and security framework that aligns with UK and global standards (Infosys, 2023p).

Key aspects of Infosys' approach to data privacy and information governance include:

1. **Data Protection Impact Assessments (DPIAs)**: Conducting thorough assessments to identify and mitigate privacy risks associated with new projects or technologies (Information Commissioner's Office, 2023b).
2. **Privacy by Design**: Integrating privacy considerations into the initial design and built of products and services (Infosys, 2023q).
3. **Data Minimization**: Collecting and processing only the necessary data required for specific purposes, in line with GDPR principles (Information Commissioner's Office, 2023c).
4. **Consent Management**: Implementing robust techniques to attain and manage user consent for data processing tasks (Infosys, 2023r).

5. **Data Security Measures**: Employing encryption, access controls, and other security measures to protect sensitive data from unauthorized access or breaches (Infosys, 2023s).
6. **Employee Training**: Providing regular training and awareness programs to ensure all employees understand their responsibilities regarding data protection and privacy (Infosys, 2023t).
7. **Third-Party Risk Management**: Assessing and monitoring the data protection practices of vendors and partners to ensure compliance with relevant standards (Infosys, 2023u).

In my role as an intern, I am required to complete mandatory data protection training and adhere to the company's data privacy and security policies. I also actively participate in discussions and initiatives aimed at enhancing our data protection practices and ensuring compliance with UK regulations.

Section D: Development and Implementation of an Ethical AI Framework

Framework Proposal and Rationale

As a Data and AI Intern at Infosys, I propose developing an ethical framework for AI use which relates to our company's values and addresses the unique challenges in our industry. This framework will be inspired by established guidelines like Google's Responsible AI practices, but tailored to Infosys' specific needs and context (Google AI, 2023).

The rationale for this framework is to ensure that Infosys remains at the forefront of ethical AI development and deployment. By establishing clear guidelines, we can mitigate risks, build trust with stakeholders, and create AI platforms that correspond with our organizational values and community standards (Floridi et al., 2020).

Key Principles, Guidelines, and Practices

The proposed ethical framework for AI at Infosys consists of the following key principles:

- **Human-Centered Design**: AI platforms should be built to enhance human capabilities, not replace them. This approach ensures that AI solutions prioritize human well-being and societal benefits (Shneiderman, 2020).
- **Fairness and Non-Discrimination**: AI platforms should be designed and utilized in a way that prevents any unjust bias against people based on protected traits. (Mehrabi et al., 2021).
- **Transparency and Explainability**: AI decision-making processes must be transparent and explainable to stakeholders, including employees, clients, and end-users (Arrieta et al., 2020).
- **Privacy and Data Protection**: AI systems must respect user privacy and comply with data protection regulations, ensuring secure handling of personal and sensitive information (Tsamados et al., 2022).
- **Accountability and Governance**: Clear accountability structures and governance mechanisms should be established for AI development and deployment (Dignum, 2019).
- **Safety and Reliability**: AI systems must be rigorously tested and monitored to ensure their safety and reliability in various operational contexts (Amodei et al., 2016).

Guidelines:

- Perform comprehensive ethical impact evaluations prior to the installation of AI systems (Wright and Veale, 2017)
- Ensure diverse and inclusive teams in AI development (West et al., 2019)
- Implement robust testing and validation processes (Brundage et al., 2020)

- Maintain clear documentation of AI systems and decision-making processes (Raji et al., 2020)
- Establish mechanisms for regular monitoring and auditing of AI systems (Koshiyama et al., 2021)
- Provide clear channels for stakeholder feedback and redress (Floridi et al., 2018)

Practices:

- Regular ethics training for all employees included in AI system built and deployment (Morley et al., 2021)
- Establishment of an AI ethics review board (Cath et al., 2018)
- Implementation of explainable AI techniques (Adadi and Berrada, 2018)
- Regular bias testing and mitigation (Holstein et al., 2019)
- Strict data governance and protection measures (Schneider et al., 2020)
- Collaboration with external experts and stakeholders for ethical guidance (Fjeld et al., 2020)
- Transparent reporting on AI ethics metrics and incidents (Raji et al., 2020)

Alignment with Organizational Values

The proposed ethical framework relates closely with Infosys' core values of client value, leadership by example, integrity, transparency, and fairness (Infosys, 2023a). By prioritizing fairness and non-discrimination, Infosys can ensure that their AI systems treat all stakeholders equitably. The emphasis on transparency and explainability aligns with their commitment to integrity and transparency, fostering trust with their clients and the public. The company's focus on human-centered AI and accountability reflects their value of leadership by example, positioning Infosys as a responsible leader in the AI industry. The framework's attention to privacy, safety, and environmental sustainability demonstrates their commitment to creating long-term value for its clients and society at large (Jobin, Ienca and Vayena, 2019). This framework addresses specific ethical concerns by providing clear guidelines for issues such as bias mitigation, data protection, and algorithmic transparency. It also establishes mechanisms for ongoing monitoring and improvement, ensuring that Infosys can adapt to emerging ethical challenges in the rapidly evolving field of AI (Hagendorff, 2020).

Implementation and Integration Plan

Implementation Strategy:

1. **Leadership Endorsement**: Secure buy-in from top management to ensure organization-wide support for the ethical AI framework (Floridi et al., 2018).
2. **Cross-Functional Ethics Committee**: Establish a dedicated committee comprising representatives from various departments to oversee the implementation and ongoing management of the ethical AI framework (Jobin et al., 2019).

3. **Policy Development**: Create detailed policies and guidelines based on the framework's principles, tailored to different AI applications and use cases within Infosys (Hagendorff, 2020).
4. **Ethics Review Process**: Implement a mandatory ethics review process for all AI projects, ensuring alignment with the framework before development or deployment (Morley et al., 2021).
5. **Stakeholder Engagement**: Engage with clients, partners, and end-users to gather feedback on the ethical ramifications of AI solutions and refine the framework accordingly (Fjeld et al., 2020).

Integration into Existing and Future AI Initiatives:

1. **Project Lifecycle Integration**: Incorporate ethical considerations at every stage of AI project lifecycles, from ideation to deployment and maintenance (Morley et al., 2020).
2. **Ethical Impact Assessments**: Conduct thorough ethical impact assessments for all AI initiatives, identifying potential risks and mitigation strategies (Wright and Veale, 2017).
   Ethical KPIs: Develop and track key performance indicators (KPIs) related to ethical AI practices, integrating them into project evaluations and team performance metrics (Askell et al., 2019).
3. **Client Education**: Develop resources and courses to educate clients on the significance of ethical AI and how Infosys' framework ensures responsible AI solutions (Floridi et al., 2018).
4. **Ethical AI Certification**: Create an internal certification program for AI practitioners who demonstrate proficiency in applying the ethical framework (Jobin et al., 2019).
5. **Open-Source Contributions**: Encourage teams to contribute to open-source AI ethics tools and frameworks, fostering a culture of responsible AI development beyond Infosys (Hagendorff, 2020).

Role of Training, Monitoring, and Continuous Improvement

Training will be essential for the successful execution of the ethical framework. We will:

- Develop a comprehensive AI ethics curriculum for all employees involved in the built and deployment of AI systems (Morley et al., 2021).
- Conduct regular workshops and seminars on emerging ethical issues in AI (Burton et al., 2017).
- Integrate ethical AI considerations into onboarding processes for new employees (Rakova et al., 2021).

Monitoring will be an ongoing process to ensure adherence to the ethical framework:

- Implement automated tools for bias detection and fairness auditing in AI systems (Bellamy et al., 2019).

- Conduct regular internal audits of AI projects for ethical compliance (Raji et al., 2020).
- Establish a reporting system for ethical concerns or incidents (Madaio et al., 2020).

Continuous improvement will be achieved through:

- Regular review and update of the ethical framework based on internal feedback and external developments (Mittelstadt, 2019).
- Collaboration with academic institutions and industry partners on AI ethics research (Jobin et al., 2019).
- Participation in global AI ethics initiatives and standards development (Fjeld et al., 2020).

By implementing this comprehensive ethical framework, Infosys can ensure responsible AI development and deployment, aligning with its organizational values and setting a standard for ethical AI practices in the industry.

References

1. Accenture, 2023. Applied Intelligence. [online] Available at: https://www.accenture.com/us-en/services/applied-intelligence-index [Accessed 5 July 2024].
2. Adadi, A. and Berrada, M., 2018. Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI). IEEE Access, 6, pp.52138-52160.
3. Almasri, F., 2024. Exploring the Impact of Artificial Intelligence in Teaching and Learning of Science: A Systematic Review of Empirical Research. *Research in Science Education*. Available at: https://doi.org/10.1007/s11165-024-10176-3 [Accessed 7 July 2024].
4. Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J. and Mané, D., 2016. Concrete problems in AI safety. arXiv preprint arXiv:1606.06565.
5. Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R. and Chatila, R., 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, pp.82-115.
6. Askell, A., Brundage, M. and Hadfield, G., 2019. The role of cooperation in responsible AI development. arXiv preprint arXiv:1907.04534.
7. AWS, 2023. Machine Learning on AWS. [online] Available at: https://aws.amazon.com/machine-learning/ [Accessed 5 July 2024].
8. Bank of England (2022). *Machine learning in UK financial services*. Available at: https://www.bankofengland.co.uk/report/2022/machine-learning-in-uk-financial-services [Accessed: 8 July 2024].
9. Bellamy, R.K., Dey, K., Hind, M., Hoffman, S.C., Houde, S., Kannan, K., Lohia, P., Martino, J., Mehta, S., Mojsilović, A. and Nagar, S., 2019. AI Fairness 360: An extensible toolkit for detecting and mitigating algorithmic bias. IBM Journal of Research and Development, 63(4/5), pp.4-1.
10. Brundage, M., Avin, S., Wang, J., Belfield, H., Krueger, G., Hadfield, G., Khlaaf, H., Yang, J., Toner, H., Fong, R. and Maharaj, T., 2020. Toward trustworthy AI development: mechanisms for supporting verifiable claims. arXiv preprint arXiv:2004.07213.
11. Burton, E., Goldsmith, J., Koenig, S., Kuipers, B., Mattei, N. and Walsh, T., 2017. Ethical considerations in artificial intelligence courses. AI magazine, 38(2), pp.22-34.
12. Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M. and Floridi, L., 2018. Artificial intelligence and the 'good society': the US, EU, and UK approach. Science and engineering ethics, 24(2), pp.505-528.
13. Centre for Data Ethics and Innovation (CDEI), 2023. AI Ethics and Safety. [Online] Available at: https://www.gov.uk/government/organisations/centre-for-data-ethics-and-innovation [Accessed 30 June 2024].
14. Choi, J.Y., Shin, J. and Choi, J., 2022. Cybersecurity management in the era of digital transformation: An empirical study of Korean small and medium enterprises. Sustainability, 14(4), p.2406.
15. Cybersecurity and Infrastructure Security Agency (CISA), 2024. Telework Guidance and Resources. Available at: https://www.cisa.gov/telework [Accessed 7 July 2024].

16. DeepMind, 2023. Research. [online] Available at: https://deepmind.com/research [Accessed 5 July 2024].
17. Department for Science, Innovation and Technology, 2023. AI Regulation: A Pro-Innovation Approach. [Online] Available at: https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach [Accessed 30 June 2024].
18. Department for Science, Innovation and Technology, Office for Artificial Intelligence and Centre for Data Ethics and Innovation, 2019. *Understanding artificial intelligence ethics and safety*. [online] GOV.UK. Available at: https://www.gov.uk/guidance/understanding-artificial-intelligence-ethics-and-safety [Accessed 8 July 2024].
19. Dignum, V., 2019. Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way. Cham: Springer International Publishing.
20. European Union, 2023. General Data Protection Regulation (GDPR). Available at: https://gdpr.eu/ [Accessed 7 July 2024].
21. Fjeld, J., Achten, N., Hilligoss, H., Nagy, A. and Srikumar, M., 2020. Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI. Berkman Klein Center Research Publication, (2020-1).
22. Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F. and Schafer, B., 2018. AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. Minds and Machines, 28(4), pp.689-707.
23. Floridi, L., Cowls, J., King, T.C. and Taddeo, M., 2020. How to Design AI for Social Good: Seven Essential Factors. Science and Engineering Ethics, 26(3), pp.1771-1796.
24. Foreign, Commonwealth & Development Office and The Rt Hon James Cleverly (2023) *UK unites with global partners to accelerate development using AI*. Available at: https://www.gov.uk/government/news/uk-unites-with-global-partners-to-accelerate-development-using-ai [Accessed: 8 July 2024].
25. GDPR (2013). *Art. 32 GDPR – Security of processing | General Data Protection Regulation (GDPR)*. [online] General Data Protection Regulation (GDPR). Available at: https://gdpr-info.eu/art-32-gdpr/.
26. General Data Protection Regulation (2018). *Art. 5 GDPR – Principles relating to processing of personal data | General Data Protection Regulation (GDPR)*. [online] General Data Protection Regulation (GDPR). Available at: https://gdpr-info.eu/art-5-gdpr/.
27. Google AI, 2023. Responsible AI Practices. [online] Available at: https://ai.google/responsibilities/responsible-ai-practices/ [Accessed 6 July 2023].
28. Hagendorff, T., 2020. The ethics of AI ethics: An evaluation of guidelines. Minds and Machines, 30(1), pp.99-120.
29. Holstein, K., Wortman Vaughan, J., Daumé III, H., Dudik, M. and Wallach, H., 2019. Improving fairness in machine learning systems: What do industry practitioners need?. In Proceedings of the 2019 CHI conference on human factors in computing systems (pp. 1-16).
30. IBM, 2024. Artificial Intelligence (AI) cybersecurity. Available at: https://www.ibm.com/security/artificial-intelligence [Accessed 7 July 2024].

31. IBM, 2023. IBM Watson. [online] Available at: https://www.ibm.com/watson [Accessed 5 July 2024].
32. IBM Security, 2023. Cost of a Data Breach Report 2023. [online] IBM. Available at: https://www.ibm.com/reports/data-breach [Accessed 30 June 2024].
33. ICO, 2024. *Data Protection Fining Guidance*. [pdf] ICO. Available at: https://ico.org.uk/about-the-ico/our-information/policies-and-procedures/data-protection-fining-guidance/?template=pdf&patch=28 [Accessed 7 July 2024].
34. ICO, 2021. ICO fines British Airways £20m for data breach affecting more than 400,000 customers. [online] GDPR Register. Available at: https://www.gdprregister.eu/news/british-airways-fine/#:~:text=The%20Information%20Commissioner's%20Office%20(ICO,adequate%20security%20measures%20in%20place. [Accessed 7 July 2024].
35. ICO (2023). *Personal Data breaches: a Guide*. [online] ico.org.uk. Available at: https://ico.org.uk/for-organisations/report-a-breach/personal-data-breach/personal-data-breaches-a-guide/.
36. Infosys, 2023a. Digital Innovation. [online] Available at: https://www.infosys.com/services/digital-innovation.html [Accessed 5 July 2024].
37. Infosys, 2023b. Infosys Nia. [online] Available at: https://www.infosys.com/products-and-platforms/nia.html [Accessed 5 July 2024].
38. Infosys, 2023c. Data and Analytics. [online] Available at: https://www.infosys.com/services/data-analytics.html [Accessed 5 July 2024].
39. Infosys, 2023d. Innovation Hubs. [online] Available at: https://www.infosys.com/about/innovation-hubs.html [Accessed 5 July 2024].
40. Infosys, 2023e. AI-Driven Decision Making. [online] Available at: https://www.infosys.com/insights/ai-automation/ai-driven-decision-making.html [Accessed 5 July 2024].
41. Infosys, 2023f. AI in Supply Chain Management. [online] Available at: https://www.infosys.com/industries/retail/insights/ai-supply-chain-management.html [Accessed 5 July 2024].
42. Infosys, 2023g. AI in Financial Services. [online] Available at: https://www.infosys.com/industries/financial-services/insights/ai-financial-services.html [Accessed 5 July 2024].
43. Infosys, 2023h. Responsible AI Framework. [online] Available at: https://www.infosys.com/about/corporate-responsibility/responsible-ai.html [Accessed 5 July 2024].
44. Infosys, 2023i. AI Ethics Training. [online] Available at: https://www.infosys.com/careers/learning-development/ai-ethics-training.html [Accessed 5 July 2024].
45. Infosys, 2023j. Bias Detection in AI. [online] Available at: https://www.infosys.com/insights/ai-automation/bias-detection-ai.html [Accessed 5 July 2024].
46. Infosys, 2023k. Explainable AI. [online] Available at: https://www.infosys.com/insights/ai-automation/explainable-ai.html [Accessed 5 July 2024].

47. Infosys, 2023l. AI-Driven Customer Experience. [online] Available at: https://www.infosys.com/services/experience-design/offerings/ai-driven-customer-experience.html [Accessed 5 July 2024].
48. Infosys, 2023m. Predictive Maintenance. [online] Available at: https://www.infosys.com/industries/manufacturing/insights/predictive-maintenance.html [Accessed 5 July 2024].
49. Infosys, 2023n. AI in Fraud Detection. [online] Available at: https://www.infosys.com/industries/financial-services/insights/ai-fraud-detection.html [Accessed 5 July 2024].
50. Infosys, 2023o. Healthcare Analytics. [online] Available at: https://www.infosys.com/industries/healthcare/insights/healthcare-analytics.html [Accessed 5 July 2024].
51. Infosys, 2023p. Data Privacy and Security Framework. [online] Available at: https://www.infosys.com/about/corporate-responsibility/data-privacy-security.html [Accessed 5 July 2024].
52. Infosys, 2023q. Privacy by Design. [online] Available at: https://www.infosys.com/services/cyber-security/insights/privacy-by-design.html [Accessed 5 July 2024].
53. Infosys, 2023r. Consent Management. [online] Available at: https://www.infosys.com/services/data-analytics/insights/consent-management.html [Accessed 5 July 2024].
54. Infosys, 2023s. Data Security Measures. [online] Available at: https://www.infosys.com/services/cyber-security/offerings/data-security.html [Accessed 5 July 2024].
55. Infosys, 2023t. Employee Data Protection Training. [online] Available at: https://www.infosys.com/careers/learning-development/data-protection-training.html [Accessed 5 July 2024].
56. Infosys, 2023u. Third-Party Risk Management. [online] Available at: https://www.infosys.com/services/risk-compliance/offerings/third-party-risk-management.html [Accessed 5 July 2024].
57. Jobin, A., Ienca, M. and Vayena, E., 2019. The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), pp.389-399.
58. Koshiyama, A., Kazim, E., Treleaven, P., Rai, P., Szpruch, L., Pavey, G., Ahamat, G., Leutner, F., Goebel, R., Knight, A. and Adams, J., 2021. Towards algorithm auditing: A survey on managing legal, ethical and technological risks of AI, ML and associated algorithms. Available at SSRN 3778998.
59. Lallie, H.S., Shepherd, L.A., Nurse, J.R., Erola, A., Epiphaniou, G., Maple, C. and Bellekens, X., 2021. Cyber security in the age of COVID-19: A timeline and analysis of cyber-crime and cyber-attacks during the pandemic. Computers & Security, 105, p.102248.
60. Madaio, M.A., Stark, L., Wortman Vaughan, J. and Wallach, H., 2020. Co-designing checklists to understand organizational challenges and opportunities around fairness in AI. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1-14).

61. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K. and Galstyan, A., 2021. A survey on bias and fairness in machine learning. ACM Computing Surveys (CSUR), 54(6), pp.1-35.
62. Mittelstadt, B., 2019. Principles alone cannot guarantee ethical AI. Nature Machine Intelligence, 1(11), pp.501-507.
63. Morley, J., Floridi, L., Kinsey, L. and Elhalal, A., 2020. From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. Science and Engineering Ethics, 26(4), pp.2141-2168.
64. Morley, J., Elhalal, A., Garcia, F., Kinsey, L., Mökander, J. and Floridi, L., 2021. Ethics as a service: a pragmatic operationalisation of AI ethics. Minds and Machines, 31(2), pp.239-256.
65. National Cyber Security Centre (NCSC), 2024. What we do. Available at: https://www.ncsc.gov.uk/section/about-ncsc/what-we-do [Accessed 7 July 2024].
66. National Institute of Standards and Technology (NIST), 2022b. Computer Security Incident Handling Guide. Available at: https://csrc.nist.gov/publications/detail/sp/800-61/rev-2/final [Accessed 7 July 2024].
67. National Institute of Standards and Technology (NIST), 2023b. Cybersecurity Supply Chain Risk Management. Available at: https://csrc.nist.gov/projects/cybersecurity-supply-chain-risk-management [Accessed 7 July 2024].
68. National Institute of Standards and Technology (NIST), 2022a. Digital Identity Guidelines. Available at: https://pages.nist.gov/800-63-3/ [Accessed 7 July 2024].
69. National Institute of Standards and Technology (NIST), 2023a. Risk Management Framework. Available at: https://csrc.nist.gov/projects/risk-management/about-rmf [Accessed 7 July 2024].
70. Open Web Application Security Project (OWASP), 2023. Web Security Testing Guide. Available at: https://owasp.org/www-project-web-security-testing-guide/ [Accessed 7 July 2024].
71. ICO, 2020. Penalty Notice: BA. [pdf] Available at: https://ico.org.uk/media/action-weve-taken/mpns/2618421/ba-penalty-20201016.pdf [Accessed 7 July 2024].
72. Palmer, D., 2020. Data watchdog issues biggest ever fine over airline cyberattack. *ZDNet*. [online] Available at: https://www.zdnet.com/article/data-watchdog-issues-biggest-ever-fine-over-airline-cyberattack/ [Accessed 8 July 2024].
73. PwC UK (2024) *Automation will impact around 30% of UK jobs by mid 2030s - but which ones?*. Available at: https://www.pwc.co.uk/press-room/press-releases/regions/northern-ireland/automation-impact.html (Accessed: 8 July 2024).
74. Queen Mary University of London. (2024). *Data Science and Artificial Intelligence Conversion Programme MSc*. Queen Mary University of London. Available at: https://www.qmul.ac.uk/postgraduate/taught/coursefinder/courses/data-science-and-artificial-intelligence-conversion-programme-msc/ [Accessed 7 July 2024].
75. Rakova, B., Yang, J., Cramer, H. and Chowdhury, R., 2021. Where responsible AI meets reality: Practitioner perspectives on enablers for shifting organizational practices. Proceedings of the ACM on Human-Computer Interaction, 5(CSCW1), pp.1-23.
76. Raji, I.D., Smart, A., White, R.N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D. and Barnes, P., 2020. Closing the AI accountability gap: Defining an end-to-

end framework for internal algorithmic auditing. In Proceedings of the 2020 conference on fairness, accountability, and transparency (pp. 33-44).
77. Royal Society, 2023. Explainable AI: the basics. [Online] Available at: https://royalsociety.org/topics-policy/projects/explainable-ai/ [Accessed 30 June 2024].
78. Royal United Services Institute (RUSI), 2023. AI-Powered Surveillance: Balancing Security and Privacy. [Online] Available at: https://rusi.org/explore-our-research/publications/ai-powered-surveillance [Accessed 30 June 2024].
79. SANS Institute, 2024. Security Awareness Training. Available at: https://www.sans.org/security-awareness-training/ [Accessed 7 July 2024].
80. Schneider, J., Handali, J. and vom Brocke, J., 2020. Increasing trust in AI-based, automated decision-making in the public sector: framing algorithmic transparency and explainability. In Proceedings of the 53rd Hawaii International Conference on System Sciences.
81. Shneiderman, B., 2020. Human-centered artificial intelligence: Reliable, safe & trustworthy. International Journal of Human–Computer Interaction, 36(6), pp.495-504.
82. TCS, 2023. Co-Innovation Network (COIN™). [online] Available at: https://www.tcs.com/what-we-do/research/co-innovation-network-coin [Accessed 5 July 2024].
83. Tech Nation, 2023. UK Tech Report 2023. [Online] Available at: https://technation.io/insights/uk-tech-report-2023/ [Accessed 30 June 2024].
84. The Alan Turing Institute, 2023. AI Ethics and Society. [Online] Available at: https://www.turing.ac.uk/research/interest-groups/ai-ethics-and-society [Accessed 30 June 2024].
85. Tsamados, A., Aggarwal, N., Cowls, J., Morley, J., Roberts, H., Taddeo, M. and Floridi, L., 2022. The ethics of algorithms: key problems and solutions. AI & SOCIETY, 37(1), pp.215-230.
86. UK AI Council, 2023. AI Roadmap. [Online] Available at: https://www.gov.uk/government/publications/ai-roadmap [Accessed 30 June 2024].
87. University of Oxford Institute for Ethics in AI, 2023. Interdisciplinary AI Ethics Education. [Online] Available at: https://www.oxford-aiethics.ox.ac.uk/education [Accessed 30 June 2024].
88. US Computer Emergency Readiness Team (US-CERT), 2023. Patch Management. Available at: https://www.cisa.gov/uscert/bsi/articles/best-practices/patch-management/patch-management-security-practice [Accessed 7 July 2024].
89. Warwick-Ching, L., 2020. British Airways fined £20m for data breach. Financial Times, [online] Available at: https://www.ft.com/content/3e2ac8fb-c8c5-4c84-b8b7-5e3f9a7b0f38 [Accessed 30 June 2024].
90. West, S.M., Whittaker, M. and Crawford, K., 2019. Discriminating systems: Gender, race and power in AI. AI Now Institute, 19.
91. Wright, D. and Veale, M., 2017. Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for. Duke L. & Tech. Rev., 16, p.18.