

## СЕКЦІЯ 5. СМАРТ-АНАЛІТИКА ТА КІБЕРБЕЗПЕКА БІЗНЕСУ

**Androshchuk D.,**

*Second-cycle student (Master's level),*

**Bilobaba O.,**

*Second-cycle student (Master's level),*

**Savenok A.,**

*Second-cycle student (Master's level),*

**Mozgalli O.,**

*Doctor of Economic Sciences, Professor,*

*Kyiv National Economic University named after Vadym Hetman*

### THE MODALITIES AND DIMENSIONS OF DIGITAL FOOTPRINTS IN THE DIGITAL WORLD

A digital footprint constitutes a highly traceable, durable archive of a data subject's activities across digital systems. This footprint expands through two distinct modalities of data generation: active and passive digital footprints.

*Active Digital Footprints* created through deliberate actions where a user explicitly transmits information, such as publishing social media posts, filling out online forms, or submitting transaction details.

*Passive Digital Footprints* compiled involuntarily without the direct input or immediate awareness of the individual. This is driven by automated infrastructure such as telemetry logging, Internet Protocol (IP) address harvesting, hardware fingerprinting, and persistent network-tracking scripts.

Within threat modeling, these trails are further categorized into public spheres and private spheres, restricted to authorized groups. However, due to systemic data breaches and automated web-scraping, the boundary between public and private footprints remains fragile. Because server-side databases continuously replicate data, transient online actions are converted into permanent, searchable archives beyond the data subject's control.

Unmonitored digital footprints significantly expand an individual's or enterprise's technical attack surface. Cybercriminals systematically harvest open-source intelligence (OSINT), such as public code repositories, professional networks, and historical credential leaks, to execute targeted intrusions completely outside traditional perimeter security controls.

Technical threat vectors include *Contextual Social Engineering*, *Credential Stuffing*, and *Physical & Environmental Vulnerabilities*.

*Contextual Social Engineering* involves threat actors analyzing professional roles and organizational structures to craft highly tailored spear-phishing lures that routinely bypass technical filters.

*Credential Stuffing* means that reused credentials harvested from dark-web repositories are automated to compromise corporate single-sign-on (SSO) gateways.

*Physical & Environmental Vulnerabilities* points out that geotagged images and public records expose physical home addresses, creating doxxing and swatting risks, while unsecured bring-your-own-device (BYOD) hardware introduces corporate network vulnerabilities.

Parallel to malicious threats is a pervasive commercial ecosystem governed by data brokers and ad-tech networks. By aggregating court filings, real-time browsing histories, and session replays, these entities build predictive behavioral profiles.

Modern machine learning algorithms process these unstructured footprints to infer highly sensitive personal traits, creating severe societal gatekeeping mechanisms (Table 1).

Table 1

## Societal Gatekeeping Mechanisms

| Sector        | Mechanism of Algorithmic Stratification                                                                                                                          |
|---------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Employment    | Automated cyber-vetting tools reject candidates based on historical posts or inferred psychological profiles.                                                    |
| Finance       | Alternative credit-scoring models evaluate social networks and purchasing behavior, restricting access to housing and loans.                                     |
| Social Equity | Marginalized populations lacking standard digital histories (e.g., consistent banking records) are systematically excluded by automated decision-making systems. |

Source: compiled by the authors

To mitigate the risks of unmanaged digital shadows, global legislative frameworks have codified strict privacy mandates. Article 25 of the European Union's General Data Protection Regulation (GDPR) establishes a binding legal obligation for *Data Protection by Design and by Default*, requiring organizations to embed structural safeguards directly into their technical architectures (European Parliament & Council of the European Union, 2016). This statute also enforces enhanced protections for minors, recognizing their heightened vulnerability to long-term identity theft and synthetic fraud stemming from parental documentation of their childhoods called "sharenting".

Organizations demonstrate regulatory accountability using standardized operational frameworks *Data Protection Impact Assessments (DPIAs)*, *ISO 31700-1:2023*, and *ISO/IEC 27701:2025*.

*Data Protection Impact Assessments (DPIAs)*. Systematic evaluations mandated under GDPR Article 35 to describe processing activities, assess proportionality, and mitigate structural risks to individual freedoms (European Parliament & Council of the European Union, 2016).

*ISO 31700-1:2023*. The first consumer-oriented standard for embedding Privacy by Design into consumer goods, dictating explicit requirements for user-centric control interfaces, clear privacy notifications, and cradle-to-grave data lifecycle management (International Organization for Standardization, 2023).

*ISO/IEC 27701:2025*. The updated international blueprint for Privacy Information Management Systems (PIMS). The 2025 revision decouples privacy management from core information security management systems (ISMS), enabling a standalone, globally recognized privacy audit (International Organization for Standardization, 2025).

Translating Dr. Ann Cavoukian's 7 *Foundational Principles of Privacy by Design* into executable engineering workflows requires moving beyond abstract policies to rigorous cryptographic and architectural implementations (Cavoukian, 2011).

Engineers must enforce strict backend controls through four distinct parameters: *Purpose Limitation* (collecting data solely for legitimate, predefined purposes), *Collection Limitation* (restricting data intake to bare operational necessities), *Use Limitation* (processing data exclusively within authorized bounds), and *Retention Limitation* (automating deletion once processing objectives are achieved).

To safeguard data during active processing, organizations deploy advanced mathematical frameworks, such as *Homomorphic Encryption* and *Differential Privacy*.

*Homomorphic Encryption* enables mathematical computations to be executed directly on encrypted ciphertext. The system outputs an encrypted result that, when decrypted, matches the output of operations performed on the plaintext, eliminating the need to expose raw data to memory.

*Differential Privacy* is a mathematical framework that provides provable privacy guarantees by adding calibrated statistical noise to dataset queries or machine learning training loops (Cummings et al., 2021; Dwork, 2008). It ensures that the presence or absence of a single individual's record does not significantly alter query outputs (Cummings et al., 2021).

Organizations implement this across two deployment models: *Global Differential Privacy*, where noise is introduced centrally at the query interface, and *Local Differential Privacy*, where individual edge devices obfuscate data before transmitting it to an aggregator (Friedman et al., 2016).

To eliminate central identity tracking, systems utilize the World Wide Web Consortium (W3C) standards for *Decentralized Identifiers (DIDs)* and *Verifiable Credentials (VCs)* (W3C, 2022). This architecture operates via a decentralized three-party model (Issuer, Holder, Verifier) anchored in distributed ledgers.

When a Holder presents a credential to a Verifier, the transaction is cryptographically authenticated via decentralized public key infrastructure (DPKI) without contacting the Issuer, effectively preventing cross-domain tracking. This is enhanced by two key privacy features, *Selective Disclosure* and *Zero-Knowledge Proofs (ZKPs)*.

In the case of the *Selective Disclosure*, users present individual attributes while withholding secondary, identifiable data.

In the case of the *Zero-Knowledge Proofs (ZKPs)*, protocols allow a prover to mathematically demonstrate the validity of a statement to a verifier without revealing any underlying data values.

Despite robust technical frameworks, engineering teams face significant architectural friction. Privacy-by-Design principles are frequently criticized for being overly abstract, making it difficult to map broad concepts onto database schemas or programmatic API contracts.

Furthermore, the rapid expansion of artificial intelligence and machine learning presents fundamental conflicts with privacy engineering:

*The Anti-Minimization Paradigm* means that deep learning models inherently require massive training datasets, directly violating data minimization rules.

*Black-Box Opacity* indicates that the non-linear mathematical structures of neural networks limit auditability and complicate compliance with the legal "right to an explanation."

*Model Memorization* hints at deep neural networks frequently memorize fragments of their training optimization matrices, making the enforcement of the "right to be forgotten" (erasure) exceptionally difficult.

To reconcile AI utility with data privacy, organizations are deploying distributed architectures. *Federated Learning (FL)* enables shared models to train across decentralized edge nodes; devices transmit optimized mathematical weight updates back to a central coordinator while retaining raw data locally. Concurrently, *Synthetic Data Generation* leverages generative models to produce artificial datasets that replicate real-world multivariate correlations. When injected with mathematical noise, these synthetic datasets protect individual identities while remaining highly effective for software testing and model training.

While long-term systemic architecture reforms remain critical, reducing immediate digital exposure requires individuals and organizations to implement personal defensive hygiene in *Email Alias Architectures*, *Two-Hop Split-Trust Networks*, *Hardware Token Authentication*, and *Hardened Environments* (Table 2).

Table 2

The Components of Personal Protective Hygiene for Reducing Immediate Digital Exposure

| Component                            | Principle of operation                                                                                                                                                                                                                                                                                                             |
|--------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Email Alias Architectures</b>     | Utilizing tokenized forwarding services to generate unique, randomized email addresses for individual platform signups, preventing cross-site profile correlation.                                                                                                                                                                 |
| <b>Two-Hop Split-Trust Networks</b>  | Moving away from standard VPN providers – which represent single points of failure – to split-trust network models. This architecture splits connections across independent hosts so that the ingress node sees the user's IP address but not their traffic, and the egress node sees the traffic but never the source IP address. |
| <b>Hardware Token Authentication</b> | Deploying phishing-resistant FIDO2 hardware keys to neutralize credential theft and SIM-swapping exploits                                                                                                                                                                                                                          |
| <b>Hardened Environments</b>         | Migrating endpoints to privacy-centric operating systems to actively block OS-level telemetry and background advertising identifiers                                                                                                                                                                                               |

Source: compiled by the authors

Summarizing the above, the transition from analog identification to digital existence has established persistent digital footprints—unique, durable trails of behavioral data generated through active choices or through automated, passive tracking. There are structural risks associated with digital footprints, including how malicious actors weaponize this data for cyber reconnaissance and how commercial brokers monetize behavioral profiles. To counter these systemic threats, we evaluate Privacy-by-Design (PbD) frameworks, international compliance standards, and technical mitigation strategies. By investigating advanced cryptographic primitives, decentralized architectures, and localized privacy tools, it is possible to demonstrate how organizations and individuals can shift from reactive compliance to proactive, mathematically grounded privacy guarantees.

## References

1. Cavoukian, A. (2011). *Privacy by Design: The 7 foundational principles*. Information and Privacy Commissioner of Ontario. <https://www.ipc.on.ca/wp-content/uploads/2013/09/pbd-implement-7foundational-principles.pdf>
2. Cummings, R., Kaptchuk, G., & Redmiles, E. M. (2021). "I need a better description": An investigation into user expectations for differential privacy. *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, 3037–3052. <https://doi.org/10.1145/3460120.3485252> Cited by: 152
3. Dwork, C. (2008). Differential privacy: A survey of results. In *International Conference on Theory and Applications of Models of Computation* (pp. 1–19). Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-540-79228-4\\_1](https://doi.org/10.1007/978-3-540-79228-4_1)
4. European Parliament and Council of the European Union. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union*, L119, 1–88. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679>
5. Friedman, A., Berkovsky, S., & Kaafar, M. A. (2016). A differential privacy framework for matrix factorization recommender systems. *User Modeling and User-Adapted Interaction*, 26, 425–458. <https://doi.org/10.1007/s11257-016-9177-7> Cited by: 140
6. International Organization for Standardization. (2023). *Privacy by design for consumer goods and services — Part 1: High-level requirements* (ISO Standard No. 31700-1:2023). <https://www.iso.org/standard/81188.html>
7. International Organization for Standardization. (2025). *Information security, cybersecurity and privacy protection — Privacy information management system — Requirements and guidelines* (ISO/IEC Standard No. 27701:2025). <https://www.iso.org/standard/84501.html>
8. W3C. (2022). *Decentralized Identifiers (DIDs) v1.0: Core data model, expression, and formal syntax* (W3C Recommendation). <https://www.w3.org/TR/did-core/>

**Galitsyn V.,**

assistant, PhD Student,

Kyiv National Economic University named after Vadym Hetman

## MULTI-LEVEL RESPONSE IN CYBER THREAT DETECTION SYSTEMS BASED ON IMMUNE RESPONSE PRINCIPLES

The increasing complexity of cyber threats requires not only effective anomaly detection but also flexible, context-aware response mechanisms. Most modern intrusion detection systems (IDS) operate using fixed thresholds, without considering the severity of the threat or the cost of response. In contrast, biological immune systems respond through a multi-level strategy that adapts to the nature of the threat. This study explores how the principles of immune response can be applied to cyber defense, with a specific focus on cost-efficiency and adjusting the response according to the threat level.