

RESTRUCTURING LEGAL PRINCIPLES OF PERSONAL DATA PROTECTION TO RESPOND TO BIG DATA CHARACTERISTICS

**Didit Darmawan, Arif Rachman Putra, Ainor Rasid, Hari Santoso, Feby Aji Laksono,
Mohamad Izzul Fahmi, Yudo Baskoro**

Correspondence: dr.diditdarmawan@gmail.com

Universitas Sunan Giri Surabaya

ABSTRACT

This doctrinal legal research examines the legal protection of data privacy within the big data ecosystem. The analysis reveals that conventional legal constructs, which rely on individual consent and purpose limitation, face fundamental obsolescence when confronted with the massive volume, high velocity, and wide variety of modern data processing. These characteristics render notice-and-consent mechanisms hollow and fragment accountability across multiple actors in the data chain. This study proposes a reconstructed normative framework that reinterpretes established Fair Information Practice Principles and introduces five supplementary principles proportionality, precaution, algorithmic fairness, collective data governance, and enhanced portability. The proposed framework shifts legal protection from a static ex-ante control model toward a dynamic, risk-based oversight regime that monitors continuous data flows and usage. Effective legal protection does not obstruct innovation; rather, it ensures that data-driven development proceeds within boundaries that respect human dignity and individual autonomy. This research contributes to legal scholarship by bridging the gap between individual privacy rights and systemic data processing realities. For policymakers, the findings offer clear guidance for drafting legislation that incorporates accountability mechanisms and technical audit provisions. The study concludes that a legal transformation is imperative, requiring not mere amendment but a fundamental reorientation of how law perceives and regulates personal data. Regulators must develop technical expertise to audit algorithms, while businesses must embed privacy principles into system architecture from the outset. This research thus provides a comprehensive roadmap for building a responsive and just data protection regime.

Keywords: Data privacy, big data analytics, legal framework, normative reconstruction, algorithmic accountability, consent mechanics, data governance.

INTRODUCTION

The development of information technology over the past two decades has generated a flow of digital data whose volume exceeds traditional processing capabilities. Digital transformation has brought major changes to the global economy, including through the

growth of electronic commerce that generates large volumes of data flows (Hartana et al., 2024). Every human activity ranging from banking transactions, information searches, and social media interactions to physical movements tracked via mobile devices leaves a digital footprint recorded and accumulated in giant databases. Big data is not merely a technical term, but a new social reality that alters how public and private institutions make decisions, design policies, and offer services. The development of information technology, particularly big data and data analytics, helps managers make more objective decisions through processing large amounts of information to support business strategies (Ali & Darmawan, 2023). Analytical capabilities applied to large datasets enable the identification of behavioral patterns, trend predictions, and service personalization with a level of precision previously unattainable (Abdeldayem, 2023). However, behind these benefits, the massive and continuous collection of data raises fundamental questions regarding the boundaries of data collectors' authority over individuals' personal information. Individuals often lose control over their own data because collection processes occur covertly and without explicitly understood consent. Information asymmetry between data owners and data collectors has become a defining feature of the contemporary digital landscape.

Privacy as a fundamental right comes under severe pressure from the logic of big data, which prioritizes value extraction from every piece of information. The right to self-determine when, how, and for what purpose personal data is used is undergoing erosion because consent mechanisms designed in the pre-digital era are no longer adequate (Ryan, 2019). Consent granted by checking a long checkbox does not reflect a conscious choice, but rather compliance with unread terms. In the big data ecosystem, aggregating diverse data sources generates new information never previously agreed upon by the individual, rendering the concept of specific purpose-based consent obsolete. Personal data that should be protected can be misused, thereby harming individuals and diminishing public trust in digital services (Aziz et al., 2023). Legal protection relying on notification and consent faces a crisis of relevance, as every data transaction involves multiple parties and shifting purposes aligned with analytical advances. This loophole is exploited by business operators to claim legitimacy for data processing under the guise of legitimate interests, while individuals lack an equal bargaining position. Consequently, privacy violations frequently occur not due to an absence of rules, but because existing rules were not designed to address the complexity of large-scale data processing.

Big data and analytics offer economic benefits while simultaneously presenting legal risks, necessitating a balance between business opportunities and legal protections (Bashori et al., 2024). Various jurisdictions have begun responding by establishing specific legal frameworks for personal data protection; however, the approaches taken vary widely and are often reactive to breach incidents (Liu, 2024). The European Union, through the General Data Protection Regulation, sets a high benchmark by granting strong rights to data subjects, including the right to be forgotten and the right to data portability. The United States favors a sectoral approach that differentiates protection based on data types and

industry sectors, resulting in no comprehensive federal personal data protection law. In Asia, several countries such as Japan and South Korea have revised their data protection laws to accommodate big data challenges, while others remain in policy formulation stages. This diversity reflects differences in cultural values, economic interests, and legal traditions across nations. However, these disparities also create issues when data flows across borders, as lower protection standards in one jurisdiction can become loopholes for data exploitation from another. As Solove (2008) noted, legal approaches to privacy often get trapped in a dichotomy between absolute protection and complete freedom, whereas social reality requires a dynamic balance that considers the interests of diverse stakeholders.

Big data not only expands data volume, but also transforms the nature of data itself from static to dynamic and from structured to unstructured. The characteristics of volume, velocity, variety, and veracity demand legal approaches capable of keeping pace with technological changes without sacrificing legal certainty (Fleckenstein & Fellows, 2018). Data from communications, transactions, and customer interactions can help companies identify specific patterns and tailor services according to customer needs and expectations (Essa et al., 2024). Overly rigid regulation will stifle innovation, whereas overly lax regulation jeopardizes individual rights. This issue is further complicated by the use of artificial intelligence and machine learning algorithms that process data in ways not entirely predictable by their designers. Automated decisions based on personal data can affect an individual's access to credit, employment, insurance, and public services without clear appeal mechanisms. In such situations, legal protection of data privacy must transcend mere rules on collection and storage to encompass data processing, sharing, and erasure. Tene and Polonetsky (2012) reminded that traditional approaches focused on restricting data collection become less effective, as the value of big data lies precisely in the ability to combine and analyze data from multiple sources. Therefore, legal frameworks need to shift from control at the point of entry toward continuous oversight of data use and data flows.

The ever-changing development of the digital economy creates new opportunities while raising questions regarding economic interests, data protection, and legal certainty for the public (Aryanto et al., 2024). In Indonesia, personal data protection gained public attention following several consumer data breach cases involving financial institutions and digital platforms. The Electronic Information and Transactions Law and its derivative regulations govern personal data aspects in a limited capacity, yet fail to provide clear definitions and effective enforcement mechanisms (Lestari & Rasji, 2024). This ambiguity creates a gray area exploited by data collectors to evade accountability. The draft personal data protection bill discussed in parliament indicates an awareness of the importance of specific regulation, but the lengthy legislative process signifies the complexity of the interests involved. In a society where mastery of information technology remains unequal, legal protection efforts frequently clash with the imperatives of digital economic development and ease of doing business. The fundamental question that arises is how the law can acknowledge the reality of big data without losing its normative function—

namely providing certainty, justice, and benefits for all citizens. This study seeks to fill the gap in legal studies by specifically examining data privacy protection from structural and normative perspectives, rather than solely from technological or economic viewpoints. Thus, an in-depth analysis of existing legal frameworks and the principles that should serve as the foundation for data protection becomes an indispensable first step.

The main problem faced in the legal protection of data privacy in the big data era stems from the inability of conventional legal instruments to reach modern data processing practices. Existing statutory regulations were largely created prior to the explosion of digital data, so concepts such as personal data, consent, and processing purpose were formulated based on the assumption that data is collected in a limited manner and for specific purposes. Big data shatters these assumptions because collected data can be repurposed repeatedly for various ends unrelated to the initial objective, often without the data subject's knowledge (Talapina, 2023). Uncertainty regarding who is responsible for data protection increases as the data processing chain involves multiple entities—from collectors, storers, and processors to data analysts. Each party can claim to act merely as an intermediary or service provider, leaving legal responsibility fragmented. On the other hand, individuals as data owners lack adequate access to information about how their data is processed, preventing them from detecting violations or asserting their rights. This structural inequality places data subjects in an extremely weak position, while business operators reap benefits from information asymmetry.

Another equally critical problem is the absence of effective law enforcement mechanisms for large-scale data privacy violations. Data breaches represent one of the primary risks accompanying technological advancements and large-scale data utilization (Mujisulistyo et al., 2024). Data leaks involving millions of records often end merely with administrative warnings or fines disproportionate to the harm caused. The burden of proof in data breach cases is extraordinarily difficult because violations occur in digital spaces where evidence is technical and easily altered. General courts face obstacles in understanding digital evidence, while specialized oversight bodies frequently lack resources to conduct thorough investigations (Spanca & Salihu, 2024). Consequently, the deterrent effect of legal sanctions remains very low, and personal data protection relies more on the good faith of business operators than on the threat of sanctions. The low effectiveness of law enforcement fosters the perception that data privacy is an unenforceable right, causing individuals to lose trust in the legal system. In the long run, this loss of trust can impede public participation in the digital economy, as people become reluctant to share data even for beneficial services. Issues of legal protection cannot be separated from those of law enforcement institutions and evidentiary procedures that are adaptive to the characteristics of big data.

Personal data has become a strategic commodity whose value equals that of natural resources, and in some cases is even more valuable because it is renewable and generates continuous insights. When personal data is exploited without protection, individuals lose autonomy over themselves, as decisions concerning their lives are dictated by digital

profiles unfamiliar to them. This loss of autonomy undermines human dignity and the principle of equality, because digital profiles often contain biases that reinforce discrimination against vulnerable groups. A society lacking strong data protection will face the threat of mass surveillance that can silence freedom of expression and political participation. Experiences in various countries show that personal data collected for commercial purposes can easily be accessed by security apparatuses, creating a surveillance ecosystem where it is difficult to distinguish between state and corporate action. This condition threatens the pillars of democracy and human rights, as privacy is not a privilege, but a prerequisite for true freedom.

Across various countries, legal systems are continually developed to balance open data utilization with public privacy protection (Karwati et al., 2024). The drive to construct a legal framework responsive to big data is also motivated by the need to maintain national digital economic competitiveness amidst global competition. Countries capable of providing credible data protection will attract foreign investment, as multinational corporations require legal certainty to process cross-border data (Dai, 2021). Conversely, countries with low protection standards will be avoided by compliance-oriented international business operators, while simultaneously becoming havens for irresponsible actors. A poor reputation in data protection can trigger sanctions from trading partners and erode consumer trust. Conversely, over-regulation without regard for industry needs can stifle innovation and create entry barriers for small and medium enterprises. Therefore, data protection policy formulation must be grounded in a balance obtained through a comprehensive examination of the legal nature of data, individual rights, and collective interests. Without such analysis, policy steps risk being misdirected and producing unintended consequences. This study aims to provide a conceptual foundation for such policy formulation through normative and comparative analyses of various legal approaches developed globally.

The research questions forming the focus of this study are how the characteristics of big data affect the current construction of personal data protection law, and what legal principles can serve as the foundation for establishing a personal data protection framework suited to the reality of big data. This study aims to produce a comprehensive legal analysis of the gaps between existing personal data protection regulations and the demands of the big data era, as well as to formulate normative principles to serve as guidelines for lawmakers and policy stakeholders.

METHODS

This normative-qualitative legal research relies entirely on a literature study without involving field data collection such as interviews, observations, or surveys. Grounded in Chynoweth's methodological framework, a doctrinal approach is applied to thoroughly examine concepts, principles, and written norms sourced from both library collections

and electronic databases. The legal materials analyzed comprise three tiers: primary legal materials consisting of statutory regulations related to data protection; secondary materials including textbooks and scientific journal articles on privacy and big data; and tertiary materials such as legal dictionaries. The search was conducted systematically across institutional repositories and accredited journals to explore the meaning of legal norms amidst technological escalation.

The data processing and analysis procedure applies the SALSA (Search, Appraisal, Synthesis, Analysis) framework by Grant and Booth (2009). The process begins with a thematic keyword-based literature search, followed by an evaluation of source credibility and relevance. Validated findings are then synthesized into groups of legal arguments aligned with the research questions. Subsequently, a comprehensive analysis is conducted to yield a coherent normative interpretation. The validity and reliability of this study are maintained through logic-consistency testing of arguments, the use of vetted references, and transparent, accountable documentation of the search process.

RESULTS AND DISCUSSION

The Influence of Big Data Characteristics on the Existing Legal Construction of Personal Data Protection

The existing legal construction of personal data protection across various jurisdictions is built upon foundational assumptions that data collection occurs in a limited manner, for explicit purposes, and with informed consent provided by individuals. This assumption originates from the era preceding the digital data explosion, when data storage and processing capacities were expensive and constrained. The law presupposed that data subjects knew reasonably well what data was collected, by whom, and for what purpose, making consent a central instrument to legitimize data processing (Giannopoulou, 2020). However, big data shatters these assumptions. The volume of data collected daily reaches scales never previously imagined, making it impossible for individuals to fully comprehend the types of data footprints they leave behind and how that data is processed. Data collection occurs continuously across every digital interaction, often unnoticed, and data once deemed irrelevant can become highly valuable upon aggregation with other datasets. Consequently, the consent mechanism serving as the backbone of legal protection loses its substance, transforming from a conscious choice into a mere formality masking individual powerlessness.

The characteristic of immense volume in big data renders every attempt to seek separate consent for each data use impractical and disruptive to user experience. Regulations mandating notice and explicit consent for every processing purpose create information noise that makes individuals increasingly indifferent to what they agree to. Practice shows that the length and complexity of privacy policies lead most users to check consent boxes without reading the contents, fulfilling formal requirements without reflecting actual

intent. This condition is termed illusory consent, wherein consent is legally deemed valid but sociologically meaningless. Laws anchored in individual consent are unable to address the reality that in a big data ecosystem, every individual action generates data subsequently processed multiple times in unpredictable ways. When consent is granted for purpose A, but the data is later utilized for entirely different purposes B, C, and D, the legal protection afforded by initial consent becomes irrelevant. Thus, volume illustrates the law's inability to accommodate the reality of continuous data flows and shifting processing purposes over time.

The velocity of data processing in big data creates friction with legal procedures constructed on a different temporal logic. Data protection regulations generally require notification to data subjects prior to or at the time of collection, along with access and opt-out opportunities before data utilization. These procedures were designed for situations where data collection occurs at a single point in time and processing unfolds over a set period. In big data, processing occurs in real-time, with algorithms constantly updating analyses as new data arrives (Talapina, 2023). This velocity makes pre-processing notification impossible, as data is processed immediately upon receipt. Consequently, the data subject's right to know and control processing lags behind a technological rhythm far faster than the legal rhythm. Appeal or objection procedures taking days or weeks become useless when data-driven decisions occur in fractions of a second. Velocity also complicates data erasure, as data already processed into profiles or predictions cannot easily be removed without compromising established analytical outputs. This temporal misalignment places the law in a reactive, perpetually delayed position, rendering protection superficial and applicable only to cases detected after harm has occurred.

Variety, encompassing structured, semi-structured, and unstructured data, complicates the legal definition of what constitutes personal data. Traditional data protection law defines personal data as information directly or indirectly identifiable to an individual, such as name, address, identification number, or photograph. This definition relies on the ability to link data to a specific individual identity. In big data, data variety enables identification through combinations of data that appear non-personal at first glance, such as movement locations, shopping habits, browsing patterns, or keystroke dynamics. Data not directly containing identity become hyper-specific when combined, identifying individuals without needing names or addresses. Laws protecting only data that directly identifies a person prove inadequate, as the majority of big data consists of anonymous or pseudonymous data. This loophole is exploited to claim such data falls outside legal protection, allowing unconstrained processing. Conversely, identity-masking techniques viewed as solutions often prove fragile, as sufficient data volume enables re-identification. Data variety also complicates the application of data minimization principles, as big data practices encourage maximal collection for future analytics, directly contradicting the principle that only necessary data should be gathered.

Data quality, or veracity in big data terminology, raises questions regarding responsibility

for the accuracy of processed data. High-volume data frequently contains errors, duplications, incompleteness, or biases from diverse sources. Data protection law grants data subjects rights to rectify inaccurate data; however, exercising this right becomes difficult when inaccuracy stems not from a single datum, but from statistical patterns or algorithmic predictions (Keskar et al., 2021). For instance, an individual may be classified as a high credit risk based on correlations that do not reflect their personal characteristics, yet rectifying their personal data does not alter the classification because it results from aggregate analysis. In such scenarios, data subjects lack avenues to demand correction because the error lies not in specific data, but in analytical methods and predictive models. Furthermore, data entering big data systems often originates from unverified sources, blurring accuracy obligations among collectors, processors, and end users. Regulations governing only data accuracy obligations fail to address issues originating from analytical processes. This condition leaves data subjects increasingly vulnerable, as profiles formed from inaccurate data impact their lives without offering effective channels for objection. Systemic inaccuracy undermines procedural fairness in data protection, depriving subjects of adequate means to defend the truth about themselves.

The economic value of big data incentivizes repeated data utilization for purposes unanticipated at collection. The purpose limitation principle adopted across nearly all data protection regimes mandates that data be used solely for notified and consented purposes. Yet the logic of big data is inverted: data is gathered in hopes of discovering new value later. Businesses resist restricting data use to avoid losing innovation opportunities that might emerge from the same data. The conflict between economic value and legal certainty becomes acute in practice, as attempts to tighten purpose restrictions are viewed as stifling analytics. Existing laws offer loopholes through legitimate interest grounds for controllers, yet this concept is so broad that it is easily misinterpreted. Lacking a clear definition of legitimate interest, any commercial goal can be claimed as such. Consequently, purpose limitation becomes meaningless as daily big data practices violate the principle without consequence. Effective legal protection requires mechanisms balancing data value with individual rights, yet current legal constructions lack adequate tools for such balancing, having been designed for a world where data value was limited and predictable.

Aggregating data from multiple sources within a big data ecosystem generates new information that cannot be traced back to original sources, creating issues surrounding ownership and responsibility for derivative information. When data from banking, e-commerce, social media, and location sensors combine, the result is a detailed individual profile never explicitly approved by the person. Law does not clearly regulate whether derivative information constitutes protected personal data and who holds responsibility for its accuracy and use (Talapina, 2023). Businesses often claim derivative information as intellectual property resulting from analytical investment, refusing to grant data subjects access to it. This claim contradicts the principle that personal data remains the

right of the subject even after processing, yet insufficient jurisprudence exists to affirm that stance. Data aggregation also creates cross-sectoral issues, where data gathered by one entity is used by another with no direct relationship to the individual, severing the chain of accountability. Individuals do not know who to hold accountable if errors exist in derivative profiles, as each entity shifts responsibility to others. This ambiguity is a direct consequence of an open ecosystem structure, differing from the closed processing systems assumed by traditional data protection law.

The involvement of multiple parties across the big data processing chain complicates the controller and processor framework that serves as the pillar of responsibility in data protection regulation. In existing law, the controller determines processing purposes and means, while the processor processes data on the controller's behalf. This division is distinct in traditional organizational structures; however, in big data, roles overlap as data is processed collaboratively by entities that can each influence processing purposes and means. Digital platforms providing services to users while selling data to advertisers and partners position themselves as controllers while acting as processors for third parties. When violations occur, data subjects cannot determine which party is actually responsible, as each points to the other as the primary controller. Laws governing joint and several liability remain rarely applied and are often constrained by jurisdiction. Processing chain complexity is further compounded by cloud computing, where data is stored and processed across data centers in different countries, shifting liabilities according to storage location. This condition creates *de facto* impunity, as no single entity feels fully accountable, leaving data subjects without clear channels to lodge objections or claims. Existing legal constructions were not designed to cover long, complex data value chains, rendering legal liability fragmented and ineffective.

Automated decision-making leveraging big data and artificial intelligence presents unresolved transparency and accountability issues within the legal framework of data protection. Corporate obligations to maintain data transparency and security, as well as satisfy rights of access, correction, and deletion, place companies as parties responsible for managing data within digital systems (Osean et al., 2024). Automated decisions affecting access to employment, credit, housing, or public services are often made by complex algorithms that cannot be simply explained to data subjects. The right to receive an explanation regarding automated decisions mentioned in several regulations proves difficult to fulfill, as machine learning models operate in non-linear ways (Mougdir, 2020). Data controllers can explain variables used, but cannot explain why a specific variable was weighted in a particular case. In practice, explanations provided are often generic, failing to help data subjects understand decisions made about them. Consequently, the right to object becomes an empty right, as subjects cannot formulate convincing arguments without understanding decision baselines. Algorithms may also contain biases unknown to designers, resulting in discrimination against specific groups, yet proving such discrimination is difficult due to non-transparent processes. Traditional data protection law

lacks instruments to audit algorithms or mandate impact assessments prior to deploying automated decisions. The absence of these tools leaves a major gap in legal protection, as decisions impacting individual lives most severely remain outside effective legal oversight.

Data subject rights guaranteed by data protection regulations such as rights of access, correction, and erasure face real technical hurdles when applied in big data environments. Personal data in big data is not stored in a single centralized database, but is dispersed across multiple systems, often in unstructured formats. To fulfill access requests, controllers must execute cross-system searches that are costly and time-consuming. Many controllers intentionally make access procedures complex and convoluted to reduce submitted requests, as each request requires non-trivial technical resources. Erasure rights, known as the right to be forgotten in some jurisdictions, become nearly impossible to execute because data absorbed into analytical systems cannot be deleted without damaging constructed models. Even if data is erased from primary databases, copies may remain in backups, system logs, or machine learning models that have absorbed patterns from the data. Legal protection providing rights normatively without clear technical obligations for fulfillment generates disparities between recognized rights and the ability to exercise them. Data subjects possess formal rights, but cannot exercise them effectively in practice. This condition reduces the credibility of legal protection and reinforces perceptions that data protection law is more symbolic than substantive.

Law enforcement regarding data privacy violations in the big data ecosystem faces serious difficulties, as violations frequently cross international borders and impact mass victims. Data breaches affecting tens of millions of individuals cannot be handled through standard civil litigation due to immense costs and complexity. Class action or representative lawsuits in some countries serve as available instruments, but procedures are drawn out and outcomes uncertain, particularly when violations are committed by multinational corporations possessing vast legal resources (Aura et al., 2024). Courts also face difficulties handling digital evidence, as proof requires scarce, expensive experts, and digital evidence integrity is frequently questioned due to ease of alteration. Data protection supervisory authorities, though granted investigative powers, frequently lack staff and budgets to process growing complaint volumes. Fines imposed in many cases are disproportionate to profits businesses gain from unlawful processing, causing fines to be treated merely as operational costs. The lack of a deterrent effect leads to repeated violations, while legal protection remains at the level of rhetoric. Thus, big data characteristics not only impact the substance of rights, but also undermine the effectiveness of enforcement mechanisms meant to serve as the ultimate guarantee of legal protection.

The characteristics of big data including volume, velocity, variety, veracity, and value are fundamentally incompatible with the assumptions underlying current legal constructions of personal data protection. Law was built upon a model of data processing that is limited, centralized, static, and predictable, whereas big data operates on a model that is massive, dispersed, dynamic, and open to exploration. This mismatch causes data protection

norms ranging from definitions of personal data, consent, purpose limitation, and data subject rights to enforcement mechanisms to become ineffective in delivering meaningful protection. Businesses can easily exploit loopholes arising from this incompatibility to evade responsibility, while data subjects lose the protection they ought to receive. Legal protection of data privacy in the big data era requires a reconstruction of the normative framework that does not merely adjust terminology, but alters the philosophical and operational foundations of how the law views personal data. Without such reconstruction, any remedial effort will produce only patches incapable of sealing structural leaks in the existing protection system. This analysis serves as a departure point for identifying new legal principles that can form the foundation for a data protection framework suited to the reality of big data, as will be elaborated in the subsequent section.

Legal Principles for a Personal Data Protection Framework Adaptive to Big Data Realities

The search for legal principles that can serve as the foundation for a personal data protection framework in the big data era must begin with the acknowledgment that hitherto used approaches are no longer adequate. The Fair Information Practice Principles developed in 1973 by the U.S. Department of Health, Education, and Welfare served as the foundation for data protection regulations worldwide. These principles encompassed prohibitions on secret data-keeping systems, individual rights to know collected data and its uses, rights to prevent data use beyond consented purposes, rights to correct data, and organizational duties to assure data reliability and prevent misuse. This framework was subsequently adopted and expanded by the Organisation for Economic Co-operation and Development through its Guidelines on the Protection of Privacy and Transborder Flows of Personal Data in 1980, revised in 2013. The OECD Guidelines contain eight principles: collection limitation, data quality, purpose specification, use limitation, security safeguards, openness, individual participation, and accountability. These principles have served as global benchmarks for establishing data protection legislation across various countries. However, as detailed in the preceding section, big data characteristics render these principles difficult to apply effectively. Consequently, a reordering of these principles is required to maintain relevance without sacrificing core values intended for protection. This reordering does not imply discarding existing principles altogether, but rather reinterpreting and expanding their scope to align with contemporary data processing realities.

The first principle requiring reinterpretation is the collection limitation principle, which in traditional formulations requires data collection to be lawful, fair, and conducted with the knowledge or consent of the data subject. In big data, data collection occurs continuously and often imperceptibly, rendering knowledge and consent requirements difficult to fulfill. Reinterpreting this principle requires shifting focus from collection itself toward continuous notification and control mechanisms. Rather than mandating advance consent for every collected data type, an adaptive legal framework can require periodic

notifications regarding collected data types and grant data subjects opportunities to opt out of specific future data uses. This approach, known as continuous opt-out, affords individuals control without sacrificing the seamless data collection defining big data. The collection limitation principle must also expand to encompass obligations for data collectors to periodically review and erase no longer necessary data, preserving data minimization principles even at scale. This reinterpretation recognizes massive data collection as an inevitability in digital ecosystems while upholding accountability for collected materials and management practices.

The purpose specification and use limitation principles represent the two principles most impacted by big data logic, as big data relies precisely on data usage for goals unspecified in advance. The purpose specification principle requires collection purposes to be stated at collection and subsequent uses limited to those purposes or non-contradictory alternatives (Forgó et al., 2017). In big data, processing purposes frequently shift as new patterns and insights emerge from identical datasets. Reinterpreting this principle involves distinguishing between primary and secondary purposes. Primary purposes form the core basis for collection and require explicit consent from data subjects, whereas secondary purposes emerge later, requiring notification and opt-out opportunities. Legal frameworks can also mandate that data controllers perform impact assessments before using data for new purposes significantly differing from initial objectives. Impact assessments function as internal oversight mechanisms ensuring new data uses generate no disproportionate risks to individual privacy rights. Under this approach, purpose limitation no longer acts as an innovation barrier while remaining an effective control instrument against data misuse. This approach aligns with ideas that data protection must be dynamic and follow data flows rather than regulating solely at entry points.

The data quality principle, requiring data to be accurate, complete, and updated as necessary, faces massive challenges in big data because volume and velocity render individual verification impractical. Reinterpreting this principle requires distinguishing between data used for individual decision-making and data used for aggregate analysis. For data used in individual decisions bearing significant impacts—such as credit determinations, employment, or service access—accuracy duties must be strictly enforced alongside easily accessible correction mechanisms. For data used in aggregate analysis, accuracy duties can translate into transparency obligations regarding data sources, processing methods, and analytical uncertainty levels. Data subjects hold rights to know that decisions affecting them rely on data of specified accuracy levels and possess opportunities to challenge those decisions. This approach recognizes that not all big data can be individually verified while maintaining accountability for inaccurate data usage. The data quality principle must also expand to cover obligations identifying and mitigating data and algorithmic biases, as bias constitutes systemic inaccuracy that frequently goes undetected. Thus, data quality is no longer understood purely as factual truth, but as procedural fairness in data processing.

The openness principle, requiring transparent general policies regarding developments, practices, and policies related to personal data, must be reinforced in the big data era because processing complexity hinders data subjects from understanding what occurs with their data. Openness cannot consist merely of long, complex privacy policies, as such policies reduce user understanding. An adaptive legal framework should mandate concise, clear, and easily understandable information for lay users, presented in formats enabling cross-service comparisons (Bruening & Culnan, 2016). Openness must also cover access to information regarding algorithms used in automated decision-making, at least regarding variables used and relative weightings. Although full algorithmic transparency may not always be feasible due to intellectual property constraints, openness regarding foundational decision logic remains a prerequisite for accountability. The openness principle must also encompass obligations to notify data subjects of personal data breaches within reasonable timeframes, enabling self-protection measures. Under this approach, openness transforms from a passive, reactive requirement into an active instrument empowering data subjects to protect their rights.

The individual participation principle, granting individuals rights to access, correct, and challenge data relating to them, faces serious technical hurdles in big data as data is dispersed across diverse systems and formats. Reinterpreting this principle requires considering standard interfaces and machine-readable data formats to facilitate data access and portability. Data subjects hold rights to receive their data in structured formats transferable to alternative service providers, preventing vendor lock-in within a single digital ecosystem. Participation rights should also expand to include inputs regarding specific data uses, including rights to withdraw previously granted consent. In big data contexts, consent withdrawal does not imply erasing all data as data may be integrated into analytical models—but at least mandates halting data usage for new purposes and removing data from active systems. Individual participation mechanisms must also encompass easily accessible, responsive grievance channels, sparing data subjects lengthy court processes merely to correct erroneous data. Under this approach, individual participation becomes an actionable, effective right rather than a formal paper promise.

The accountability principle, obligating data controllers to comply with data protection principles, becomes increasingly paramount in the big data era because processing complexity makes external oversight difficult. Accountability within an adaptive framework must translate into comprehensive documentation obligations covering entire processing lifecycles from collection, storage, processing, and sharing to erasure (Vedder & Naudts, 2017). This documentation must be available for audit by independent supervisory authorities and, within defined limits, interested data subjects. Accountability also encompasses mandatory periodic data protection impact assessments, particularly before deploying new processing systems or utilizing data for novel purposes. These impact assessments must identify privacy risks, implemented mitigation measures, and determinations regarding risk acceptability. Data controllers

must also appoint data protection officers responsible for ensuring compliance and serving as contact points for data subjects and supervisory bodies. Accountability is no longer merely ex post but also ex ante, as controllers must demonstrate compliance before breaches occur rather than after detection. This approach shifts evidentiary burdens from data subjects and oversight authorities to data controllers, who possess superior knowledge of their own systems.

Beyond reinterpreting existing principles, a big data-adaptive protection framework requires new principles absent from traditional formulations. The first necessary addition is the proportionality principle, requiring data controllers to balance processing interests against data subject privacy rights. Proportionality mandates that data processing align with legitimate objectives, where benefits derived do not fall short of risks posed to individual privacy. This principle provides room for case-by-case evaluations, preventing controllers from claiming any processing is lawful merely due to commercial interests. Proportionality also dictates that collected data be necessary, prohibiting controllers from gathering data beyond stated purpose requirements. In practice, proportionality can be implemented through mandatory assessments prior to initiating new processing operations and periodic reviews thereafter. This principle provides a normative basis for overseeing excessive data collection and unnecessary usage practices.

The second required new principle is the precautionary principle, obligating data controllers to anticipate potential processing risks and implement preventive measures before harm materializes. Big data characteristics often render risks invisible until breaches occur, making reactive approaches inadequate. The precautionary principle requires controllers to systematically identify risks, including data leaks, data misuse, algorithmic discrimination, and reputational damage to data subjects (Batey, 2022). Required preventive steps encompass data encryption, access controls, employee training, and internal processing policy development. The precautionary principle also mandates emergency response plans for breach incidents, minimizing harm and ensuring rapid notification to data subjects. This approach shifts data protection paradigms from reactive to proactive, aligning with accelerating, complex technological advances.

The third new principle is algorithmic fairness, which recognizes that automated decisions based on big data can produce discrimination and injustice undetected by traditional data protection mechanisms. Algorithmic fairness obligates data controllers to test algorithms for bias and discrimination, taking concrete steps to mitigate identified biases. Testing must occur prior to deployment and periodically thereafter, as bias can emerge from new data or shifting user behaviors. Test results must be documented and available for audit by supervisory bodies. Algorithmic fairness also encompasses obligations to provide meaningful explanations to data subjects regarding automated decisions affecting them, including variables used and decision baselines. Explanations need not reveal complete algorithmic code, but must suffice to enable data subjects to understand and challenge decisions. This principle addresses gaps in traditional frameworks that focus solely on data

itself, without regard to how data is processed to make decisions.

The fourth new principle is collective data governance, recognizing that personal data protection cannot be placed entirely on individuals given their limited capacity and resources. Big data involves mass-scale processing affecting groups and society as a whole, not merely individuals. Collective data governance mandates representing collective interests in data processing decision-making through mechanisms like public consultations, stakeholder participation, and data ethics boards. This principle also obligates controllers to consider societal impacts of data processing beyond individual data subjects. In practice, this principle can be implemented through mandatory social impact assessments before deploying new processing systems and involving civil society organizations in data protection oversight. This approach recognizes privacy as a social value rather than solely an individual right, requiring collective participation for its protection.

The fifth new principle is expanded data portability, which not only grants individuals rights to transfer data between service providers, but also mandates inter-system interoperability. In big data, data value increases when combined with cross-source datasets, making portability a key instrument to prevent excessive data concentration within single entities. Expanded portability requires controllers to provide data in structured, machine-readable formats and support interoperability standards enabling cross-system data exchanges. This principle also includes individual rights to request direct data transmission to alternative service providers without unnecessary technical barriers. Effective data portability enhances digital market competition, allowing users to switch easily to providers offering superior data protection (Laje & Schmidt, 2024). Thus, portability protects individual rights while driving improved protection standards through market mechanisms.

Implementing these new principles requires fundamental changes in data protection institutional structures. Supervisory authorities must be granted broader powers, including technical audit authority over data processing systems and algorithms, as well as powers to impose proportionate sanctions for violations. Supervisory bodies must receive adequate resources, including technical expertise in data and artificial intelligence, as big data oversight requires understanding evolving technologies. Dispute resolution forms a vital component of modern economic activities and data protection, as data use or breaches create conflicting interests, interpretations, and claims among parties (Junaidi et al., 2024).

An adaptive legal framework must also offer swift, low-cost dispute resolution mechanisms, as traditional civil litigation is too slow and expensive for mass-victim data breaches. Mechanisms such as class actions, collective complaints, and mandatory mediation serve as effective alternatives. Furthermore, international cooperation is essential, as data flows unimpeded across borders and breaches frequently span multiple jurisdictions. Adaptive legal frameworks should facilitate cross-border cooperation through mutual recognition mechanisms, mutual legal assistance, and protection standard harmonization.

A personal data protection framework adaptive to big data realities requires combining reinterpreted existing principles with new principles unknown to traditional formulations. Reinterpretation is necessary because existing principles—such as collection limitation, purpose specification, and individual participation—retain core normative value but require adaptation to big data characteristics (Batey, 2022). Adding new principles such as proportionality, precaution, algorithmic fairness, collective data governance, and expanded portability is necessary to address protection dimensions uncovered by old principles. This combination yields a comprehensive normative framework responsive to technological developments without abandoning foundational values. This framework protects individual rights while promoting responsible, equitable data governance. Thus, legal protection of data privacy in the big data era is achievable, provided the law adapts and renews itself to modern demands. The principles formulated in this section serve as guidelines for lawmakers, supervisory authorities, and business operators to realize effective, equitable data protection.

CONCLUSION

The characteristics of big data cause the legal construction of personal data protection which relies on individual consent and purpose limitation to lose its effective reach. The massive volume of data, real-time processing velocity, and variety of formats result in consent mechanisms becoming empty formalities and render the rights of data subjects technically difficult to fulfill. Existing laws lack instruments to address automated data processing, fragmented chains of responsibility, and derivative information generated from data aggregation. Consequently, a reinterpretation of the Fair Information Practice Principles is required, alongside the addition of the principles of proportionality, precaution, algorithmic fairness, collective governance, and expanded portability. The synthesis of reconstructing old norms and establishing new norms yields a normative framework capable of accommodating big data realities without abandoning the foundational values of privacy protection. This framework shifts the protective orientation from entry-point control toward continuous oversight over data flows and utilization. Effective legal protection in the big data era is not an effort to stifle innovation, but rather an effort to ensure that innovation occurs within boundaries that respect human dignity and self-determination. The law must move in tandem with technology while maintaining its normative function as a safeguard of fundamental rights. Therefore, reconstructing the legal framework becomes an indispensable imperative if society is to enjoy the benefits of big data without sacrificing basic freedoms.

For supervisory authorities, the implication is a need to enhance technical capacity and human resources to audit data processing systems and algorithms. Digital business entities gain guidance that compliance with data protection is no longer merely a formal administrative requirement, but demands a comprehensive overhaul in system design and risk management. Civil society and academics obtain a conceptual framework to conduct

oversight and offer critiques of implemented data processing policies. Furthermore, this study paves the way for cross-disciplinary collaboration among legal studies, computer science, and social sciences to formulate solutions to increasingly complex privacy issues.

REFERENCES

- Abdeldayem, A. M. (2023). The Transformative Power of Big Data: Leveraging Insights for Economic Growth, Environmental Preservation, and Societal Well-being. *International Journal of Green Management and Business Studies*, 3(2). <https://doi.org/10.56830/ijgmbs12202303>
- Ali, R., & Darmawan, D. (2023). Big Data Management Optimization for Managerial Decision Making and Business Strategy. *Journal of Social Science Studies*, 3(2), 139-144.
- Aryanto, E. A., Darmawan, D., & Irfan, M. (2024). Business Legal Framework and Challenges in Implementing Social Responsibility Towards Local Communities in Large Investment Projects. *Journal of Social Science Studies*, 4(1), 129-136.
- Aura, A., Priowirjanto, E. S., & Dewi, C. T. I. (2024). Privacy Protection Guarantee for Data Hacking Victims According to Indonesian Law. *Padjadjaran Journal of International Law*, 8(2). <https://doi.org/10.23920/pjil.v8i2.1823>
- Aziz, A., Darmawan, D., Khayru, R. K., & Wibowo, A. S. (2023). Effectiveness of Personal Data Protection Regulation in Indonesia's Fintech Sector. *Journal of Social Science Studies*, 3(1), 23-28.
- Bashori, B., Hardyansah, R., & Darmawan, D. (2024). Legal Analysis of Big Data and Analytics in Preventing Discrimination and Protecting Insurance Customers. *Journal of Social Science Studies*, 4(1), 185-192.
- Batey, C. E. (2022). *Big Data and Data Protection*. In Research Handbook on Privacy and Data Protection Law. Edward Elgar Publishing. <https://doi.org/10.4337/9781786438515.00024>
- Bruening, P. J., & Culnan, M. J. (2016). Through a Glass Darkly: From Privacy Notices to Effective Transparency. *North Carolina Journal of Law & Technology*, 17(4), 515.
- Chynoweth, P. (2008). *Legal Research*. In A. Knight & L. Ruddock (Eds.), *Advanced Research Methods in the Built Environment*. Wiley-Blackwell.
- Dai, Y. (2021). Data-Related Legislation and its Implications for a Country's Competitiveness: The Perspective of the People's Republic of China. *Journal of Data Protection & Privacy*, 4(2), 146-157. <https://doi.org/10.69554/ommt2534>
- Essa, N. E., Mardikaningsih, R., & Ismail, A. B. (2024). Service Transformation Through Technology the Role of AI and Big Data in Service Science. *Journal of Social Science Studies*, 4(1), 193-200.
- Fleckenstein, M., & Fellows, L. (2018). *Big Data and Data Management*. In Modern Data Strategy. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-68993-7_2
- Forgó, N., Händl, S., & Schütze, B. (2017). The Principle of Purpose Limitation and Big Data. In New Technology, Big Data and the Law. Singapore: Springer Singapore. https://doi.org/10.1007/978-981-10-5038-1_2
- Giannopoulou, A. (2020). Algorithmic Systems: The Consent is in the Detail?. *Internet Policy*

- Review, 9(1), 1-19. <https://doi.org/10.14763/2020.1.1452>
- Grant, M. J., & Booth, A. (2009). A Typology of Reviews: An Analysis of 14 Review Types and Associated Methodologies. *Health Information and Libraries Journal*, 26(2), 91-108.
- Hartana, P. C., Mardikaningsih, R., & Halizah, S. N. (2024). Legal Construction and Enforcement Challenges of Credit Card Data Hacking Crimes in E-Commerce Systems. *Journal of Social Science Studies*, 4(2), 503-516.
- Junaidi, J., Hardyansah, R., & Putra, A. R. (2024). Business Dispute Resolution Through Arbitration and Mediation Within the National Legal Framework. *Journal of Social Science Studies*, 4(1), 423-436.
- Karwati, K., Hardyansah, R., & Saktiawan, P. (2024). Legal Analysis of Open Banking and Bank Customer Data Privacy Rights in Indonesia. *Journal of Social Science Studies*, 4(1), 93-104.
- Keskar, V., Yadav, J., & Kumar, A. (2021). Taxonomy of Data Inconsistencies in Big Data. 9(1), 353-363. <https://doi.org/10.17762/ITII.V9I1.139>
- Laje, A., & Schmidt, K. (2024). The Right to Data Portability as a Personal Right. *Laws*, 13(4), 47-47. <https://doi.org/10.3390/laws13040047>
- Lestari, E., & Rasji, R. (2024). Legal Study on Personal Data Protection Based on Indonesian Legislation. *Awang Long Law Review*, 6(2), 471-477. <https://doi.org/10.56301/awl.v6i2.1206>
- Liu, Z. (2024). The Development and Challenges of Data Protection Laws. *International Law Research*, 13(1). <https://doi.org/10.5539/ilr.v13n1p38>
- Mouqdir, S. (2020). Artificial Intelligence in a Privacy-Concerned World: Automated Decision-Making and the GDPR. *Journal of Data Protection & Privacy*, 3(4), 393-402. <https://doi.org/10.69554/bppn8465>
- Mujisulistyo, Y. F., Darmawan, D., & Dirgantara, F. (2024). Reconstruction of the Legal Mechanism for Consumer Rights Recovery Regarding Personal Data Leaks in the Financial Technology and E-Commerce Sectors in Indonesia. *Journal of Social Science Studies*, 4(1), 75-84.
- Osean, R., Negara, D. S., & Putra, A. R. (2024). Digital Transformation in Manufacturing and Legal Aspects of Data Protection and Workers. *Journal of Social Science Studies*, 4(1), 21-32.
- Ryan, P. (2019). Protecting Privacy in a World of Big Data. 85-100. <https://doi.org/10.4324/9781351104845-11>
- Solove, D. J. (2008). *Understanding Privacy*. Harvard University Press.
- Spanca, F., & Salihu, A. (2024, October). *Unveiling the Consequences of Data Breaches: Risks, Impacts, and Mitigation in the Digital Age*. In 2024 International Conference on Electrical, Communication and Computer Engineering (ICECCE). IEEE. <https://doi.org/10.1109/icecce63537.2024.10823432>
- Talapina, E. V. (2023). Big Data and Human Rights: Toward Legal Regulation. *Gosudarstvo i pravo*, (7), 129-138. <https://doi.org/10.31857/s102694520026812-4>
- Tene, O., & Polonetsky, J. (2012). Big Data for All: Privacy and User Control in the Age of Analytics. *Northwestern Journal of Technology and Intellectual Property*, 11(5), 239-273.
- Vedder, A., & Naudts, L. (2017). Accountability for the Use of Algorithms in a Big Data Environment. *International Review of Law, Computers & Technology*, 31(2), 206-224. <https://doi.org/10.1080/13600869.2017.1298547>