Because of our increasing reliance on digital services and worries about data privacy, analyzing privacy issues in technological businesses such as Microsoft 365, Google, and IBM is critical. The European Union's General Data Protection Regulation (GDPR) has a substantial impact on how these businesses manage personal data, mandating an awareness of their compliance[1]. The National Institute of Standards and Technology (NIST) publishes AI guidelines centered on ethics and dependability, making it critical to evaluate these companies' adherence to these standards, particularly for AI-powered goods[2]. Privacy Impact Assessments (PIAs) are critical for identifying and minimizing privacy risks in new initiatives, since they provide insight into how tech businesses deal with potential privacy issues[3].

The increased public desire for transparency in data collecting, use, and protection underscores the importance of this research in assessing these companies' transparency and trustworthiness. Understanding how firms negotiate current privacy restrictions helps assess their preparation for future difficulties as AI and machine learning improve. Given the global nature of data, a global view on privacy is required, taking into account varied legal and cultural standards. Examining the privacy practices of these tech behemoths can provide best practices for other firms, improving their privacy strategies. This analysis is critical for understanding these organizations' commitment to data security, ethical AI use, and readiness for the growing digital landscape, as well as for contributing to industry-wide privacy benchmarks and best practices[4].

1. GDPR (General Data Protection Regulation):

The General Data Protection Regulation (GDPR) in the EU imposes strict rules on personal data processing, emphasizing Data Minimization and Purpose Limitation to collect data only when necessary and for stated purposes. It mandates express consent for data processing and transparency in data usage, storage, and sharing. GDPR empowers individuals with rights like data access, rectification, erasure, and portability. It requires Data Protection Impact Assessments (DPIA) for high-risk activities and a swift breach notification system to mitigate risks promptly[5].

2. NIST AI (National Institute of Standards and Technology - Artificial Intelligence):

The National Institute of Standards and Technology (NIST) emphasizes responsible AI development, focusing on mitigating privacy issues through its AI Risk Management Framework. This framework offers a structured approach for identifying and managing risks, including privacy concerns, in AI systems[6]. NIST advocates for the ethical development of AI, integrating privacy considerations from the onset and providing guidelines and best practices. These efforts ensure AI technologies are developed and deployed with a strong focus on privacy and ethical standards[7].

3. Privacy Impact Assessment (PIA):

A Privacy Impact Assessment (PIA) is a critical process for businesses to identify and mitigate privacy risks in their projects. It involves a detailed analysis of how personal information is collected, used, stored, and shared[8]. This process helps in spotting potential privacy issues and compliance gaps, enabling the development of strategies to manage these risks. PIAs emphasize minimizing privacy impacts and include stakeholder involvement, often through public consultations, making them essential tools in privacy risk management[9].

Tools:

Microsoft copilot:

Microsoft Copilot, potentially developed using large language models like OpenAI's GPT, assists in coding and document creation by emulating human-like text. It's trained with diverse data from books, websites, and coding repositories like GitHub. Microsoft enforces privacy policies to handle personal data, with data protection varying by the tool's application and user interaction. Measures may include data anonymization and security, although the degree of personal data usage depends on user engagement[10].

Google:

Google employs a range of artificial intelligence and machine learning technologies across its various services, including its renowned search engine, voice assistants, and other products. These technologies are powered by algorithms designed to process and understand vast datasets, enabling Google to deliver relevant results and services efficiently. A significant aspect of Google's operation involves the collection of extensive data, which encompasses search queries, website visits (tracked through Google Analytics), and location information (when location services are enabled). While Google asserts that it implements various measures to safeguard user privacy, such as anonymizing and encrypting data, the company has faced criticism and scrutiny over its extensive data collection practices and the potential risks they pose to user privacy[11].

IBM:

IBM's AI solutions, such as Watson, are employed across multiple industries for data analysis, natural language processing, and pattern recognition. Data is gathered from a variety of sources, including business-specific and public data, to train AI models. IBM prioritizes data security and user privacy, especially for its enterprise clients, adhering to strict data protection regulations, although specifics vary by service[12].

Data Life cycle for Microsoft copilot, Google, IBM:

Data life cycle under GDPR:

AI systems such as Microsoft's Copilot, Google's AI, and IBM's AI solutions must follow strict data management rules under the GDPR framework. Data collection is controlled by legal, fair, and transparent norms, which necessitate explicit communication and agreement from data subjects. Data must be processed only for particular, valid objectives, ensuring accuracy and limiting the scope of processing. Data storage entails secure processing and stringent retention limitations, with mechanisms for data subjects to request deletion under the 'right to be forgotten'. Data sharing, particularly across borders, necessitates proper safeguards, whereas data security necessitates powerful risk-aversion techniques. Finally, secure destruction or anonymization of end-of-life data is required, and organizations must undertake regular audits and report any data breaches to authorities[13].

Data life cycle under PIA:

The PIA framework, on the other hand, focuses on examining the impact of data practices on privacy. It starts with assessing how personal data is acquired and verifying that it is required for the intended purpose. To minimize privacy threats, the framework examines the implications of data processing, storage procedures, retention policies, and management. It also examines data sharing procedures to ensure they are in accordance with privacy rules and legal obligations. Security audits are essential for protecting data from unauthorized access and other dangers. To identify new threats and assure continued compliance with privacy rules, the PIA framework emphasizes constant monitoring and periodic reassessment.

Data life cycle under NIST AI:

Under the NIST AI Framework, the data life cycle for AI systems such as Microsoft Copilot, Google AI, and IBM AI includes numerous critical stages. It begins with data collected from various sources, which is then processed and prepared for AI application. This data is then efficiently saved and maintained. The following stage comprises data analysis and modeling utilizing AI algorithms to ensure that insights are derived accurately. Following that, these models are rigorously tested and validated to ensure performance and reliability. Once validated, these models are deployed and incorporated into a variety of systems, necessitating constant monitoring and maintenance to assure their continued usefulness. Furthermore, a heavy emphasis is maintained throughout the cycle on data governance and compliance to conform to legal and ethical norms, a vital part of the NIST AI Framework in managing AI technology ethically and efficiently.

AI Governances strategies:

AI governance policies for technology companies such as Microsoft copilot, Google, and IBM are focused on ensuring that their AI systems are developed and deployed in a responsible and ethical manner. Here's a rundown of some of the primary tactics they employ:

AI governances for Microsoft (Copilot):

For the developments of all AI technology , including the Copilot program, Microsoft has built a comprehensive ethical AI framework. Fairness, dependability, safety, privacy, security, inclusion, openness, and responsibility are among the concepts behind this paradigm. Microsoft has established a special AI Ethics Committee to oversee AI advances in order to assure adherence to these standards. Furthermore, the company highlights the significance of user input and control, allowing users to influence how AI systems such as Copilot interact with their data and tasks. Microsoft is committed to ensuring that people have clear information about how these systems make judgments and process data in its AI activities. Furthermore, Microsoft actively promotes and advances ethical AI practices through partnerships and collaborations with academic and industrial partners.

AI governances for Google:

Google has outlined a set of AI principles aimed at ensuring that their technology provides social benefit, safety, privacy, scientific difficulty, and responsibility. To fulfill these criteria, they created an independent AI Ethics Board that is in charge of analyzing AI initiatives and assessing their ethical implications and societal impact. Google also promotes open-source initiatives as a means of increasing openness and involving the broader community in AI development. Recognizing the need of ongoing learning and progress in the field, Google is actively involved in developing and financing research in crucial areas of AI ethics such as machine learning fairness, interpretability, and privacy.Furthermore, the company focuses a heavy emphasis on staff training and raising awareness about ethical AI development and usage, ensuring that their workforce is well-equipped to handle the complexity of AI responsibly.

AI governances for IBM:

IBM emphasizes trust and transparency in AI, emphasizing the necessity of explainability in AI decision-making processes. This focus is highlighted further by their creation of the AI justice 360 Toolkit, a complete resource designed to detect and eliminate bias in AI models, emphasizing their commitment to AI justice. IBM has built a systematic governance framework that guides its operations in this field to ensure ethical AI development and implementation. Furthermore, IBM understands the need of

collaboration with external organizations in creating responsible AI governance and is actively involved in this endeavor with policymakers and regulatory agencies. IBM has established an AI Ethics Board to reinforce their commitment to ethical AI. This internal organization is responsible for examining AI initiatives to ensure they adhere to IBM's ethical rules and procedures.

Test methodology:

The initial stage of our process is to collect the frameworks that will be used. We discovered a PIA framework, a GDPR template, and NIST's AI framework to help us with our investigation. We shall use our frameworks to gain insight into the policies of this organization. We have entered the information of the three frameworks into a spreadsheet in order to compare them across our three firms. Following that, we will scan all three sources and enter any relevant information into the spreadsheet. This summary will be used to identify common and/or contrasting tendencies across the organizations, which we will highlight in our final report. With our summarized observations, we will recommend activities for any organization in this field to do in order to comply with privacy guidelines.

Research Motivation:

The issue stated in our topic "Privacy Impact Assessments of Technology AI Companies" focuses on the critical analysis of privacy issues within major technological companies such as Microsoft Copilot, Google, and IBM, in light of increasing reliance on digital services and data privacy concerns. The document underlines the necessity of understanding how these corporations manage personal data, particularly in light of the European Union's General Data Protection Regulation (GDPR). It also emphasizes the importance of following the National Institute of Standards and Technology's (NIST-AI) requirements, especially for AI-powered goods.

The purpose of this research is to analyze the transparency and trustworthiness of these companies' data collecting, usage, retention, and protection procedures. This evaluation is critical for understanding how these companies negotiate existing privacy requirements while also preparing for future issues as AI and machine learning technologies improve. The global nature of data necessitates a global view on privacy, taking into account diverse legal and cultural standards.

In addition, the paper covers the significance of Privacy Impact Assessments (PIAs) in detecting and mitigating privacy risks in new initiatives, as well as providing insights into potential privacy issues and compliance gaps within major digital organizations. The research aims to provide best practices for other businesses to improve their privacy strategy and contribute to industry-wide privacy benchmarks.

Overall, the problem statement emphasizes the importance of analyzing these IT companies' privacy practices in order to understand their commitment to data protection, ethical AI use, and preparation for the shifting digital landscape.

Privacy Risks

The General Data Protection Regulation (GDPR) and the National Institute of Standards and Technology (NIST) provide frameworks for addressing personal data privacy and security concerns. GDPR, which is applicable in the European Union, establishes criteria for lawful and transparent personal data processing, highlighting the risk of improper handling. Organizations must put in place strong security measures to protect personal data from breaches while also maintaining its availability, confidentiality, and integrity. GDPR also mandates Data Protection Impact Assessments (DPIA) for high-risk operations and requires specific informed consent for data processing.

NIST, a United States-based agency, focuses on cybersecurity and promotes stringent security safeguards. It advises against lax access controls and encryption, both of which can allow illegal data access. NIST highlights the necessity of incident response plans in addressing security vulnerabilities quickly. Continuous monitoring of security measures is also critical for detecting and resolving security issues as well as lowering the chance of undiscovered privacy breaches. Both the GDPR and the NIST aim to protect personal data through standards, consent, and strong security measures, but their techniques and scopes differ according to their various jurisdictions and focuses.

Result:

PIA Result:

The spreadsheet "Privacy Comparisons" provides an analytical summary of privacy practices in Personal Information Assessment (PIA) across Microsoft Copilot, Google, and IBM. It demonstrates that IBM does not handle personal data, whereas Microsoft Copilot and Google do. While Microsoft Copilot admits to collecting personal information, Google does not, and IBM's position is unclear. Microsoft Copilot and IBM are participating in the types of personal data, but Google's stance is unknown. In terms of third-party involvement, Microsoft Copilot and IBM, unlike Google, operate independently of third parties. All three companies agree on the collection of non-personal data, demonstrating a shared methodology for collecting

non-identifiable information. This comparison demonstrates the major tech giants' various privacy policies and procedures.

GDPR Result:

The GDPR sheet compares how Microsoft Copilot, Google, and IBM adhere to the General Data Protection Regulation (GDPR) rules, with a focus on data privacy procedures. While the sorts of data obtained differ and are not uniformly disclosed for each organization, the sheet exposes diverse data collection methodologies. Microsoft Copilot collects a wide range of information, including personal information, use information, and payment information. Google, on the other hand, is concerned with personal, demographic, and device data. Similarly, IBM collects personal information, service interaction data, and device information. Data collecting methods also differ: Microsoft Copilot captures data during account creation or service use, Google during service registration or interaction, and IBM during service use or event registration. All three organizations collect data from client interactions via surveys, feedback forms, and other means. This report provides insights into their GDPR compliance, emphasizing their distinctive data handling techniques and displaying their dedication to privacy and data protection in a fast-changing digital context.