# WIP: Privacy and Data Awareness Education in the Artificial Intelligence Age

1st Susan S. Conrad
*Marymount University*
Arlington, VA
sconrad@marymount.edu

2nd Diane Murphy
*Marymount University*
Arlington, VA
dmurphy@marymount.edu

***Abstract*—This work-in-progress innovative practice paper describes how integrating data privacy education into a computer science/technology or engineering curricula will prepare students to adopt a privacy by design (PbD) approach when designing, building, and maintaining data driven systems. Artificial Intelligence (AI) is at the forefront of technological advancements and is revolutionizing everyday lives through aspects such as smart healthcare, self-driving cars, voice activation, and automated decision making. With exponential AI market growth, the need to better train AI models to behave in ways that emulate or complement human behavior necessitates large data sets often containing personal health, biometric, location and other personally identifiable information (PII). The data used for training AI applications has great transformative potential to improve our lives, but it offers a significant privacy risk to individuals. Using data without consent is a violation of global data protection laws and could result in financial and/or criminal penalties. The need to educate students about privacy is clear, but the literature is vague in addressing how to integrate privacy education into engineering, technology, or computer science programs. Given the complex and ever-changing privacy requirements, this paper discusses a roadmap for how data privacy can be integrated into AI, computer science and engineering curricula to help students comply with laws and ethical challenges when developing AI applications.**

***Index Terms*—privacy, data privacy, privacy laws, Artificial Intelligence, education, ethics, machine learning**

## I. Introduction

Artificial Intelligence (AI) has become ubiquitous with the term innovation, as technology solutions look to include AI components into every aspect of life from monitoring glucose levels to hiring employees. This explosion of AI has been enabled by advancements in computer chip design and the availability of large datasets. The trifecta of chips, data and human ingenuity are the necessary ingredients for AI to dominate technology-infused innovations. However, as with any transformational change, trade-offs are often necessary to accommodate effective progress. One AI trade-off for progress comes at the expense of data privacy. Large datasets are collected from an array of sources, many without any restrictions or preservation of an individual's anonymity. For example, some datasets may contain personal healthcare information without individuals' consent. Other data may contain data such as name, address, and gender that clearly identifies an individual and potentially exposes them to personal physical or reputational harm. Another concern is anonymized data that when combined with other data may identify the individual [1]. Also, some data may be from children, in clear violation of the Children Online Privacy Protection Act (COPPA). Given the significance of datasets for the training of AI applications, ensuring that the data has not violated the privacy of individuals is regulated by many countries, especially in the European Union (EU).

Aside from personal privacy concerns, source reliability, incorporated intellectual property, and data quality influence the applicability of a model to reliably perform tasks such as pattern recognition, predictive analytics and decision making. Developers and domain experts must evaluate the suitability of a data source by looking for data anonymization, deidentification, omissions and other data issues. Knowing from where the data was collected, as well as the process for collecting and managing the data, is essential for adhering to extensive and varied international legal requirements. AI developers, together with domain experts, must be able to assess the data quality for authenticity and accuracy before training an AI application. Failure to use data that is balanced, accurate, and unbiased, negatively impacts an AI application to effectively perform in accordance with behavioral norms and may be discriminatory [2]. Ensuring that the provenance of the data is known, and that the data is used legally, protects AI developers, domain experts, and users from legal repercussions and helps maintain public trust in AI applications.

Security must also be considered. Technologies such as data lakes and data fabric, with their collection of diverse data from multiple sources in various locations, are becoming commonplace as sources of data for AI model development. Consequently, there is increased possibility of intentional data modification and data poisoning [3].

A clear motivation for AI developers to become educated on data privacy is to be compliant with legal regulations regarding data use. Many AI applications process vast amounts of personal data, and this presents unique challenges and risks for the success of AI innovations and individual privacy. To answer the question why teaching privacy-by-design and data awareness are important to AI developers, a theoretical framework, the Data Privacy Education Framework, has been developed. It consists of four components: privacy laws and compliance, data provenance, ethics and transparency, and Privacy Enhanced Technologies (PETs). This paper will provide a

roadmap of concepts and learning objectives that engineering, technology, and computer science programs should include in their curriculum.

## II. LITERATURE REVIEW

Privacy by design (PbD), originally proposed as a concept for embedding privacy protections into technology products, has evolved into a accountability model for organizations to protect the privacy of an individual's data [4]. The popularity of the concept dates back to the 1990s led by Dr. Ann Cavoukian who identified seven foundational principles for effective PbD [4]. Since then, others have proposed privacy methodologies such as GDPR, NIST and MITRE. Yet these initiatives have not been readily adopted by engineering, technology, and computer scientists. In a study of more than 80 technologists, the engineers interviewed claimed that user privacy requirements were difficult to integrate into design, and many thought it was beyond the responsibilities of their job [5]. Corporate departments with embedded privacy responsibilities tend to take a hands-off approach to consumer privacy requirements, claiming that ambiguous definitions of privacy and privacy laws could not be uniformly implemented into technologies. Privacy began to be a compliance check-the-box approach with little substance and great variability depending upon the technologist [6] [5]. In another study of 168 developers, the use of privacy methodologies occurred only when it was mandated by the organizational leaders and often technologists put forth minimal effort to meet privacy requirements [7].

While there has been research discussing privacy education for engineering and computer science programs, little seems to have been done to integrate privacy methodologies. As far back as 2013, researchers have proposed ways that such programs can add privacy methodologies to existing courses. However, it appears that few academic institutions have added significant privacy components to the system lifecycle design models being taught [8]. The concern for privacy methodologies embedded within AI models is even more pronounced as machine learning (supervised and unsupervised) rely on large datasets to train models without evaluating the data for PII [9]. Contributing to the literature for privacy education, best practices, PbD, and Privacy Enhanced Technologies (PET) are scattered through the literature [10] [11] [12]. However, the literature is scarce when it comes to defining privacy learning objectives and measuring outcomes when teaching PbD methodologies to technology and engineering students. .

## III. PRIVACY BY DESIGN CONCEPTS

As engineering, technology, and computer science curricula adapt to include education on privacy, several topic areas are considered essential to properly prepare students with the necessary knowledge to modify their design lifecycle so that privacy is planned into the design and not an after-thought, especially for those working with AI. Topics include: (A) privacy laws and compliance policies; (B) data provenance and ownership; (C) privacy ethics; (D) privacy enhancing technologies (PETs).

### *A. Privacy Laws and Compliance Policies*

AI engineers and domain experts must be aware of the various privacy laws surrounding the data being used to train their AI models. These laws dictate definitions of private data, user rights and penalties for violating these rights. The European Union (EU) leads the way by defining eight rights to safeguard a user's privacy as defined in the General Data Protection Regulations (GDPR) [13]. These rights include topics such as: consent to collect data, consent for cookies, disclosures for sharing of data, retention of data, user access to this data, the right to delete data, prohibition of automated decision-making, right to appeal and privacy-by-design. Failure to adhere to these laws can have enormous financial consequences. Some examples of fines for failing to follow GDPR are: Meta €1.2 Billion (2023), TikTok €345 (2023), Amazon €746 (2021), and WhatsApp €225 (2021) [14].

Currently the U.S. does not have a uniform definition of PII, nor is there a national privacy law, except the 4th Amendment, which protects the right of people to be "secure in their persons, houses, papers, and effects, against unreasonable searches and seizures" [15]. There are sector specific privacy laws, such as the 1996 Health Information Portability Accountability Act (HIPAA) for health records, the 1970 Gramm-Leach-Bliley (GLB) for financial data, the 1970 Fair Credit Reporting Act (FCRA) for credit information, the 1974 Family Education and Rights Privacy Act (FERPA) for student academic records, and the 1998 Children's Online Privacy and Protection Act (COPPA) for children under 13 [16]. However, these niche laws seek to protect only specific data and not an individual's entire personal identity. Several states have enacted their own privacy laws and have differing requirements. For example, Illinois deals only with biometrics, while Delaware focuses on stricter laws protecting children's data [17]. This patchwork of ever-changing state, national and international laws differ in terms of regulations, penalties and scope making it challenging for organizations to remain privacy compliant, particularly if their business crosses multiple jurisdictions. The fluidity of privacy laws/regulations necessitates that educators prepare students to be legally compliant and observant of the evolving privacy landscape.

More recently, the EU enacted the 2024 AI Act, effective on August 1, 2024, so identifying regulation based on levels of risk, i.e., the probability an individual can experience harm from an AI application using their data, without consent, to train the AI application. The law impacts developers and providers of AI applications and requires that AI application designers implement privacy-by-design practices at each phase of the AI lifecycle, including: (1) proactively include privacy in the design, (2) make privacy the default option, (3) privacy embedded in the design at all development phases, (4) transparency of algorithms, (5) user-centric design, (6) data privacy in the full-functionality, and (7) end-to-end security. The data used to train AI models must comply with the quality

criteria defined by the AI Act, and the application should be transparent to adequately interpret the system's output appropriately.

### *B. Data Provenance and Ownership*

Developers are responsible to know where the data they collect originated and what permissions the user has granted for this data. Data with PII or health information that has not been given with user consent is a violation of GDPR, CCPA, Virginia and other laws and can result in severe penalties. When evaluating data there are four specific areas that developers must understand: data collection; data use and disclosure, data retention; and data deletion.

Data Collection: At the point of collection, users must provide their consent to share their data. During the data collection phase, the data collector must stipulate for what purpose the data is being collected and how the data may be shared. Failure to provide consent means that the user does not want their data to be collected, and thus it is not legally available to be used in a dataset to train machine learning models. Types of data consist of: (1) first party data, (2) surveillance data, (3) repurposed data, and (4) third-party data. Each data type legally requires users to provide consent before data can be used.

Use and Disclosure: A privacy notice, also termed a privacy statement, describes how the collection organization plans to use, disclose, store, and retain the information. Without providing the specifics about what will happen with the data, the users cannot appropriately give consent. AI developers and domain experts need to know that users have consented to the use of the data, along with the limits of how the data can be used, before using the datasets in machine learning model development. Repurposing or disclosing data in ways other than intended by the user may cause privacy harm and may be illegal.

Data Retention: Data should only be retained for the minimum amount of time necessary. The data must be secured and must have the availability to be corrected. Since users can correct their data, AI developers must consider how their AI models will be impacted if after using the data to train AI models, the data is found to be incorrect. If new uses for the data are uncovered, the users must be notified and provide consent or there may be legal ramifications.

Data Deletion: Under many privacy laws users have the right to delete their information. This means data cannot be available for training AI models if the user has requested the data be deleted or the original data privacy notification retention period has expired.

AI developers must take additional precautions to preserve the privacy of individuals' data used to train AI models. Often training datasets include data collected by data brokers and other third-party sources that may lack traceability and anonymity [18]. Without traceable datasets, AI applications may use distorted data sets that unbalance the application's ability to systematically be free of impartiality when making decisions and its outcomes may adversely affect the lives of people. The transparency of data provenance (origination), their pathway from this origin, as well as the authenticity of the datasets are critical to the usefulness and accuracy of any AI application.

One pivotal issue with any dataset is that it may contain personal data obtained without the individual's consent. In addition, not all data must be protected with the same degree of diligence. Categorizing various data types and defining PII is overly complex based upon the context and characteristics of the data and the date and time it was generated. Labeling, a key component of categorization, is fraught with individual interpretation and ambiguity, minimizing impartiality of the definition [19]. When evaluating data for AI, Figure 1, the Privacy Action Model, can assist developers in assessing risks and identifying appropriate actions to mitigate problems.

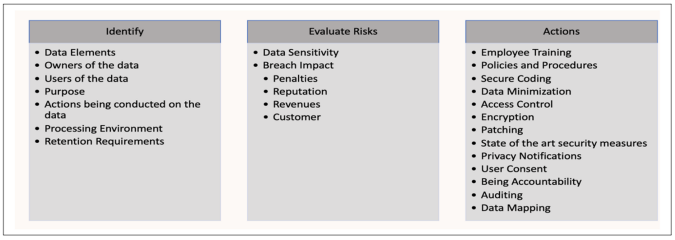


Fig. 1. Privacy Action Model. [12]

Privacy laws and regulations are designed to protect individuals throughout each phase of the data lifecycle. Prior to collecting, data organizations must provide individuals with their privacy notification policy and give users an opportunity to consent to have their data collected and/or shared. AI engineers and domain experts must collaborate with internal stakeholders from every corner of their organization to understand the types of data collected and only use the data as intended by the individual [16]. Each phase of the data lifecycle requires data collectors to evaluate the data's sensitivity, purpose, processing, storage and sharing requirements [10]. The NIST privacy framework specifies several tools to guide AI developers in how to design and use data for AI models [20]. Each phase has unique requirements and will vary depending upon the sensitivity of the data. Given the evolving privacy landscape, strong educational programs are imperative to prepare students for careers in AI.

### *C. Ethics and Transparency*

For an AI application to be considered ethical, it must adhere to a set of principles that ensure its development, deployment, and operation respect human rights and values, promote fairness, and prevent harm. This involves addressing multiple dimensions of how AI applications impact individuals and society as a whole [21]. Bias, transparency, freedom from surveillance, and accountability are critical components to creating an ethical and transparent AI model.

Biased and discriminatory are often ethical and operational criticisms of AI applications. The potential for AI to perpetuate or even exacerbate existing societal biases runs high when data

is unchecked, and its provenance is unknown. AI applications learn from historical data which may contain biases based on race, gender, or socioeconomic status. When such biased data is used to train AI models, the resulting decision, whether related to job hiring, loan approval, or law enforcement, can be unfairly skewed against certain groups, thereby reinforcing inequalities [22] .

Transparency in AI is intricately linked to the issue of bias and is vital for building trust and accountability in AI applications. Many AI models, particularly those based on deep learning, operate as "black boxes" where the decision-making process is opaque. This lack of visibility can make it difficult for users and regulators to understand how decisions are made, posing significant challenges in contexts where fairness and accountability are critical [23]. For instance, in healthcare, an AI application's recommendation for a patient treatment plan needs to be fully explainable to ensure it aligns with medical standards and individual patient needs.

Freedom from surveillance is a Fourth Amendment right. Once that freedom has been violated, a person's sense of safety and mental health can be impacted . AI applications that process vast amounts of personal data can lead to invasive levels of surveillance, if not effectively managed. This raises questions about individual autonomy and the control people have over their own data and life.

Accountability and explainability also should be integrated into AI designs. The EU AI Act requires this functionality be baked into AI applications. Clear mechanisms to hold designers, developers, and deployers of AI applications accountable for how their applications operate are part of this law and provide a regulatory framework that can address any harm caused by AI applications [18].

### *D. Privacy Enhancing Technologies (PETs)*

PETs are integral elements for PbD in which a developer takes a "privacy-first" approach and integrates technical and procedural measures to safeguard individual rights and mandate 'data protection by design and by default'[12]. PETs seek to minimize data collection and only store the least possible PII needed based on a holistic approach to privacy throughout the product lifecycle.

There are three distinct categories of PETs which focus on 'data protection by design and default' utilizing technology tools, compliance regulations and best privacy practices [12] [20]. These include PETs that: (1) reduce the amount of identifiable data being processed about an individual which address the "data minimization principle"; (2) technology practices that hide and shield which address the security principle; and (3) split or control access to personal data which address both the data minimization and security principles [12]. Technical application of PETs includes diverse types of encryptions, differential privacy, pseudonymization, obfuscation, synthetic data, anonymization, and use of federated learning for machine learning.

## IV. TEACHING PRIVACY BY DESIGN

Preparing students on privacy concepts starts with laws, regulations, policies, and procedures. Laws and regulations have a global reach as data may be collected in one location and may reside in a server at another location, while the end-user may be in yet a third location. To be complaint, AI developers and domain experts need data awareness, including the location of data from collection to application, data provenance, and data ownership, when determining how the data may be used and shared. Given the complexities associated with the data ecosystem, an outline of key concepts which should be incorporated into the curriculum are given below.

### *A. Understanding Privacy Laws and Compliance*

The first module discusses relevant laws impacting AI applications and provides an overview of privacy laws and terminology. This module analyzes international privacy laws, focusing on the EU's General Data Protection Regulations (GDPR) as a model. Since the United States does not currently have a national privacy law, the course should focus on state laws such California and US sector laws that are dedicated to protecting privacy such as the Healthcare Information Portability and Accountability Act (HIPAA). In addition, analysis of related legislative initiatives such as the Patriot Act, Freedom Act and Foreign Intelligent Surveillance Act (FISA) stimulate discussions among students as they debate national security concerns with privacy rights.

As educators look to integrate legal and regulatory privacy concepts into engineering, technology, and computer science courses, an example of learning objectives may include: (1) identify PII data and know what protections are required; (2) understand international privacy laws, especially the GDPR; (3) discuss sectoral privacy laws such as HIPPA, COPPA, etc.; (4) analyze US state privacy laws. Learning activities that can be incorporated into existing courses could include case studies of privacy litigation especially those involving Google, Microsoft and Meta. This activity will encourage students to use the cognitive process from Bloom's Taxonomy of "remember," "understand," "analyze," and "evaluate."

### *B. Examining Data and Data Ownership*

Understanding the origins of data and the transfer of ownership from individual to data controller to data processor is an integral step in protecting an individual's data privacy and compliance with the laws. This module will focus on the data lifecycle following data through the phases of collection, usage and disclosure, retention, and deletion. Students will learn about user consent, cookies and other collection and tracking methods to understand how data can be misused without a user's consent.

Learning objectives for this module includes: (1) understanding the lifecycle; (2) examining user notifications and consent documents; (3) applying the Privacy Action Model when developing technology products; (4) connecting legal requirements with data usage practices. Activities to help students meet the learning objectives may include examining

websites and consent agreements to understand what companies want to do with personal data; developing simple technology products to track the data flow, examining risks along the way; and examining large datasets to analyze origin and content of data.

### C. Evaluating Ethics and Transparency

The third module discusses using technology in an ethical and transparent manner. Privacy and AI decisions can greatly impact people's lives. Deep fakes disparaging a person's reputation; applications determining a person's creditworthiness or surveillance cameras intended to keep individuals safe are examples of privacy and AI ethical concerns. Decisions about privacy protection and safety are often made by technologists who lack a familiarity with privacy laws. Large datasets used for training may be biased, contain PII, or be discriminatory. Questions about origination of the data, user consent, and PII cleansing may present an ethical dilemma for AI developers when faced with tight deadlines. Bias, discrimination, and unfairness can be a by-product of non-representative data used in AI applications. A lack of transparency in how decisions are made can result in inaccurate decision-making and even personal harm.

Learning outcomes in this module include students learning how to analyze: (1) biases in data; (2) algorithm transparency; and (3) accountability in data output. Case studies can be used to examine known ethical issues with AI technologies.

### D. Incorporating Privacy Enhancing Technologies (PETs)

This module discusses the concept of privacy-by-design by integrating software and hardware solutions that minimize data collection issues; hide data, provide security, limit access, and minimize risk. Students will pursue the following learning objectives: (1) hiding techniques, such as encryption, to protect privacy but not impact the utility of the data; (2) data splitting techniques that use a systems and data architectures approach to processing, managing, and storing personal data [12]; (3) data generating techniques that weaken the relationships between data by adding or modifying additional data to the data stream, providing differential privacy (noise), and use synthetic data; (4) other topics including cookie collection, web tracking, location tracking and surveillance.

Learning activities to reinforce the learning for these techniques would include hands-on exercises to implement and test the results. Students would practice adding differential privacy, encryption to data, and test the security of the data. Students would also examine methods to distribute data across multiple servers to minimize risk.

## V. Conclusion

With more data collected and stored in a digital format, the vulnerability of data privacy is more ever-growing. Digital devices, aimed at making life easier, are collecting data about every aspect of a user's life without regard to personal privacy or potential harm. Teaching AI developers and programmers about the ethical and legal consequences of collecting, processing, storing, retaining and deleting an individual's data should be required in the engineering design process. Protecting privacy needs to be baked into technology designs to ensure compliance with privacy laws and respect for the individual. Future studies evaluating technology student attitudes and behaviors towards PbD, data privacy, data ethics, data usage, and the integration of PETs to protect privacy will inform the direction of engineering and computer science curriculums for the next generation. This paper provides an evolving roadmap showing how privacy and data awareness may be incorporated into engineering and computer science classes to benefit students.

## References


[1] "How is Big Data Collected by Companies?," Computools.

[2] D. R. Cox, C. Kartsonaki, and R. H. Keogh, "Big data: Some statistical issues," Stat. Probab. Lett., vol. 136, pp. 111–115, May 2018, doi: 10.1016/j.spl.2018.02.015.

[3] J. Snow, "As Generative AI Takes Off, Researchers Warn of Data Poisoning," WSJ.

[4] A. Cavoukian, "Privacy by Design The 7 Foundational Principles".

[5] A. E. Waldman, "Designing Without Privacy," Mar. 31, 2017, Rochester, NY: 2944185.

[6] K. A. Bamberger and D. K. Mulligan, "PRIVACY ON THE BOOKS AND ON THE GROUND," Stanford Law Rev., vol. 63.

[7] A. Senarath, M. Grobler, and N. A. G. Arachchilage, "Will They Use It or Not? Investigating Software Developers' Intention to Follow Privacy Engineering Methodologies," ACM Trans Priv Secur, vol. 22, no. 4, p. 23:1-23:30, Nov. 2019, doi: 10.1145/3364224.

[8] J. Vaidya, B. Shafiq, D. Lorenzi, and N. Badar, "Incorporating Privacy into the Undergraduate Curriculum," in Proceedings of the 2013 on InfoSecCD '13: Information Security Curriculum Development Conference, Kennesaw GA USA: ACM, Oct. 2013, pp. 1–7. doi: 10.1145/2528908.2528918.

[9] B. Liu, M. Ding, S. Shaham, W. Rahayu, F. Farokhi, and Z. Lin, "When Machine Learning Meets Privacy: A Survey and Outlook," ACM Comput. Surv., vol. 54, no. 2, pp. 1–36, Mar. 2022, doi: 10.1145/3436755.

[10] [S. S. Conrad, "Integrating Data Privacy Principles into Product Design: Teaching 'Privacy by Design' to Application Developers and Data Scientists," J Comput Sci Coll, vol. 38, no. 3, pp. 132–142, Nov. 2022.

[11] M. Alshammari, "A Principled Approach for Engineering Privacy by Design," Ph.D., University of Oxford (United Kingdom), England, 2019.

[12] ICO.UK, "Data protection by design and default," Guide to Accountability.

[13] GDPR, "What are the GDPR consent requirements?," GDPR.eu.

[14] DPM, "Luxembourg DPA issues €746 Million GDPR Fine to Amazon," Data Privacy Manager.

[15] D. Solove and P. M. Schwartz, Privacy Law Fundamentals 6th Edition. International Association of Privacy Professionals, 2022.

[16] P. Swire and D. Kennedy-Mayo, U.S. Private Sector Privacy, 3rd ed. IAPP Publisher, 2020.

[17] Folks, Andrew, "US State Privacy Legislation Tracker."

[18] C. Novelli, M. Taddeo, and L. Floridi, "Accountability in artificial intelligence: what it is and how it works," AI Soc., Feb. 2023, doi: 10.1007/s00146-023-01635-y.

[19] L. Ticong, "What Is Data Classification? Your Ultimate Guide," Datamation.

[20] Nist.gov, "Privacy Framework," NIST.

[21] N. Talagala, "AI Ethics: What It Is And Why It Matters," Forbes.

[22] Abrams, Zara, "Addressing equity and ethics in artificial intelligence," https://www.apa.org.

[23] N. Green, "An AI Ethics Course Highlighting Explicit Ethical Agents," in Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, in AIES '21. New York, NY, USA: Association for Computing Machinery, Jul. 2021, pp. 519–524. doi: 10.1145/3461702.3462552.