Bocconi

WORKING PAPER SERIES

WORKING PAPER
N. 271
APRIL 2026

DATA-DRIVEN SUPERVISION AND THE REDUCTION OF SUPERVISORY BURDEN A CONCEPTUAL AND OPERATIONAL FRAMEWORK FOR NATIONAL COMPETENT AUTHORITIES

Andrea Gentilini

Università Bocconi
BAFFI
Centre on Economics,
Finance and Regulation

Data-Driven Supervision and the Reduction of Supervisory Burden

A Conceptual and Operational Framework for National Competent Authorities

Andrea Gentilini[a][1]

[a] Commission de Surveillance du Secteur Financier, Luxembourg
Email : andrea.gentilini@cssf.lu

March 2026

Abstract : Financial supervisors are doing more with less. Over the past fifteen years, mandates have multiplied — covering prudential soundness, conduct, market integrity, operational resilience, and systemic risk — while supervised populations have grown and budgets have not kept pace. The resulting capacity gap is structural, not cyclical. This paper proposes data-driven supervision (DDS) as a scalable response. DDS converts regulatory reporting data into reproducible risk indicators, composite scores, and automated triage pathways, concentrating expert judgment where it adds most value. Because marginal costs per entity are low and infrastructure is reusable across supervisory missions, DDS generates economies of scale that traditional inspection-based approaches cannot. The paper develops formal frameworks for indicator design and triage, documents how regulatory reporting enables population-scale risk detection, and shows how a shared data infrastructure can serve multiple supervisory objectives simultaneously. Governance challenges — model risk, false positives, preservation of human judgment — are addressed directly, and a phased implementation roadmap for National Competent Authorities is presented. The analysis is candid about limits: DDS depends on data quality, complements rather than replaces traditional supervision, and requires active management of implementation risk.

Keywords: data-driven supervision; supervisory burden; risk-based supervision; supervisory convergence; regulatory reporting; data governance, suptech; model risk governance

JEL Classification Numbers: C81, E58, G18, G28, G23, K23.

---

[1] *The views expressed are those of the author and do not necessarily reflect those of the CSSF. The author thanks Marco D'Errico for helpful comments on earlier drafts, and practitioners at national competent authorities for discussions informing the institutional context. All remaining errors are my own.*

1. Introduction: The Supervisory Capacity Challenge

1.1 The Structural Tension

Financial supervision in the European Union is characterised by a structural tension between expanding mandates and constrained supervisory capacity. Over roughly the past fifteen years, post-crisis reforms and subsequent regulatory initiatives have broadened supervisory objectives beyond traditional prudential soundness and conduct-of-business oversight to encompass systemic-risk surveillance, market integrity, operational resilience, and cross-border supervisory convergence within the Single Market.

National Competent Authorities (NCAs) must now apply these multidimensional mandates to increasingly granular and high-frequency financial activity, including:

- Derivatives and securities financing transactions
- High-frequency and algorithmic trading
- Complex outsourcing and third-party risk chains
- Cross-border group structures and integrated markets

At the same time, budgetary and staffing trajectories have not scaled proportionally with the growth in supervised entities, market innovation, and reporting obligations, creating a persistent supervisory capacity constraint (ESMA, 2025a).

Authorities are asked to supervise more entities, process more data, and monitor more risk channels under tighter expectations of timeliness, proportionality, and legal defensibility, without commensurate increases in inspection capacity or expert headcount.

This imbalance is structural rather than temporary. Traditional supervisory models—relying on periodic on-site inspections, ad hoc data requests, qualitative risk assessments, and bilateral supervisory dialogue—produce high marginal costs per additional supervised entity and exhibit limited scalability when transaction volumes and systemic interconnections increase. International comparative assessments show that such models struggle to provide systematic, documented coverage of large populations and to support cross-jurisdictional consistency, especially when supervisory discretion is exercised under resource pressure (A2ii et al., 2022).

1.2 Policy Context: Simplification and Efficiency

The supervisory burden problem is embedded in a broader EU policy agenda that prioritises simplification and regulatory efficiency. The European Commission's Better Regulation and REFIT initiatives frame simplification not as deregulation, but as achieving regulatory goals at minimum administrative and compliance cost, including through digitalisation and rationalised reporting frameworks.

Strategic diagnoses of Europe's competitiveness—such as the Draghi and Letta reports—identify regulatory fragmentation, duplicative requirements, and administrative complexity as structural impediments to investment, scale formation, and effective Single Market functioning in financial services. Within this framing:

- Inefficient supervisory processes are themselves economic frictions.
- Heterogeneous evidence demands and divergent interpretations across NCAs increase cost and uncertainty for supervised entities.

- Fragmented supervisory practices can undermine the intended benefits of harmonised EU legislation.

Supervisory burden reduction is therefore a strategic necessity rather than a discretionary optimisation exercise. More efficient supervisory processes reduce compliance costs for firms, improve consistency and predictability, and free scarce supervisory resources for higher-value diagnostic and remediation activities.

1.3 Data-Driven Supervision as Strategic Response

Against this backdrop, data-driven supervision has emerged as a central strategic response. DDS is understood as a supervisory operating model in which regulatory reporting data are systematically transformed into quantitative indicators that enable:

- **Continuous population-level monitoring** in place of episodic, sample-based oversight.
- **Risk-based prioritisation and triage** rather than undifferentiated allocation of attention.
- **Structured escalation pathways** in place of ad hoc engagement decisions.
- **Automated evidence compilation** instead of manual, case-by-case assembly.

This reconfiguration can be thought of as a change in the supervisory **production function**: automated population-wide screening and indicator-based triage substitute for a portion of labour-intensive, entity-by-entity analysis[2]. The shift has three interlocking implications:

- **Epistemic:** Supervisory knowledge is increasingly produced through reproducible, quantitative indicators derived from harmonised datasets rather than solely through qualitative assessments.
- **Economic:** Supervisory attention is allocated through structured, risk-sensitive rules rather than primarily through discretionary case selection.
- **Institutional:** Consistency and convergence are operationalised through shared indicators, thresholds, and verification functions rather than only through soft-law coordination and peer review.

1.4 Empirical Momentum

Empirically, the transition toward DDS is already visible across multiple jurisdictions. International surveys document that authorities are investing in supervisory technology (suptech) applications for market surveillance, conduct-risk analytics, and prudential early-warning indicators, frequently leveraging transaction-level reporting frameworks such as EMIR and MiFIR.

ESMA's reports (ESMA, 2024a and ESMA, 2025b) on the quality and use of data show that EU securities regulators increasingly rely on large-scale reporting datasets for:

- Market-abuse detection and market-integrity monitoring
- Data-quality assessment and feedback to reporting entities
- Cross-market and cross-venue analysis of trading patterns

[2] *See International Association of Insurance Supervisors, Access to Insurance Initiative, & Financial Stability Institute. (2022) and EBA (2025)*

Similar trends are evident at other EU institutions and national authorities, including the ECB's data strategies for banking supervision[3], the EBA's work on AML/CFT suptech (EBA, 2025), and national experiments with integrated data platforms and analytics teams[4].

1.5 Research Question

The central research question of this paper is:

How can data-driven supervision reduce the supervisory burden faced by NCAs while simultaneously strengthening risk-based supervision and enabling supervisory convergence?

The thesis advanced is that DDS achieves this by re-engineering the supervisory production function along three dimensions:

1. **Continuous measurement** of supervisory-relevant phenomena based on regulatory reporting.
2. **Quantitative, risk-sensitive prioritisation** using metrics that are explicitly materiality- and uncertainty-aware.
3. **Automated generation of standardised evidence packs** to support engagement and enforcement.

This transformation is not claimed to reduce absolute resource requirements in the short run, given substantial fixed costs of data infrastructure, governance, and skill development. Rather, the argument is that DDS enables a structurally lower marginal supervisory cost per additional supervised entity or risk channel and reduces duplicative information demands on firms, thereby delivering burden reduction in a sense consistent with the EU's efficient-regulation agenda (ESMA, 2025a, 2026 and A2ii et al., 2022.).

From a theoretical standpoint, the paper conceptualises supervision as an inference and control problem under measurement error: supervisors seek to infer latent risk states and systemic externalities from noisy regulatory data and to implement interventions that minimise expected loss subject to legal and resource constraints. Measurement theory provides a framework for understanding how data quality, indicator design, and model specification condition the reliability of supervisory inferences. Decision theory and risk-based supervision principles formalise the allocation problem: scarce supervisory resources must be assigned to entities and activities in a way that reflects probability, impact, and uncertainty, with explicit recognition of non-linearities arising from network-based risk transmission and liquidity amplification. Systemic-risk modelling, particularly in network and multi-layer frameworks, shows that supervisory mismeasurement for highly interconnected or systemically important institutions can generate disproportionately large system-wide welfare losses, underscoring the need for materiality-aware metrics and exposure-weighted indicators within DDS architectures[5].

Operationally, the paper illustrates how regulatory reporting data—using EMIR-style derivatives reporting as a paradigmatic case—can be transformed into a common supervisory language through the design of indicators, verification functions, and decision rules that map quantitative signals to supervisory actions. This mapping is central for three reasons. First, it enables population-scale supervision with targeted depth: all entities are screened continuously, but only a subset triggers intensive engagement, thereby aligning supervisory effort with risk. Second, it provides a concrete

[3] European System of Central Banks / European Central Bank (ECB), Recent advances in modelling systemic risk using network analysis, 2010.

[4] See Banca d'Italia, EMIR data analytics for research, financial stability and supervision, 2024, for a review of national initiatives.

[5] See Aldasoro et al. (2016), ESMA (2024a, 2025b), Cheh et al. (2014), ECB (2010).

mechanism for supervisory convergence: when different NCAs apply the same indicators and decision trees to harmonised datasets, they can interpret regulatory requirements in a more consistent and transparent manner without constraining legitimate differences in enforcement strategy. Third, it supports legal defensibility by generating reproducible, documented evidence that can be reviewed ex post in administrative and judicial proceedings, thereby strengthening due-process guarantees and equal-treatment claims[6].

The contribution of this paper is both conceptual and policy-practical. Conceptually, it integrates measurement theory, decision theory, and systemic-risk modelling into a unified framework for understanding DDS as a structural response to supervisory burden and convergence challenges. Policy-practically, it offers an architecture for metrics and prioritisation that is compatible with existing European reporting frameworks and supervisory mandates, and it articulates a governance and model-risk approach tailored to the needs of NCAs and EU-level authorities. In doing so, the paper advances the proposition that burden reduction, efficient regulation, and supervisory convergence are not competing objectives. Properly designed, DDS can make them mutually reinforcing outcomes of a coherent, data-driven supervisory regime.

2. Conceptual Foundations of Data-Driven Supervision

Data-driven supervision rests on conceptual building blocks that connect supervisory practice to formal inference under uncertainty, measurement theory, decision theory, and systemic-risk modelling. These foundations are not purely academic; they directly inform how NCAs design indicators, prioritise actions, and structure convergence mechanisms around shared data and metrics (FSB, 2020).

2.1 Supervision as Inference Under Measurement Error

Supervision can be formalised as an inference problem in which authorities seek to recover latent supervisory states from imperfect regulatory data. Let X_t denote the true (unobserved) supervisory state of a supervised population at time$t$—including compliance posture, solvency and liquidity risk, conduct behaviour, and operational resilience—and let Y_t denote the set of regulatory reporting outputs and other structured data available to the supervisor.

The relationship[7] can be written as:

$$Y_t = f(X_t;\, \theta) + \epsilon_t \quad , \tag{1}$$

where $f(\cdot)$ represents the transformation induced by the legal reporting framework and its technical implementation, θ captures institutional parameters of reporting processes, and ε_t denotes measurement error (missing values, misclassification, stale timestamps, inconsistent identifiers).

Supervisory decisions are modelled as actions $a \in A$ (for example, maintain baseline monitoring, initiate desk-based review, conduct on-site inspection, impose remedial measures, or open enforcement proceedings) chosen to minimise expected supervisory loss:

$$a^*(t) = arg \min_{a \in A} E[L(a, X_t) | Y_t] \, , \tag{2}$$

[6] *See Choy, A., et al. (2013), Banca d'Italia (2025), EFAMA (2025), ESMA (2025b).*

[7] *See Appendix A for the mathematical notation and formal framework.*

where $L(a,X_t)$ encodes the social loss associated with different combinations of actions and true states (for example, undetected misconduct, systemic events, or over-enforcement)[8].

In this framework, DDS improves supervision along two key margins:

- **Enhancing the informativeness of Y_t** through better data quality management, standardisation, and integration (reducing ε_t)[9].
- **Deploying more sophisticated inference mappings $R(\cdot)$** from data to risk estimates, including non-linear aggregations, network-based measures, and machine-learning models, subject to governance and explainability constraints[10].

From a supervisory design perspective, this implies that DDS must include robust data-quality governance[11], reproducible indicator pipelines[12], and structured use of quantitative outputs as inputs to professional judgement rather than as automatic decision rules[13].

2.2 Measurement Theory and Regulatory Data

Regulatory reporting systems are designed measurement instruments and must be analysed as such. The IMF's Data Quality Assessment Framework (IMF, 2003) and related statistical literature emphasise that the usefulness of data for policy depends on several dimensions:

- **Relevance and validity**: alignment between reported variables and the underlying legal and economic concepts supervisors intend to observe.
- **Accuracy and reliability**: freedom from significant errors and stability over time and across reporters.
- **Coherence and comparability**: internal consistency and harmonisation across datasets, institutions, and jurisdictions.
- **Timeliness and periodicity**: delivery within decision-relevant time frames.

In the EU context, these principles have been progressively imported into supervisory practice, for example through ESMA's EMIR and MiFIR data-quality reports[14] and the ECB's work on risk-data aggregation[15]. DDS internalises this measurement-theoretic perspective in three ways:

---

[8] *This decision theoretic representation underpins the logic of risk-based supervision: when resources are scarce, supervisory intensity should be allocated where the expected marginal reduction in loss is highest, conditional on the information contained in Y_t. See FSB (2020) and FSI (2019).*

[9] *This improves the signal content of regulatory data*

[10] *Formally, let $\hat{R} = R(Y_t)$ denote the vector of supervisory risk metrics extracted from Yt. Because $R(\cdot)$ typically involves non-linear transformations (e.g. netting, exposure aggregation, centrality computations), small errors in inputs—especially identifiers and counterparty flags—can generate large distortions in inferred risk, particularly for systemically important nodes. This sensitivity motivates the strong emphasis in ECB/SSM and FSB work on risk data aggregation and data quality as preconditions for reliable supervision.*

[11] *Elements such as validation rules, reconciliation checks, outlier detection are a first-class component, not a downstream "cleaning" step.*

[12] *Such that a given input dataset Y_t deterministically yields the same $\hat{R}$, enabling auditability and legal defensibility.*

[13] *Supervisory judgement remains central in specifying the loss function $L(\cdot)$ and in interpreting $\hat{R}$, but is increasingly applied after a structured, data driven screening stage.*

[14] *ESMA's data quality reports on EMIR and MiFIR transaction reporting explicitly assess dimensions such as completeness, accuracy, and consistency (e.g. matching between counterparties and trading venues), and link data quality weaknesses to reduced supervisory capability for market abuse detection and risk monitoring.*

[15] *The ECB's and SSM's work on risk data aggregation (RDARR) similarly frames banks' internal data capabilities as a foundation for effective supervision and risk management.*

1. **Design of reporting frameworks:** New or revised templates are assessed not only for legal coverage but for measurability, comparability, and suitability for automated indicator computation.
2. **Data-quality indicators as supervisory metrics:** Persistent data-quality deficiencies (such as high rejection rates or systematic mis-reporting of key fields) are treated as risk signals in their own right.
3. **Information-to-burden ratio:** Reporting frameworks are evaluated in terms of supervisory information delivered relative to reporting burden imposed, leading to an emphasis on re-use and standardisation.

For supervisory practice, this implies that DDS projects must incorporate multi-layer validation, feedback loops to firms, and escalation policies where repeated failures are treated as governance weaknesses rather than purely technical issues (ESMA, 2025b).

2.3 Decision Theory, Risk-Based Allocation, and Proportionality

Within the inference framework above, risk-based supervision can be seen as an application of decision theory under resource and legal constraints[16]. Let R_t denote a vector of risk indicators derived from Y_t, and let $c(a)$ denote the resource cost of supervisory action a. Supervisors seek decision rules $\pi(R_t)$ that:

- Minimise expected supervisory loss $E[L(a, X_t)|Y_t]$
- Satisfy resource constraints $E[c(a)] \leq C$;
- Respect legal principles of proportionality and equal treatment

In practice, this yields a family of triage and prioritisation rules in which:

- Entities above certain risk-score thresholds receive intensified supervision.
- Entities with medium scores are monitored through enhanced off-site analysis.
- Entities with consistently low scores remain under baseline monitoring.

International case studies of suptech-enabled risk-based supervision show that such structured triage frameworks improve alignment between supervisory attention and underlying risk, reducing both under- and over-supervision relative to complaint-driven or size-based selection alone.

For NCAs, the decision-theoretic foundation translates into concrete design choices:

- **Specification of risk scores** (e.g. SPMU-type metrics combining severity, persistence, materiality, and uncertainty);
- **Calibration of thresholds** based on historical outcomes and capacity constraints;
- **Formal stopping rules**, where automated monitoring is deemed sufficient unless indicators breach defined thresholds for a given duration.

These design choices must be documented and periodically reviewed, both to ensure effectiveness and to support legal defensibility in cases where supervised entities challenge the proportionality or consistency of supervisory actions.

2.4 Systemic-Risk Modelling, Transmission, and "Materiality-Aware" DDS

[16] *Appendix B further details the economics of Data Driven Supervision and its cost-structure. In addition, see ESMA (2026) and Toronto Center (2021) for technological innovation implication on financial supervision and more precisely on risk-based supervision.*

Risk-based supervision in a systemically important financial system cannot be purely micro-prudential. Systemic-risk models show that the impact of distress at a given institution depends not only on its standalone risk profile but also on its position in financial networks and on market-wide liquidity conditions (BIS, 2018).

Network-based contagion models demonstrate that losses can propagate along interbank exposures, derivatives positions, and collateral channels in highly non-linear ways: a small shock at a highly central node can generate disproportionately large system-wide losses, while similar shocks at peripheral nodes may be largely absorbed. Liquidity-amplification mechanisms further show how margin calls, collateral requirements, and fire-sale dynamics propagate stress through funding and market-liquidity networks.

For DDS, these findings imply that supervisory metrics must be **materiality-aware**, therefore:

- Indicators should be exposure-weighted and network-adjusted, not purely frequency-based.
- Data elements critical for network reconstruction (for example, counterparty identifiers, UTI/LEI, clearing-member flags) must receive particular data-quality scrutiny.
- Prioritisation frameworks should allocate more weight to risks at entities whose failure or disorderly behaviour would generate large externalities, consistent with ESRB and FSB approaches to systemic-risk monitoring.

A materiality-aware DDS ensures that automation and triage do not come at the expense of missing low-frequency, high-impact systemic events, but instead align micro-supervisory attention with macroprudential risk-transmission channels[17].

Burden reduction through DDS cannot come at the price of blinding supervisors to systemic threats. Effective macroprudential oversight depends on reliable, granular information about individual institutions and, crucially, about the network of exposures and funding links that connect them. Those are precisely the data that DDS pipelines ingest, validate, and transform into materiality-aware and network-sensitive indicators, so the same infrastructure that lowers marginal supervisory cost also strengthens the visibility of system-wide vulnerabilities.

- **2.5 SupTech and DDS as Institutional Technology**

International assessments of suptech emphasise that technological tools can enhance oversight, surveillance, and analytical capabilities, but cannot substitute for sound supervisory judgement[18]. DDS can be seen as the institutionalisation of these suptech capabilities within supervisory operating models and governance structures[19].

---

[17] *Supervisory translation of these results can be made explicit through DDS design:*

- *Systemic importance scores (e.g. adapted from G SII/G SIB methodology) are integrated into triage;*
- *Network based early warning indicators (e.g. sudden increases in centrality or concentration of exposures) trigger cross sectoral or macroprudential escalation.*
- *Stress test inputs are derived from EMIR and MiFIR style datasets, enabling scenario analysis that feeds back into micro supervisory prioritisation.*

In practical terms, a materiality aware DDS ensures that burden reduction through automation and triage does not come at the expense of missing low frequency, high impact systemic events. Instead, it aligns micro supervisory attention with macroprudential risk transmission channels.

[18] *The FSB's report on supervisory and regulatory technology (FSB, 2020) notes that suptech solutions can improve oversight, surveillance, and analytical capabilities, generating near real time indicators of risk and enabling more efficient risk-based supervision. The BIS and IAIS similarly document how integrated platforms and advanced analytics allow supervisors to (i) automate data ingestion and validation; (ii) integrate multiple data sources (regulatory reporting, market data, complaints, social media) and (iii) support risk based prioritisation and thematic work through dashboards and scorecards.*

[19] *See FSB (2020).*

Key institutional features include[20]:

- **Dedicated data and analytics functions** with clear mandates and integration into supervisory decision-making.
- **Formal model governance** for indicators and risk scores, including validation, performance monitoring, and documentation.
- **Training and guidance** to ensure supervisors understand both capabilities and limitations of DDS outputs and can challenge models where necessary.

For NCAs, DDS is therefore not a plug-in solution, but a multi-year institutional change programme aligned with ESMA's and the ECB's data strategies and the broader move towards evidence-based, outcome-focused supervision.

3. Regulatory Reporting as an Enabler of Population-Scale Supervision

Traditional supervisory approaches face fundamental scalability constraints: inspections, qualitative file reviews, and bilateral supervisory dialogue are inherently resource-intensive and episodic, making it difficult to monitor large and heterogeneous populations systematically. As mandates expand to cover prudential, conduct, market-integrity, and operational-resilience dimensions, NCAs often supervise hundreds or thousands of entities with limited staff and budgetary resources (FSB, 2020).

This creates a structural trade-off:

- **Superficial breadth:** broad coverage but insufficient depth per entity.
- **Selective depth:** deep analysis of a subset of entities, with unavoidable blind spots elsewhere.

Regulatory reporting frameworks such as EMIR, MiFIR, COREP, and FINREP fundamentally alter this trade-off by generating population-level datasets that enable NCAs to observe virtually all supervised entities and transactions simultaneously. When leveraged through DDS, these datasets support a **"monitor all, investigate few"** operating model, in which continuous automated screening identifies risk signals and concentrates scarce supervisory resources on the highest-priority cases[21].

3.1 Population-Scale Coverage: "Monitor All, Investigate Few"

In traditional supervision, capacity constraints force supervisors to select a subset of entities for deep review based on priors (for example, size, business model, previous issues) or triggers (for example, complaints, market events), creating blind spots in unsampled parts of the population. DDS eliminates much of this selection problem by enabling continuous, automated monitoring of the entire supervisory population.

Each entity's regulatory reporting is ingested, validated, and processed through a standardised indicator pipeline that generates metrics capturing compliance status, risk exposure, data quality, and trends. These metrics function as a first-layer screen, allowing supervisors to identify[22]:

- Outliers relative to peer distributions
- Persistent deficiencies and unresolved issues

---

[20] *See IMF (2025) and Toronto Centre (2018).*

[21] *See ESMA Reports on Quality and Use of Data for an illustration of EMIR and MiFIR data use, e.g. ESMA (2024a, 2025b).*

[22] *For an illustration of data use see ESMA Reports on Quality and Use of Data, e.g. ESMA (2025b). See also FSB (2020) and World Bank (2018).*

- Emerging risks visible in changes or trends over time

Formally, let $E=\{e_1,e_2,\ldots,e_N\}$ denote the supervised population, and let $M(e,t) \in R^k$ denote a vector of k supervisory indicators computed for entity e at time t. DDS constructs a composite score[23]:

$$Score\ (e,t)\ =\ \sum_{i=1}^{k} \pi_i(t) \cdot g(M_i(e,t))\,, \tag{3}$$

where $\pi_i(t)$ are weights reflecting supervisory priorities (for example, higher weight for systemic institutions or consumer-facing firms), and $g(\cdot)$ is a non-linear transformation (for example, threshold function, percentile rank, or persistence penalty).

Entities are then ranked, and supervisory actions are allocated according to a pre-defined triage ladder:

- **Top tier:** high scores, persistent issues, or high materiality → immediate engagement, on-site inspection, or formal remediation.
- **Middle tier:** medium scores or rising trends → enhanced desk-based monitoring, targeted data requests, or thematic review.
- **Baseline tier:** low scores, stable behaviour, and good data quality → automated monitoring only.

This approach achieves full coverage with targeted depth: every entity is observed continuously, but only a subset receives intensive supervisory attention, optimising scarce expert resources[24].

3.2 Standardised Metrics Reduce Marginal Supervisory Effort

A major source of supervisory burden—for both NCAs and supervised entities—is bespoke, case-by-case interpretation of regulatory requirements. When each supervisory interaction requires ad hoc data requests, custom analysis, and individualised judgement, marginal cost per entity remains high and consistency is difficult to achieve[25].

DDS addresses this problem by establishing **standardised metrics and verification functions** that apply uniformly across the supervised population. For a given regulatory requirement R, a verification function $V_R(Y)$ maps reported data Y into quantitative compliance or risk indicators, enabling supervisors to:

- Compare entities against peers (benchmarking).
- Reproduce calculations over time (auditability).
- Predict likely supervisory outcomes based on historical patterns and calibrated thresholds.

As documented in ESMA's data quality and use reports (ESMA, 2024a), NCAs increasingly rely on common indicators derived from EMIR and MiFIR datasets to assess data quality, identify reporting errors, and detect anomalies such as missing counterparty identifiers, inconsistent valuations, or unusual transaction patterns. These indicators are computed systematically for all entities, creating a shared

---

[23] See Appendix A for more details.

[24] The World Bank's case studies of suptech for market conduct supervision document how automated transaction monitoring systems allowed regulators to screen all firms and detect manipulation patterns that would not have surfaced under manual review. See World Bank (2018).

[25] For a deeper review of the limitations and burdens associated to supervisory reporting requirements see the Fitness Check of EU Supervisory Reporting Requirements (European Commission, 2019), the UK Finance response to the European Commission Fitness Check on Supervisory Reporting (UK Finance, 2019) as well as ECB Simplification of the European prudential regulatory, supervisory and reporting framework (2025).

supervisory language that supports convergence across NCAs and reduces the need for repeated bilateral clarifications.

The economic benefit is straightforward: once an indicator pipeline is designed, validated, and implemented, the marginal cost of applying it to an additional entity or time period is near-zero, whereas manual processes require proportional staff effort for each new case[26] . Table 1 illustrate the comparison between traditional supervision and data-driven supervision across 6 dimensions.

Table 1. DDS contribution to supervisory convergence

Dimension	Traditional supervision	Data-driven supervision
Fixed costs	Low (existing staff, basic IT)	High (data infrastructure, analytics, training)
Marginal cost per entity	High (manual review, bespoke requests)	Low (automated indicator computation)
Scalability	Poor (near-linear cost growth)	Strong (declining average cost with scale)
Consistency	Variable (discretion, resource pressure)	High (standardised metrics, reproducible logic)
Coverage	Selective (sample or complaint-driven)	Population-wide (continuous monitoring)
Evidence generation	Manual (staff compile case files)	Automated (evidence packs from data pipelines)

Source: Author synthesis based on FSB (2020), A2ii, FSI, IAIS (2022), and World Bank (2018) analyses.

This cost structure implies that burden reduction is a dynamic phenomenon: initial DDS implementation may increase total supervisory costs because of investment in infrastructure and skills, but as supervised populations grow and indicators serve multiple missions, cumulative savings and coverage improvements become substantial.

3.3 Evidence Automation Reduces Time-to-Action and Due-Process Costs

Supervisory engagements, remediation programmes, and enforcement actions require documented evidence showing scope, severity, persistence, and legal basis of identified issues. Traditionally, evidence compilation is labour-intensive: staff manually extract transaction records, reconcile data sources, prepare timelines, and draft narratives, often under tight deadlines (Galán and Svetiev, 2025).

DDS reduces this burden through **automated evidence-pack generation**. Once an entity triggers supervisory escalation, the indicator pipeline can easily produce a standardised dossier containing:

- Indicator values and trends over a defined historical window.

[26] *Leading to a near linear cost profile that does not scale efficiently.*

- Record-level examples illustrating anomalies or potential breaches (transactions, missing reports, outlier valuations).
- Peer benchmarks showing how the entity compares to similar firms.
- Data-quality metrics (completeness, rejection rates, reconciliation status) to contextualise reliability.
- Legal anchors linking indicators to specific regulatory provisions (for example, EMIR Article 9, MiFIR transparency obligations).

These evidence[27] packs:

1. **Accelerate supervisory engagement:** Supervisors can quickly review key facts and decide on next steps without manually reconstructing cases.
2. **Enhance legal defensibility:** Standardised, reproducible evidence supports proportionality arguments and reduces the risk of successful appeals based on inconsistent treatment or procedural defects.
3. **Lower due-process costs:** Clear documentation reduces time spent in dialogue, administrative hearings, and judicial review for both NCAs and supervised entities.

Automated evidence generation also improves traceability and accountability, as each pack is time-stamped, versioned, and traceable to specific data inputs and indicator definitions, facilitating audit and internal quality assurance[28].

3.4 Dynamic Supervision: Earlier Detection, Lower Downstream Costs

A structural feature of supervisory burden is that late detection of problems increases remediation and enforcement costs. Issues that persist undetected typically escalate in severity, generate consumer harm or market distortions, and require more intensive interventions (for example, repeated on-site inspections, formal remediation plans, enforcement proceedings).

DDS supports earlier detection through continuous monitoring and trend analysis, allowing supervisors to identify deterioration before it becomes critical:

- **Change-point detection algorithms** can flag sudden increases in data-quality errors, transaction volumes, or risk metrics.
- **Persistence tracking** records how long specific issues have been present, escalating entities with repeated or unresolved deficiencies.
- **Peer-divergence indicators** highlight entities whose behaviour deviates significantly from peers, signalling potential business-model shifts or emerging risks.

By intervening earlier—often through targeted supervisory letters or guidance rather than full enforcement—NCAs can achieve compliance at lower cost to both authorities and firms, reducing the volume and intensity of cases that escalate into resource-intensive enforcement[29]. Evidence from the

---

[27] ESMA's work on EMIR and MiFIR reporting illustrates this dynamic: NCAs increasingly use automated data-quality dashboards and error reports to provide feedback to supervised entities and, where necessary, to support formal follow-up actions or sanctions for persistent mis-reporting. See ESMA (2024a, 2025b).

[28] This transparency strengthens supervisory legitimacy and aligns with Better Regulation principles emphasising proportionality, due process, and equal treatment.

[29] This "supervisory triage" reduces the volume of cases requiring intensive engagement and shortens the time-to-resolution for those that do escalate, thereby lowering the overall burden on supervisory systems.

economics of regulation supports the view that early-stage interventions are more cost-effective than late-stage enforcement, both in resource terms and in social-welfare outcomes (Word Bank, 2025)

3.5 Supervisory Convergence Through Shared Verification Logic

Fragmentation in supervisory interpretation and evidence standards imposes multiplicative compliance burdens on cross-border entities and complicates cooperation among NCAs. When different authorities define and detect risks in divergent ways, firms must reconcile heterogeneous expectations and parallel data requests, and joint supervisory action is more difficult.

DDS mitigates this fragmentation by enabling NCAs to apply the **same indicators and verification functions** to the same harmonised datasets (EMIR, MiFIR, and others), achieving convergence as an analytical property of the supervisory system rather than solely through procedural coordination[30]. When different NCAs compute the same metrics and observe the same quantitative signals, their supervisory interpretations at the screening and prioritisation stage are more likely to align, even if enforcement outcomes remain proportionate to local context.

ESMA's peer-review work documents that authorities using common data-quality indicators and risk dashboards exhibit greater consistency in follow-up actions and that divergences are easier to diagnose and address systematically[31]. This form of convergence improves supervisory effectiveness and reduces compliance burden, as firms face more predictable and harmonised supervisory expectations across jurisdictions.

4. Quantitative Risk-Based Supervisory Allocation: Metrics, Triage, and Decision Architecture

Risk-based supervision is a foundational principle of modern financial oversight and is recognised by ESMA, IOSCO, the ECB, and national authorities as essential in resource-constrained environments[32]. Implementing risk-based supervision at scale requires explicit, quantitative frameworks for measuring risk, ranking entities, and triggering supervisory actions, rather than relying solely on qualitative judgement.

DDS operationalises risk-based supervision through supervisory metrics, triage models, and decision architectures that link data inputs to supervisory outputs in a structured, reproducible manner. This section sets out the analytical foundations, describes the data architecture that supports them, and explains the governance and auditability requirements that make DDS outputs defensible and trustworthy. Appendix D further details data architecture and validation layers.

4.1 Supervisory Metrics as Risk Signals: Formalising Severity, Persistence, Materiality, and Uncertainty

Supervisory metrics are quantitative indicators derived from regulatory data that function as proxies for compliance quality, risk exposure, governance robustness, and systemic relevance. In DDS, metrics replace or augment purely qualitative assessments as the first-layer risk screen, enabling continuous population-wide monitoring (FSB, 2020).

---

[30] Appendix C details supervisory convergence. Furthermore, see Galán and Svetiev,(2025) and ESMA (2025d, 2025e).

[31] See ESMA Peer Review Methodology (2020, rev. 2024), ESMA (2026) Principles on Risk Based Supervision, and EBA Decision adopting a methodology for the conduct of peer reviews (2020)

[32] See IOSCO (2009), ECB Banking Supervision – Supervisory Methodology 2025; ESMA Strategy 2023-2028 (2022).

Within this framework, the **SPMU** (Severity, Persistence, Materiality, Uncertainty) concept structures metric design across four dimensions[33]:

1. **Severity** (s_i): the magnitude or intensity of the risk signal, such as the size of a capital shortfall, the volume of suspicious transactions, the extent of valuation discrepancies, or the number of missing regulatory reports.
2. **Persistence** (p_i): the duration over which an issue has been observed, typically measured as the number of consecutive reporting periods in which a threshold is breached or an anomaly is detected. Persistence signals structural rather than transient problems and warrants escalation.
3. **Materiality** (m_i): the systemic or market significance of the entity, calibrated using measures such as total assets, market share, network centrality (e.g. derivatives exposures, interbank funding), or designation as systemically important (G-SII, O-SII). Materiality-weighting ensures that supervisory attention reflects not only standalone risk but also potential externalities.
4. **Uncertainty** (u_i): a measure of data quality, model confidence, or analytical ambiguity (e.g. high rejection rates in reporting, incomplete identifier coverage, contradictory signals across datasets). High uncertainty may justify additional data collection or conservative escalation even when severity appears moderate.

A composite **prioritisation score** can then be constructed as:

$$Priority\,(e, \mathrm{t}) \;=\; \sum_{i=1}^{k} \pi_i(t) \cdot g(M_i(e,t))\;, \tag{4}$$

where $M_i(e,t)$ is metric i for entity e at time t, $\pi_i(t)$ are policy weights reflecting supervisory priorities (e.g. higher weight for consumer-facing firms, systemic institutions, or specific risk types under thematic focus), and $g(\cdot)$ is a non-linear transformation function (e.g. threshold indicator, percentile rank, or penalty for repeated breaches)[34]. Table 2 provides illustrative examples of supervisory metrics based on EMIR- and MiFIR-style regulatory reporting.

Table 2. Supervisory metrics based on EMIR- and MiFIR-style reporting, illustrative example

Metric category	Indicator	SPMU dimension(s)	Source data
Data quality	Rejection rate (% of reports failing validation)	Severity, Persistence	EMIR/MiFIR submission logs
Data quality	Counterparty reconciliation breaks (% unmatched UTIs)	Severity, Uncertainty	EMIR trade repository data
Compliance	Missing or late reports (count, % of expected submissions)	Severity, Persistence	EMIR/MiFIR submission logs

[33] *For a more detailed review see ESRB (2020, 2025), FSB (2020), ECB Banking supervision (2024),*

[34] *This formalisation aligns with the ECB's Supervisory Review and Evaluation Process (SREP) methodology for banking supervision, which combines anchor scores (quantitative, automated computations) with constrained supervisory judgement (qualitative adjustments within pre-defined ranges) to produce risk ratings and capital add ons. Similarly, IOSCO's guidelines on risk based supervision emphasise that risk ratings should integrate both quantitative indicators (financial ratios, transaction volumes, complaints) and qualitative factors (governance, complexity), with explicit documentation of weighting and scoring logic to ensure consistency and transparency.*

Market integrity	Valuation outliers (transactions priced >X% from market benchmark)	Severity, Materiality	MiFIR transaction reports
Market integrity	Unusual trading patterns (volume spikes, clustering, timing anomalies)	Severity, Uncertainty	MiFIR transaction reports
Systemic relevance	Derivatives exposure concentration (% of market, centrality in network)	Materiality	EMIR exposure data
Prudential	Margin call escalation (frequency, size relative to capital)	Severity, Materiality	EMIR collateral/margin data
Conduct	Consumer complaint rate (per £ AUM or per client)	Severity, Persistence	Conduct reporting, complaints DB

Source: Author synthesis based on ESMA (2025b, 2025c), ECB Banking Supervsion (2024), IOSCO (2009), and ESRB (2020, 2025) frameworks.

These metrics are **not supervisory decisions**, but **inputs to structured triage**. They provide supervisors with a quantitative basis for ranking entities and identifying outliers, but supervisory judgement remains necessary to interpret context, assess root causes, and determine proportionate actions.

4.2 Risk-Based Triage and Escalation: From Metrics to Actions

Triage is the process by which metrics are transformed into differentiated supervisory actions. The core insight is that not all entities require the same supervisory intensity at all times; by stratifying the population according to risk signals, NCAs can deploy scarce expert resources where expected marginal loss reduction is highest.

A standard triage model defines **escalation tiers** linked to ranges of the composite priority score:

- **Tier 1 (Baseline monitoring)**: Entities with low priority scores, stable trends, and good data quality remain under **automated population monitoring (i.e. routine off-site supervision)**. Supervisory interaction is limited to standard reporting cycles and data-quality feedback loops.
- **Tier 2 (Enhanced monitoring)**: Entities with moderate scores or emerging adverse trends are subject to **desk-based supervisory review**, which may include targeted data requests, peer benchmarking, or structured supervisory dialogue to understand the drivers of the risk signal.
- **Tier 3 (Intensive supervision)**: Entities with high scores, persistent issues, or high materiality receive **on-site inspections, formal remediation plans, or thematic reviews**. Supervisory engagement is frequent, documented, and linked to explicit milestones for risk reduction.
- **Tier 4 (Enforcement process)**: Entities that fail to remediate despite Tier 3 engagement, or that exhibit serious misconduct or governance failures, are escalated to **formal enforcement proceedings**, supported by automated evidence packs and legal analysis.

This **escalation ladder** mirrors the logic of the ECB's SREP, in which higher overall SREP scores (1 = low risk, 4 = high risk) trigger more intensive supervisory measures, higher capital requirements (Pillar 2 Requirement, P2R), and, where necessary, early intervention or enforcement. The SREP methodology explicitly documents anchor scores, adjustment rules, and thresholds to ensure horizontal consistency across supervised banks and to support accountability in internal and external review[35].

IOSCO's guidelines similarly emphasise that **triage rules must be transparent, documented, and periodically reviewed** to avoid arbitrary treatment and to enable supervised entities to understand how their risk profile influences supervisory intensity[36]. Academic work on monitoring and enforcement further shows that structured, rule-based screening mechanisms outperform discretionary selection in large populations, particularly when information is dispersed and resource constraints are binding (see Angelova et al., 2023).

4.3 Supervisory Data Architecture: From Reporting to Actionable Signals

Implementing quantitative risk-based supervision at scale requires a robust supervisory data architecture that ingests, validates, transforms, and delivers data to decision-makers in a timely, reliable, and auditable manner. This architecture can be conceptualised as a multi-stage pipeline:

1. **Ingestion**: Regulatory reports (e.g. EMIR, MiFIR, COREP, FINREP) are received from supervised entities and trade repositories, typically via standardised technical protocols (ISO 20022 XML, FIX, API endpoints).
2. **Canonicalisation**: Raw inputs are transformed into a **canonical data model** with standardised identifiers (LEI, ISIN, UTI), harmonised taxonomies, and consistent timestamps, enabling cross-dataset integration and reproducibility.
3. **Validation and reconciliation**: Data pass through a **rule engine** that applies schema validation, business-rule checks, and cross-dataset reconciliations (e.g. matching EMIR counterparty reports, reconciling MiFIR transaction reports with transparency data).
4. **Feature engineering**: Validated data are aggregated and transformed into **supervisory metrics** (indicators, scores, peer benchmarks, trend signals) using pre-defined computational logic and statistical models.
5. **Scoring and triage**: Metrics feed into composite prioritisation models that rank entities and assign them to escalation tiers according to triage rules.
6. **Case management and evidence packs**: For entities requiring engagement, the system automatically generates **evidence dossiers** containing indicator histories, record-level examples, peer comparisons, and legal anchors.
7. **Dashboards and supervisory outputs**: Supervisors access results via interactive dashboards, alerts, and reports that support both population-level oversight and entity-specific deep dives.

This architecture is analogous to the RDARR expectations imposed on banks by the ECB and BCBS[37], but applied to supervisory authorities' own infrastructures[38].

---

[35] *See ECB Banking Supervision (2025)*

[36] *See IOSCO (2009, 2017).*

[37] *Which emphasise accuracy, completeness, timeliness, and auditability of risk data as foundational to sound risk management and effective supervision.*

[38] *DDS cannot be reliable if data pipelines at NCA level are opaque, untested, or non-reproducible.*

4.4 Governance, Versioning, and Auditability Requirements

For DDS to be sustainable and legally defensible, it must be embedded in strong model and data governance frameworks[39].

Key requirements include:

1. **Versioning**: All components of the DDS pipeline—data schemas, validation rules, indicator definitions, scoring weights, and triage thresholds—must be **versioned and change-controlled**, such that any output can be traced to a specific configuration and re-executed if needed.
2. **Traceability and drill-down**: Supervisors must be able to **drill down** from aggregate scores to the underlying metrics, and from metrics to the specific data records that drove them, enabling full transparency and supporting supervisory dialogue with entities.
3. **Reproducibility**: DDS computations should be **deterministic.** Given the same input data and pipeline version, the system must produce identical outputs, ensuring that results are auditable and not subject to hidden randomness or undocumented judgement.
4. **Model validation and performance monitoring**: Indicators and scoring models must undergo periodic **validation** to assess their predictive accuracy, stability, and fairness (e.g. whether they systematically over- or under-weight certain business models or entity types). Champion–challenger approaches, back-testing, and peer review are standard techniques.
5. **Documentation and explainability**: All models and indicators must be **documented** in accessible language, with clear statements of purpose, methodology, data sources, limitations, and known biases, consistent with principles of explainable AI and supervisory accountability.
6. **Human oversight and challenge**: Automated outputs are **inputs to supervisory decision-making, not substitutes for it**. Supervisors must retain the authority and capacity to challenge, override, or contextualise model outputs based on qualitative knowledge, evolving risks, or proportionality considerations.

These requirements align with the ECB's SREP methodology[40] and ESMA's expectations for data strategies and convergence[41], which stress transparency, consistency, and ongoing evaluation.

4.5 Materiality-Aware Metrics and Systemic-Risk Integration

Burden reduction through DDS cannot be pursued in isolation from financial-stability objectives. If automation and triage simply made supervision cheaper while increasing the chance of missing systemic vulnerabilities, they would be a false economy. In practice, the opposite should hold: effective macroprudential oversight requires exactly the kind of reliable, granular information on individual entities and their interconnections that DDS pipelines are designed to produce, and this information can be embedded directly into materiality-aware, systemic-risk-informed supervisory metrics. As discussed in section 2.4, burden reduction through DDS must not come at the expense of missing systemic risks. Supervisory metrics and triage rules must therefore be materiality-aware and systemic-risk-informed, drawing on ESRB (2018, 2020 and 2025) and FSB (2020) frameworks. In this sense, DDS should ease

[39] *See FSB (2020, 2024), ECB SREP Methodology (2024).*

[40] *Automated anchor scores provide a common starting point, but supervisors can adjust within pre-defined ranges, with each adjustment documented and subject to horizontal consistency review.*

[41] *Common tools and methodologies must be transparent, well-governed, and subject to ongoing evaluation to maintain trust and effectiveness. See Appendix C for more details on convergence.*

supervisory workload while simultaneously sharpening, not weakening, supervisors' view of systemic vulnerabilities, because it builds on exactly the granular, interconnected data that macroprudential oversight requires. Key principles includes:

- **Exposure-weighted indicators**[42]: Metrics should reflect not only the frequency or severity of issues but also the scale of exposures and the centrality of entities within financial networks (e.g. derivatives exposures, funding interconnections, common asset holdings).
- **Interconnectedness and contagion measures**: Network-based indicators (e.g. eigenvector centrality, CoVaR, fire-sale spillover indices) capture the potential for distress transmission and amplification, allowing supervisors to prioritise entities whose failure would have disproportionate system-wide effects.
- **Liquidity and funding risk indicators**: Monitoring of maturity mismatches, margin dynamics, collateral adequacy, and portfolio similarity across institutions provides early-warning signals of procyclical stress and liquidity spirals.
- **Dynamic thresholds and stress scenarios**: Triage thresholds can be **state-contingent**, increasing supervisory intensity during periods of elevated systemic stress or when early-warning indicators breach pre-defined levels.

For instance, an EMIR-based DDS might assign higher priority to entities that are central nodes in the derivatives network, have large concentrated exposures to other systemic institutions, or exhibit persistent data-quality problems in counterparty identifiers that directly impair systemic-risk monitoring.

This **materiality-aware triage** ensures that DDS allocates supervisory attention not only to individual compliance or risk issues, but also to entities whose distress or disorderly behaviour could generate significant externalities, aligning micro-supervision with macroprudential objectives.

5. Supervising Large Populations with Consistency: Coverage, Comparability, and Convergence

A defining benefit of DDS is the ability to supervise entire populations systematically and consistently, rather than relying solely on sample-based oversight or complaint-driven intervention. This addresses three core challenges: comprehensive coverage, comparability across entities and over time, and supervisory convergence across jurisdictions.

5.1 Population-Wide Supervision: Eliminating Blind Spots

Traditional approaches force NCAs into a choice between selective depth[43] and superficial breadth[44], each creating blind spots or insufficient depth. DDS breaks this trade-off by enabling population-wide continuous monitoring combined with risk-based depth: every entity, regardless of size, business model,

---

[42] Exposure-weighted scoring is detailed in Appendix A.

[43] Intensive supervision of a subset (often based on size, prior issues, or complaints), leaving the remainder under minimal or reactive oversight and creating blind spots where risks can accumulate undetected.

[44] Light-touch monitoring across all entities, with insufficient depth to detect complex or emerging risks.

or prior history, is screened against common indicators derived from regulatory reporting, but only a subset receives intensive attention[45].

The operational logic can be summarised as **"monitor all, investigate few"**:

- All entities are scored and ranked continuously.
- Only those triggering risk thresholds or exhibiting persistent problems are escalated.
- Smaller or previously low-profile entities still appear in dashboards and can surface through data-driven signals.

This reduces blind spots and allows supervisors to detect deterioration at entities that would otherwise have remained off radar in complaint- or size-driven models[46].

5.2 Consistency, Comparability, and Reproducibility

Quantitative, data-driven metrics promote consistency, comparability, and reproducibility—core requirements for fairness and convergence.

- **Consistency:** Similar entities in similar circumstances are assessed using identical criteria and thresholds, reducing reliance on unstructured discretion at the initial screening stage and supporting equal treatment under the law[47].
- **Comparability:** DDS generates peer distributions for each indicator, allowing benchmarking of an entity against its peers and over time[48].
- **Reproducibility:** Given the same input data and the same indicator definitions, identical outputs are produced, which is critical for internal audit, external review, and legal defensibility[49].

These properties are enhanced when NCAs share analytical infrastructures or at least share indicator definitions and verification logic. Academic work on peer review and supervisory convergence emphasises that consistency is more robust when embedded in shared analytical infrastructures rather than achieved solely through procedural coordination or soft-law guidance. Galán and Svetiev's analysis of ESMA's peer-review methodology shows that convergence efforts have evolved from formal legal compliance checks toward deeper qualitative assessments of supervisory practices, with an increasing recognition that common tools, metrics, and evidence standards are necessary to achieve consistent supervisory outcomes across diverse jurisdictions[50].

IOSCO's principles on cross-border cooperation (IOSCO, 2010) similarly note that divergent supervisory approaches—even under harmonised legal frameworks—create fragmentation, regulatory

---

[45] This universal coverage is achieved through automated data pipelines that ingest, validate, and transform regulatory data into supervisory metrics without requiring manual intervention for routine cases.

[46] This model mirrors suptech implementations documented by the World Bank (2018), FSB (2020), and BIS (2018), in which automated transaction monitoring and risk dashboards allow regulators to screen entire markets while concentrating human expertise on anomalies, outliers, and high-risk cases.

[47] Automated indicator pipelines apply identical validation rules, aggregation logic, and scoring formulas to all entities, ensuring that differences in supervisory signals reflect actual differences in compliance or risk rather than heterogeneous supervisory practices or resource availability.

[48] DDS generates peer distributions for each indicator, enabling supervisors to identify outliers and to contextualise individual entity performance relative to norms. This benchmarking capability is central to ESMA's data-quality work on EMIR and MiFIR (ESMA, 2025).

[49] This determinism is achieved through versioning of indicator logic, documentation of data transformations, and traceability of all computations back to source records. Reproducibility supports due process by enabling supervised entities and external reviewers to verify how scores and escalation decisions were derived, reducing the risk of arbitrary or opaque supervisory actions.

[50] See Galan and Svetiev (2025).

arbitrage risks, and compliance inefficiencies, and that enhanced **information sharing and common analytical frameworks** are essential for effective global oversight.

5.2 Indicator Families for Quantitative Risk-Based Allocation

DDS relies on families of indicators capturing multiple supervisory dimensions of supervisory concen. While specific indicators vary by mandate and dataset, the following categories are broadly applicable across regulatory reporting regimes[51]:

1. **Compliance and data-quality integrity**: metrics capturing the completeness, accuracy, and timeliness of regulatory reporting, such as rejection rates, missing fields, late submissions, reconciliation breaks, and identifier inconsistencies. Persistent data-quality failures are supervisory signals in their own right, often indicating weak governance or systems.
2. **Behavioural compliance risk**: indicators of rule breaches (e.g. OTC derivative contracts not timely confirmed), non-compliance patterns, recurrence of issues, and governance weaknesses (e.g. repeated failures to remediate, inadequate internal controls).
3. **Materiality and systemic relevance**: exposure-weighted metrics, market share, network centrality (e.g. in derivatives or interbank funding networks), and systemic importance designations (G-SII, O-SII). These ensure that supervisory prioritisation reflects not only standalone risk but also potential externalities.
4. **Operational resilience signals**: indicators of operational fragility, such as system outages, reporting latency spikes, incident patterns, and concentration in critical third-party service providers. These metrics are increasingly relevant as operational risk and cyber resilience rise in supervisory importance.
5. **Market integrity signals**: transaction-level anomalies, valuation outliers, suspicious trading patterns (clustering, timing, volume spikes), and cross-venue inconsistencies, derived from MiFIR transaction reporting and transparency data.

The EMIR reporting framework provides a particularly strong foundation for DDS because its **dual-sided reporting** (both counterparties report the same transaction) enables **reconciliation and truth-testing**: discrepancies between counterparty reports signal data-quality issues, potential misreporting, or substantive disagreements about contractual terms. While this reconciliation capability is unique to certain frameworks, the conceptual structure—translating regulatory data into risk signals through pre-defined indicator families—is portable across supervisory mandates[52].

5.3 The SPMU Framework: Formalised Composite Score

To operationalise risk-based prioritisation at scale, DDS constructs a composite **prioritisation score** that integrates multiple indicator dimensions. The **SPMU framework** (Severity, Persistence, Materiality, Uncertainty) provides a structured approach.

Let e denote a supervised entity and T a supervisory observation window. For each risk or compliance indicator i, define:

- s_i: **severity**, measuring the intensity or magnitude of the issue (e.g. number of breaches, volume of affected transactions, size of capital shortfall);

---

[51] *See ESMA Report of data quality and use of (2025), ESMA Principles on third-party risks supervision (2025), ESRB (2025)*

[52] *See CSSF-CAA Report (2024).*

- $p_i(e,T)$: **persistence**, capturing the duration or recurrence of the issue across reporting periods (e.g. number of consecutive periods in which a threshold is breached);
- $m_i(e,T)$: **materiality**, reflecting the systemic or market significance of the entity (total assets, market share, network centrality);
- $u_i(e,T)$: **uncertainty**, representing data quality, model confidence, or analytical ambiguity (e.g. high rejection rates, incomplete identifiers, conflicting signals);

Furthermore, define $\pi_i(e,T)$ as policy weight or **prevalence**, adjusting for current supervisory priorities (e.g. higher weight for consumer-facing activities, cyber resilience, or thematic focus areas).

The composite prioritisation score is then:

$$DDSScore\,(e,\mathrm{t}) = \sum_i s_i \cdot g(p_i(e,t)) \cdot h(m_i(e,t)) \cdot u_i(e,t) \cdot \pi_i(e,t)\,, \tag{5}$$

where $g(\cdot)$ and $h(\cdot)$ are non-linear transformation functions that penalise persistence (e.g. exponential or threshold-based) and weight materiality (e.g. systemic importance tiers).

This formulation achieves several objectives:

- **Multidimensionality**: the score integrates standalone risk, systemic externalities, and data reliability, avoiding over-reliance on any single dimension.
- **Transparency and explainability**: each component is defined, documented, and traceable, supporting supervisory dialogue and legal defensibility.
- **Policy flexibility**: weights π_i can be adjusted to reflect evolving supervisory priorities (e.g. increased focus on operational resilience, climate risks, or retail conduct) without requiring wholesale redesign of the indicator pipeline.

The SPMU score provides a **common language** for supervisory prioritisation across different business models, jurisdictions, and supervisory missions, functioning as a structured input to triage decisions rather than an automated decision rule.

5.4 Decision Rules: Structured Escalation Gates

Once entities are scored, supervisory actions are allocated according to pre-defined **escalation tiers**. A stylised decision tree can be formalised as follows (aligned with Section 4.2):

Let $A \in \{0,1,2,3,4\}$ denote supervisory action intensity:

- $A = 0$: **Baseline monitoring** (automated only, no active engagement).
- $A = 1$: **Enhanced off-site monitoring** (desk-based review, data requests, supervisory letter).
- $A = 2$: **Structured engagement** (on-site inspection, remediation plan, enhanced reporting).
- $A = 3$: **Intensive supervision** (formal measures, capital add-ons, restrictions).
- $A = 4$: **Enforcement interface** (sanctions, license withdrawal, public disclosure).

Decision gates based on SPMU scores and specific indicator thresholds:

Gate 1 (Foundational data integrity):
If foundational data-quality indicators (e.g. rejection rate, reconciliation break rate, missing LEI/UTI) exceed a critical threshold τ_1 , then escalate immediately to $A \geq 1$ (structured engagement to restore data integrity).

Gate 2 (Persistence of critical issues):
If compliance or risk issues persist for more than H periods (e.g. three consecutive reporting cycles), then escalate to $A \geq 2$ (formal remediation plan with timelines and milestones).

Gate 3 (Materiality and systemic relevance):
If entity materiality score $m_i(e,T)$ exceeds threshold M (e.g. systemic importance designation, top decile of market exposure), then apply higher escalation tier $A \geq 2$ or intensify existing engagement, with coordination to macroprudential authorities if appropriate.

Gate 4 (Inadequate remediation response):
If entity fails to implement agreed remediation within specified period, or if supervisory dialogue reveals governance failures or misconduct, then escalate to enforcement interface $A = 4$, with automated evidence pack generation supporting formal proceedings.

This decision-tree structure **standardises the escalation pathway**, reducing ad hoc supervisory discretion at the triage stage and ensuring that entities understand how their risk profile translates into supervisory intensity. The approach is consistent with the ECB's SREP methodology, which explicitly defines trigger points, escalation criteria, and documentation requirements to support horizontal consistency and legal defensibility.

5.5 Supervisory Convergence Through Common Measurement

Supervisory convergence has traditionally been pursued through legal harmonisation, soft-law guidelines, peer reviews, and coordination fora. While necessary, these tools have proven insufficient to achieve consistent supervisory outcomes, as ESMA's peer-review work and IOSCO's cross-border cooperation studies document[53].

DDS offers a **complementary convergence mechanism** grounded in **shared measurement and interpretation**. When NCAs apply the **same verification functions** $VR(Y)$ to the **same harmonised datasets** (e.g. EMIR, MiFIR), they generate **identical supervisory signals** for identical data patterns, thereby converging at the level of **risk detection and prioritisation** even if enforcement outcomes differ based on proportionality and local context.

This functional view of convergence — convergence as a **property of the analytical infrastructure** rather than solely a procedural aspiration — has several advantages:

1. **Operational rather than aspirational**: Convergence is embedded in the tools NCAs use daily, not limited to periodic reviews or coordination meetings.
2. **Scalable and continuous**: Common indicators and dashboards operate in real time across all entities, supporting ongoing alignment rather than episodic harmonisation efforts.
3. **Transparent and evidence-based**: When NCAs discuss supervisory cases, they reference common metrics and peer benchmarks, enabling structured dialogue and reducing discretionary divergence.
4. **Consistent with national discretion**: Convergence at the screening and triage stage does not constrain enforcement proportionality or policy choices, preserving subsidiarity while reducing fragmentation in risk assessment.

[53] *See IOSCO (2010), ESMA (2025d, 2025f).*

ESMA's strategy[54] explicitly frames supervisory convergence as requiring a **common supervisory culture** and **shared tools and methodologies**, recognising that legal harmonisation alone cannot ensure consistent outcomes. The evolution of ESMA's peer-review methodology—from compliance checklists to qualitative assessments of supervisory practices and root-cause analysis of divergences—reflects this insight.

Similarly, IOSCO's principles on cross-border cooperation emphasise that **mutual trust and effective information sharing** depend on NCAs having confidence in each other's supervisory capabilities and evidence standards, which in turn depends on **common analytical frameworks and data quality**. FSB (2025) and IOSCO (2025) works on crypto-assets further highlights the risks of regulatory divergence and the need for **enhanced coordination, common definitions, and consistent oversight approaches** to prevent arbitrage and ensure global financial stability.

Table 3 summarises how DDS contributes to convergence at different stages of the supervisory process:

Table 3. DDS contribution to supervisory convergence

Supervisory stage	Traditional convergence tools	DDS contribution
Legal framework	Directives, Regulations, Level 2/3 acts	Not directly applicable (legal convergence is prerequisite)
Risk detection	Guidelines, Q&As, peer reviews	**Common indicators** and **verification functions** applied to harmonised datasets → identical signals
Prioritisation	Peer pressure, best-practice sharing	**SPMU scoring** and **triage rules** provide quantitative basis for consistent resource allocation
Evidence compilation	Ad hoc supervisory practice, case studies	**Automated evidence packs** ensure reproducible, standardised documentation
Supervisory dialogue	Colleges, bilateral contacts	**Common dashboards** and **peer benchmarks** enable structured, data-grounded discussions
Enforcement	Peer review of sanctions, ESMA opinions	Preserved national discretion, but evidence base is more consistent and defensible

Source: Author synthesis based on ESMA (2022), IOSCO (2009, 2010), and FSB (2020) frameworks.

Convergence at the measurement and screening stage does not eliminate national discretion in enforcement but reduces unjustified divergence in how similar risks and behaviours are initially detected and prioritised.

[54] See ESMA (2022).

5.6 Burden Reduction Through Convergence

A frequently overlooked dimension of burden reduction is the **reduction in duplicative and inconsistent supervisory demands on cross-border entities**. When NCAs apply divergent indicators, thresholds, and evidence requirements, firms face multiplicative compliance burdens: they must reconcile heterogeneous supervisory interpretations, respond to parallel data requests, and maintain separate compliance documentation for each jurisdiction.

DDS addresses this inefficiency by enabling **supervisory convergence at the measurement stage.** When NCAs rely on common indicators derived from existing regulatory reporting, the need for bespoke, jurisdiction-specific data requests is substantially reduced. Firms benefit from more **predictable and standardised supervisory expectations**, while NCAs benefit from more **comparable and mutually intelligible supervisory evidence**, facilitating cooperation in colleges and joint inspections.

DDS-enabled convergence reduces this by:

- **Standardising what is measured and how** across jurisdictions.
- **Reducing bespoke data requests** by re-using harmonised reporting datasets.
- **Providing a common evidential language** for colleges, enhanced supervisory cooperation and joint actions.

DDS foster a dual burden reduction, internal (for NCAs) and external (for supervised firms). Authorities benefit from comparable and mutually intelligible evidence, while firms benefit from more predictable supervisory expectations—directly supporting the EU's simplification and burden-reduction agenda. Indeed, this dual burden reduction is directly aligned with the European Commission's simplification agenda and with the competitiveness framing articulated in the Draghi and Letta reports. Fragmentation in supervisory measurement is itself an **economic friction** that raises costs, slows cross-border integration, and reduces contestability; conversely, convergence through shared analytical infrastructures lowers these costs without compromising supervisory objectives[55].

In this sense, **DDS functions as an institutional coordination device**: it provides a common supervisory language grounded in data, enabling NCAs to align simplification, efficiency, and convergence within a single operational framework that supports both burden reduction and supervisory effectiveness.

6. Multi-Mission Supervision and Forward-Looking Analytics

6.1 Shared Infrastructure for Multiple Supervisory Missions

A key structural advantage of DDS is its ability to support multiple supervisory missions—prudential, conduct, market integrity, operational resilience, and systemic risk—using shared data, pipelines, and governance. Historically, these missions were often supported by siloed data systems and teams; DDS enables re-use and integration[56].

Moreover, DDS is inherently **dynamic and forward-looking.** Unlike static, backward-looking compliance checks, DDS employs rolling windows, trend analysis, change-point detection, and

---

[55] See European Commission's Finess Check (2019), Draghi (2024) and Letta (2024).

[56] This multi mission capability fundamentally alters the economics of supervision by enabling reuse and cross domain integration, reducing duplicative data requests, and facilitating holistic risk assessment at the entity and system level.

early-warning indicators to identify deterioration **before risks crystallise into losses or enforcement cases**[57].

Traditional supervisory structures often compartmentalise missions by legal mandate or organisational unit: conduct teams focus on consumer protection and market abuse, prudential teams on solvency and liquidity, operational risk teams on ICT and third-party resilience, and macroprudential authorities on systemic interconnections. While specialisation is necessary, fragmentation creates inefficiencies: entities report similar data to multiple units, supervisors lack integrated views of risk profiles, and cross-cutting issues (e.g. operational failures affecting both conduct and prudential resilience) are detected late or not at all[58].

DDS addresses this fragmentation by establishing a **common data layer and analytical infrastructure** that supports multiple supervisory lenses applied to the same underlying datasets[59].

- **Conduct supervision** can use transaction-level MiFIR data to detect mis-selling patterns, pricing anomalies, or unfair treatment of retail clients, while simultaneously using EMIR data to monitor suitability and appropriateness in derivatives sales.
- **Prudential supervision** can use the same EMIR exposure data to assess counterparty credit risk, leverage, and margin adequacy, feeding into capital adequacy assessments and stress testing.
- **Operational resilience supervision** can analyse reporting timeliness, rejection rates, and system outage patterns as proxies for ICT robustness and third-party risk, consistent with ECB and Bank of England operational resilience frameworks.
- **Market integrity supervision** can screen for market manipulation, insider trading, and transparency breaches using MiFIR transaction reports, cross-referenced with ownership data and suspicious transaction reports.
- **Systemic-risk surveillance** can reconstruct derivatives networks, assess concentration and interconnectedness, and monitor liquidity and margin dynamics to identify amplification channels and contagion risks.

The **marginal cost of extending DDS to a new mission** is substantially lower than building separate infrastructures for each mandate: the primary incremental effort lies in designing mission-specific indicators and decision rules, not in replicating data ingestion, validation, and storage layers. This cost structure strengthens the business case for DDS investment and encourages a **holistic supervisory view** that integrates micro-prudential, conduct, and macroprudential considerations[60]. This re-use is a central

---

[57] See Yildrim and Sanyal (2022) and BCBS (2025) for a review on early warning indicators and supervisory effectiveness.

[58] See the new supervisory approach developed by Central Bank of Ireland (CBoI 2025a, 2025b, 2025c).

[59] Examples of multi mission indicators:

- *Prudential + conduct: Data quality metrics, reporting deficiency counts, remediation backlogs.*
- *Market integrity + prudential: Valuation outliers and trading anomalies derived from MiFIR reports.*
- *Operational resilience + prudential: Incident frequencies, outages, and third-party concentration metrics.*
- *Systemic risk + prudential + market integrity: Network centrality measures, exposure concentrations, common asset holdings.*

[60] The ECB's SREP methodology for operational and ICT risk explicitly adopts this integrated perspective, combining operational risk assessments, ICT risk evaluations, and operational resilience conclusions within a unified scoring framework that draws on multiple data sources and supervisory modules. Similarly, Central Bank of Ireland's "Transforming Regulation and Supervision" programme emphasises cross sectoral risk integration (conduct, AML/CFT, financial resilience) and the use of horizontal analytics to support risk-based prioritisation across missions.

mechanism of burden reduction: NCAs invest once in infrastructure and skills and deploy these across multiple missions with low marginal cost.

6.2 Forward-Looking, Dynamic Monitoring

Traditional supervisory reporting is often **static and backward-looking**: it captures the state of an entity or portfolio at a point in time (typically quarter-end or year-end) and is analysed retrospectively, often with significant lags between data submission and supervisory review. This approach struggles to detect emerging risks or rapid deterioration, limiting supervisors' ability to intervene proactively.

DDS transforms this paradigm because DDS is inherently forward-looking and dynamic rather than static and retrospective. DDS introduces **dynamic, forward-looking analytics** that continuously monitor trends, detect changes in risk trajectories, and generate **early-warning signals** before issues escalate into crises or enforcement actions. Key techniques include:

1. **Rolling-window indicators**: Metrics computed over moving time windows (e.g. 90-day, 180-day) that smooth noise while highlighting persistent trends, enabling detection of gradual deterioration or sudden shocks.
2. **Trend analysis and change-point detection**: Statistical methods that identify inflection points where risk profiles shift significantly (e.g. sudden increases in leverage, liquidity drawdowns, data-quality errors), triggering supervisory alerts.
3. **Persistence tracking**: Monitoring the duration and recurrence of issues (e.g. repeated breaches of thresholds, ongoing data-quality failures) to distinguish transient anomalies from structural problems requiring remediation.
4. **Peer divergence indicators**: Comparing entities to their peers over time to identify those whose risk profiles are deteriorating relative to benchmarks, even if absolute values remain within regulatory limits.

Furthermore, machine-learning and econometric models trained on historical data can support early-warning models aiming at predicting credit defaults, liquidity stress, operational failures, or conduct breaches, providing lead times of several months for supervisory intervention[61].

This allows NCAs to move towards proactive, preventive supervision, engaging with entities earlier in the risk cycle and potentially reducing the need for more intrusive interventions later. For NCAs, forward-looking supervision reduces burden by enabling **earlier, lower-cost interventions**[62]**.**

7. Governance, Model Risk, and Implementation

7.1 Limitations of data-driven supervision

While data-driven supervision offers substantial burden-reduction potential and supports supervisory convergence, it is **not a panacea**. DDS introduces new risks, depends on critical preconditions (data availability and quality), and must be implemented carefully to avoid failure modes that increase rather than reduce supervisory burden.

[61] See BCBS-BIS (2025) and Galytix (2022)

[62] Supervisory letters and desk-based engagement can correct issues before they require intensive on-site inspections, formal remediation programmes, or enforcement, thereby lowering both internal supervisory costs and external compliance burdens on firms.

The efficacy of a DDS framework is constrained by several interrelated limitations. **Data quality**—encompassing availability, coverage, completeness, and accuracy—constitutes the primary boundary: measurement error, missing observations, and systematic biases in regulatory reporting cannot be entirely eliminated and will attenuate the reliability of derived indicators[63]. Equally, DDS is **complementary** rather than substitutive to traditional supervisory practice; qualitative judgements, on-site inspections, and stakeholder engagement remain indispensable for interpreting signals, resolving ambiguities, and capturing contextual information that automated processes cannot infer. Implementation imposes further constraints in the form of **organizational complexity and cost**: establishing robust infrastructure, governance arrangements, and staff capabilities requires sustained investment. Finally, **model and measurement risk** is inherent to any automated system—false positives and negatives will occur, and model performance may degrade over time without systematic recalibration.

Mitigating these limitations requires a deliberate, governance-centred approach. Data governance and multi-layer validation architectures should be instituted to detect, correct, and document data deficiencies, supported by systematic feedback mechanisms with reporting entities. Model risk must be managed explicitly through rigorous validation protocols, continuous monitoring, and comprehensive documentation of assumptions, limitations, and performance metrics. A phased implementation strategy—beginning with pilots, accompanied by clear communication and targeted training—reduces operational risk and facilitates organisational learning. Throughout deployment, transparency about uncertainty and the known limits of DDS outputs is essential to ensure that automated indicators inform, rather than supplant, supervisory judgement. Appendix E further illustrate the limitation of DDS and the mitigation strategies[64].

This is why any attempt to reduce supervisory burden through DDS must be designed to enhance, rather than dilute, the macroprudential view that depends on granular, entity-level and network-level data.

7.2 Governance Architecture

The successful implementation and sustainable operation of data-driven supervision depend critically on **organisational adaptation, robust governance frameworks, and the preservation of supervisory judgement**. While DDS offers substantial burden-reduction potential through automation and scalability, it also introduces new risks—particularly **model risk, data risk, and governance risk**—that, if poorly managed, can erode trust, generate false positives, and ultimately **increase rather than reduce supervisory burden[65]**.

Effective DDS requires **new competencies, organisational structures, and supervisory cultures** that integrate data science, analytics, legal expertise, and domain knowledge. Traditional supervisory organisations often structured around legal compliance, on-site inspection teams, and sector-specific units, must evolve to accommodate cross-functional collaboration and data-centric workflows[66]. Data

---

[63] Empirical evidence from ESMA's data quality reports on EMIR and MiFIR shows that even mature, legally mandated reporting frameworks exhibit persistent deficiencies: missing counterparty identifiers (LEI), inconsistent valuations, incomplete transaction timestamps, and reconciliation breaks between counterparty reports. These deficiencies directly impair DDS capabilities, particularly for systemic risk monitoring (which depends on accurate network reconstruction) and market integrity surveillance (which depends on complete transaction histories). See ESMA (2024a, 2025b, 2025c)

[64] See ECB – Banking Supervision (2024, 2025).

[65] See FSB (2020), ECB – Banking Supervision (2024).

[66] DDS implementation demands diverse skill sets:

- *Data scientists and quantitative analysts: to design, validate, and maintain supervisory indicators, risk models, and early warning systems.*
- *Software engineers and data engineers: to build and operate data pipelines, dashboards, and integration layers.*

Driven culture can be further developed by ESAs and ECB, with peer learning, working groups where experiences are shared and all NCAs benefit from this mutual exchange.

DDS fundamentally shifts supervisory culture from **episodic, inspection-driven oversight** to **continuous, data-driven monitoring and engagement**. This transition requires:

- **Comfort with quantitative methods**: supervisors must understand indicator logic, recognise model limitations, and challenge outputs constructively.
- **Structured decision-making**: moving from ad hoc judgement to rule-based triage and escalation pathways, with explicit documentation of reasoning.
- **Collaborative workflows**: supervisory staff work alongside data scientists, sharing domain knowledge and refining analytical approaches iteratively.

DDS must be supported by a clear governance architecture that defines roles, responsibilities, and processes across data, models, and supervisory use.

7.3 Model-Risk Management

While DDS offers efficiency gains, it also introduces **model risk**—the risk that indicators, scoring models, or early-warning systems are mis-specified, biased, unstable, or gamed, leading to incorrect supervisory decisions. Poor governance of DDS can generate **false positives** (unnecessary escalations), **false negatives** (missed risks), **alert fatigue** (overwhelming supervisors with noisy signals), and **loss of credibility** (supervised entities challenging opaque or inconsistent models).

Robust governance frameworks are therefore essential to ensure DDS outputs are **reliable, explainable, and defensible**.

Supervisory indicators and scoring models must be subject to formal governance processes[67]:

1. **Documentation**: Each indicator must have a formal specification document covering: legal anchor, data sources, computation logic, thresholds, interpretation guidance, known limitations, and change history.
2. **Version control**: All changes to indicator definitions, weights, or thresholds must be versioned, approved by a governance committee, and communicated to stakeholders (internal supervisory teams and, where appropriate, supervised entities).
3. **Validation**: Indicators must be validated before deployment and periodically re-validated, assessing predictive accuracy, stability, sensitivity to parameter choices, and potential for gaming.
4. **Sensitivity analysis**: Understanding how indicator outputs change with variations in data quality, thresholds, and model assumptions is critical for interpreting signals and avoiding over-reliance on brittle metrics.
5. **Performance monitoring and recalibration**: Ongoing monitoring of indicator performance (e.g. hit rates, false-positive rates, time-to-event distributions) enables detection of **model**

- *Data governance specialists: to manage data quality, metadata, lineage, and compliance with data protection regulations.*
- *Legal and supervisory experts: to ensure indicators are legally anchored, proportionate, and consistent with enforcement frameworks.*
- *Change management and training: to upskill existing supervisory staff and embed data driven approaches into daily workflows.*

Recruiting and retaining these competencies is challenging for NCAs, which often compete for talent with private sector firms offering higher compensation.

[67] *See ECB – Banking Supervision (2024), ECB – Banking Supervision (2025).*

drift (degradation of predictive power over time due to changing market conditions or entity behaviour) and triggers recalibration or retirement of outdated indicators.

7.4 Explainability, transparency, and legal defensibility

For DDS outputs to be legally defensible and trusted by supervised entities, they must be **explainable and transparent**. This requirement is particularly acute when indicators drive enforcement actions or capital add-ons, as entities have rights to understand the basis of supervisory decisions and to challenge them through administrative and judicial review.

Explainability principles for supervisory indicators:

1. **Clear causal logic**: indicators should link directly to observable risk drivers and regulatory requirements, not be "black box" correlations.
2. **Decomposable scores**: composite scores should be decomposable into constituent components, enabling supervisors to explain **why** an entity received a particular score and which factors dominate.
3. **Record-level drill-down**: supervisors must be able to trace aggregate indicators back to specific transactions, reports, or data fields, providing concrete evidence for supervisory dialogue.
4. **Plain-language interpretation**: technical indicator outputs should be accompanied by supervisory narratives that translate quantitative signals into actionable insights and proportionate next steps.

These principles align with broader regulatory and academic discussions on algorithmic accountability, which emphasise that automated decision systems — especially in high stakes public sector contexts — must be transparent, contestable, and subject to human oversight[68].

7.5 Preservation of Supervisory Judgement

A fundamental misconception about DDS is that it mechanises supervision, replacing human judgement with algorithmic decision-making. In reality, **DDS repositions judgement from initial detection to interpretation, proportionality assessment, remediation design, and enforcement discretion**.

Under traditional supervision, supervisors exercise judgement at every stage: selecting which entities to inspect, interpreting compliance with rules, deciding on follow-up actions, and designing remediation. This is resource-intensive and subject to cognitive biases (availability heuristics, anchoring, confirmation bias) that can produce inconsistent outcomes.

DDS **automates detection and evidence compilation**, freeing supervisors to concentrate on:

- **Contextual interpretation**: understanding **why** indicators are elevated (systemic factors, business model changes, data issues, intentional non-compliance);
- **Proportionality**: calibrating supervisory intensity to entity-specific circumstances, balancing risk signals with governance quality, track record, and cooperation;
- **Remediation design**: working with entities to develop tailored, effective remediation plans that address root causes rather than symptoms;

[68] *See FSB (2020).*

- **Enforcement discretion**: deciding when escalation to formal enforcement is warranted, considering public interest, deterrence, and precedent.

This division of labour between automation and judgement is consistent with academic insights on **risk, uncertainty, and decision-making under ambiguity**, which emphasise that formal models perform best when **combined with expert judgement**, not used as substitutes. Frank Knight's classic distinction between measurable risk (amenable to quantification) and unmeasurable uncertainty (requiring judgement) remains relevant: DDS handles the former, while supervisors manage the latter[69].

DDS is most effective when **combined with traditional supervisory methods**, not used as a substitute. Quantitative indicators excel at **detection, screening, and prioritisation** but are limited in capturing **qualitative factors** (governance culture, strategic intent, management competence) that critically influence risk. The optimal supervisory model is therefore **hybrid**: DDS provides continuous, scalable, objective oversight, while traditional tools provide depth, nuance, and human engagement[70].

8. Implications of Data Driven Supervision

8.1 Policy implications for National Competent Authorities

DDS provides NCAs with a structural response to the supervisory burden arising from expanding mandates, growing populations of supervised entities, and constrained resources. By automating detection, triage, and early evidence assembly, DDS enables continuous oversight of large, heterogeneous populations while concentrating scarce human effort on high-priority cases—an outcome especially salient in jurisdictions with many smaller firms that would otherwise receive little scrutiny[71].

DDS operationalises longstanding aspirations for risk-based supervision. It does so through quantitative risk indicators, composite scoring, and transparent decision rules that link measured severity, persistence, materiality, and uncertainty to supervisory intensity. This systematic approach is in line with the ESMA (2026) Principles on risk-based supervision and reproduces and extends frameworks such as the Central Bank of Ireland's approach to supervision by supplying the analytical infrastructure needed for consistent, repeatable risk categorisation and engagement protocols.

Economically, DDS exhibits high fixed costs but low marginal costs per additional entity or monitoring period; once platforms, indicators, and governance are in place, incremental supervision becomes substantially cheaper than inspection-centric models. That cost profile makes DDS a credible pathway to sustainable supervision amid demographic and fiscal pressures on European public administrations[72].

Finally, DDS enhances supervisory consistency, transparency, and legal defensibility by producing standardised indicators, reproducible evidence, and documented decision logic. These features not only support proportionate resource allocation in line with optimal allocation principles but also strengthen the perceived fairness and legitimacy of supervisory action.

---

[69] *The ECB's SREP framework operationalises this balance through constrained judgement: quantitative anchor scores provide a homogeneous starting point, but supervisors adjust within pre-defined ranges based on qualitative factors, with all adjustments documented and subject to horizontal review. This approach ensures consistency without rigidity and accountability without opacity. See also De Nederlandsche Bank (2015).*

[70] *See ECB – Banking Supervision (2024).*

[71] *An effective and implementable data driven supervision allows NCAs to operationalize ESMA Strategy 2023-2028, with a reduction of costs and an increased convergence.*

[72] *DDS allows NCAs to adress some of the limitations and burdens identified by the industry with reference to supervisory financial reporting. The more efficient use of data, the ability to NCAs to supervise all entities with appropriate indicators together with the possibility to efficiently achieve supervisory convergence are key gains market players benefit through the DDS. See UK Finance (2019) and European Commission (2019).*

8.2 Implications for European supervisory strategy and convergence

At the European level, DDS advances the strategic objectives of ESMA, the ECB, the ESRB, as well as those of IOSCO and the FSB. by operationalising supervisory convergence, improving cross-border coordination, and reinforcing macroprudential monitoring. As noted in earlier sections, DDS embeds shared interpretations of regulatory requirements into common indicators and verification functions applied to harmonised datasets; this convergence manifests semantically (aligned data fields, definitions, and reporting logic), methodologically (uniform indicator definitions, thresholds, and persistence criteria), and procedurally (consistent escalation pathways and evidential standards)[73].

By translating soft-law guidance into concrete analytical artefacts—common indicator libraries and interoperable platforms—DDS reduces supervisory fragmentation for cross-border groups, lowering compliance costs and regulatory uncertainty and facilitating coordinated action in supervisory colleges and crisis management groups. The ECB's SSM illustrates the benefits at scale: common SREP methodologies, horizontal analytics, and integrated data governance enable consistent oversight across multiple jurisdictions without requiring full institutional centralisation[74]. For macroprudential authorities, DDS supplies the micro-level data infrastructure and analytics that underpin systemic risk assessment: EMIR derivatives data support network contagion and margin monitoring[75], MiFIR transaction data inform liquidity and procyclicality analysis[76] , and prudential reporting (COREP, FINREP) permits exposure aggregation[77], concentration analysis, and stress-test inputs. By improving the quality, timeliness, and comparability of these sources through systematic validation and feedback, DDS strengthens the empirical foundations of macroprudential policy in line with scholarly emphasis on data granularity and observability[78].

8.3 Implications for supervised entities

DDS reshapes both the structure and predictability of supervisory engagement, with direct consequences for firms' compliance strategies, risk management, and resource allocation. Under DDS, selection for supervisory attention is driven by observable, data-based indicators rather than opaque discretionary processes; this transparency allows firms to anticipate supervisory priorities, to monitor and remediate issues internally, and to prepare more focused, evidence-based responses when engagement occurs. As a result, regulatory uncertainty is reduced and interactions become more predictable (FSB, 2020).

DDS can also lower external compliance costs. Reliance on existing regulatory reports diminishes the need for bespoke data requests, automated evidence compilation shortens supervisory cycles, and harmonised datasets enable reuse across conduct, prudential, and resilience missions—reducing duplicative reporting. Realising these efficiencies, however, depends on deliberate design choices: platforms and indicator libraries must be configured explicitly to minimise firms' reporting burden rather than solely to optimise internal supervisory workflows.

Finally, DDS strengthens incentives for investment in data quality, internal controls, and governance. Persistent data deficiencies become observable supervisory signals that trigger escalation, aligning

---

[73] *See ESMA works on supervisory convergence and on the use of transaction reporting (e.g. ESMA, 2025b, 2025d, 2025f).*

[74] *See ECB- Banking Supervision (2024, 2025).*

[75] *See Banca d'Italia (2025).*

[76] *See ESRB (2025).*

[77] *See ESRB (2025) and ECB – Banking Supervision (2024).*

[78] *See ECB (2015).*

firms' private incentives with supervisory objectives and potentially reducing adversarial dynamics between supervisors and supervised entities (ESMA, 2025b).

9. Conclusion

DDS responds directly to the structural capacity challenge facing EU supervision: mandates and complexity have grown while resources have not scaled proportionately. Rather than a mere technological upgrade, DDS constitutes a reconfiguration of the supervisory production function — grounded in measurement theory, decision theory, and systemic-risk analysis — that converts regulatory reporting into reproducible indicators, risk scores, and automated decision pathways. This enables continuous, population-scale coverage with targeted depth, while preserving the role of professional judgement by structuring, auditing, and scaling supervisory decisions.

DDS directly advances the European policy agenda set out in the Draghi and Letta reports, the European Commission's simplification initiatives, and ESMA's strategic priorities (ESMA, 2022). It supports **competitiveness** by reducing regulatory fragmentation and compliance costs, thereby facilitating scale formation in financial services; it advances **simplification** by implementing a "less is more" approach that extracts greater supervisory value from existing data rather than imposing new reporting burdens; it promotes **Single Market integration** through supervisory convergence and reduced cross-border fragmentation; and it underpins **sustainable regulation** by lowering marginal supervisory costs per entity and enabling NCAs to sustain effectiveness under fiscal and demographic constraints[79].

Key operational benefits include:

- Population coverage with targeted intensity;
- Consistent, transparent risk-based allocation of resources;
- Lower marginal costs and scale economies for ongoing supervision;
- Convergence through shared indicators and tools;
- Multi-mission reuse of harmonised data and analytics;
- Earlier detection of emerging risks, reducing remediation and enforcement costs.

Realising these benefits requires disciplined implementation: sustained investment in data infrastructure, analytics, skills, and governance; phased rollouts with demonstrable early wins and continuous learning; robust model-risk management to preserve trust; and coordinated convergence driven by ESMA and peer engagement. With sound governance, peer coordination, and persistent attention to data quality, DDS can become a cornerstone of a more efficient, convergent, and resilient European supervisory architecture that safeguards financial stability while supporting competitiveness and market integration.

[79] *These benefits address some of the limitations identified in the European Commission's Fitness Check (2019).*

Appendix A – Mathematical Notation and Formal Frameworks

A.1 Inference Under Measurement Error

Supervisory data are noisy signals of true risk and compliance states:

$$Y_t = f(X_t;\, \theta) + \epsilon_t\ . \tag{6}$$

Where $f(\cdot)$ represents the transformation induced by the legal reporting framework and technical implementation (reporting templates, taxonomies, validation rules), θ captures institutional parameters of reporting processes (systems, outsourcing arrangements, internal controls), and ε_t denotes measurement error (missing values, misclassification, stale timestamps, inconsistent identifiers).

Supervisor chooses actions to minimise expected supervisory loss from action a given the true state X_t.

$$a^*(t) = arg \min_{a \in A} E[L(a, X_t)|Y_t]\ . \tag{7}$$

Risk-based DQ supervision seeks to allocate effort where the expected marginal reduction in supervisory loss is highest. This provides a principled foundation for prioritisation, thresholds, and escalation.

A.2 SPMU Composite Score

Composite score integrating severity, persistence, materiality, and uncertainty:

DDS operationalises risk-based prioritisation through **composite score** and **decision trees**. As formalised in Section 5.4, a general composite score combines severity, persistence, materiality, and uncertainty:

$$DDSScore(e, \mathrm{t}) = \sum_i s_i \cdot g\big(p_i(e,t)\big) \cdot h\big(m_i(e,t)\big) \cdot u_i(e,t) \cdot \pi_i(e,t)\ . \tag{8}$$

Where:

- s_i: **severity**,
- $p_i(e,T)$: **persistence**,
- $m_i(e,T)$: **materiality**,
- $u_i(e,T)$: **uncertainty adjustment**,
- $\pi_i(e,T)$: **policy weight**, adjusting for current supervisory priorities
- $g(\cdot)$ and $h(\cdot)$: non-linear transformations (for example, penalties, weights).

A.3 Triage Decision Rules (Illustrative Formalisation)

Let $A \in \{0,1,2,3,4\}$ denote supervisory action intensity.

- $A = 0$: **Baseline monitoring**.
- $A = 1$: **Enhanced off-site monitoring.**
- $A = 2$: **Structured engagement.**
- $A = 3$: **Intensive supervision**.
- $A = 4$: **Enforcement interface**.

Example:

If $DataQualityScore\ (e,t)\ >\ \tau_1 \Rightarrow\ A \geq 1$.

If $\sum_{s=t-H}^{t} 1\{Issue_{,}(e,s) > threshold\} \geq H \Longrightarrow A \geq 2$.

If $m_i(e,t) > M \Rightarrow A \geq 2$ or higher.

If $t_{remediation} - t_{agreement} > T_{max}$ and issues remain $\Rightarrow$ A=4.

A.4 Indicator families and metric templates

DDS operationalises supervisory priorities through **indicator families** that capture different dimensions of risk and compliance. While specific indicators vary by regulatory mandate, the following categories are broadly applicable:

A.4.1 Compliance-Integrity Metrics

Let $R_e(T)$ denote the set of reports submitted by entity e during period T, and let $V_s(r)$ denote the set of validation rules applied to report r. Define severity tiers s∈{Critical,High,Medium,Low} and compute tier-specific error rates:

$$ErrRate_s(e,T) \ = \frac{\sum_{r\in R_e(T)} \sum_{v\in V_s(r)} 1\{validation\ v\ fails\}}{\sum_{r\in R_e(T)} |V_s\ (r)|}\ , \tag{9}$$

Supervisory interpretation: Persistent elevated critical or high error rates indicate control weaknesses or inadequate reporting infrastructure; burst patterns (sudden spikes) suggest change-management failures or system disruptions.

A.4.2 Timeliness and operational stability

Define reporting latency as $Latency(r)=t_{submit}(r)-t_{event}(r)$

and compute the late-reporting rate:

$$LateRate\ (e,T;\lambda) = \frac{|\{r: Latency\ (r)\ >\ \lambda\}|}{|R_e(T)|}\ , \tag{10}$$

where λ is a regulatory or supervisory threshold (e.g. T+1 for transaction reporting).

Supervisory interpretation: Latency spikes are operational resilience signals, indicating system stress, capacity constraints, or third-party failures; sustained lateness indicates structural reporting fragility requiring remediation.

A.4.3 Coherence and plausibility (economic logic)

Define economic plausibility constraints c_j (e.g. price must be within market bounds, notional must align with typical trade sizes, margin must be consistent with risk) and compute a plausibility score:

$$Plaus\ (i) = \sum_j w_j \cdot 1\{c_j(i)\ violated\}, \tag{11}$$

where w_j reflects the severity weight of constraint $_j$.

Supervisory interpretation: High plausibility scores flag potential data errors, misreporting, or economic anomalies (e.g. off-market pricing potentially linked to manipulation or valuation model failures). Constraints should be segmented by business model and product class to avoid false positives.

A.4.4 Materiality-Weighted Non-Compliance

Let $w(r)$ denote a materiality weight for report r (e.g. transaction notional, position valuation, exposure size). Compute:

$$MWNCR(e,T) = \frac{\sum_{r \in R_e(T)} 1\{non - compliant\} \cdot w(r)}{\sum_{r \in R_e(T)} w(r)}\,. \qquad (12)$$

Supervisory interpretation: This ensures the NCA focuses on defects that matter, prioritising large, systemically significant, or consumer-facing activities over minor technical issues.

A.4.4 Remediation-Backlog Dynamics

Track the **defect backlog** B_t for entity e :

$$B_{t+1} = B_t + N_t - S_t \qquad (13)$$

where N_t is the number of new defects identified in period t, and S_t is the number resolved.

Supervisory interpretation: Persistent backlog growth predicts higher future supervisory burden (repeated engagement, escalation) and indicates inadequate governance capacity or insufficient commitment to remediation. This metric informs decisions about intensifying supervision or imposing formal remediation plans.

A.5 Exposure-Weighted Scoring: Design Rationale and Formal Specification

A foundational design principle of data-driven supervision is that supervisory metrics must reflect both materiality and systemic relevance — not merely the frequency of non-compliance or the magnitude of risk in isolation. Treating all reporting deficiencies as equivalent regardless of the economic significance of the underlying activity would systematically misdirect supervisory effort toward peripheral entities and minor infractions, while underweighting defects with genuine macroprudential consequences. This concern is directly reflected in ESMA's 2026 Principles on Risk-Based Supervision, which establish that risk scores must be constructed as a function of both probability of materialisation and impact on markets or entities, with supervisory prioritisation and resource allocation driven by the resulting composite score rather than by case counts (ESMA, 2026). The exposure-weighted scoring framework translates this principle into a formal structure that amplifies supervisory priority for entities whose deficiencies carry disproportionate system-wide implications.

Motivating example. Consider two entities that each report errors in counterparty identification — a data-quality failure with direct implications for systemic-risk monitoring. Entity A is a mid-market financial firm with €500 million in derivatives exposures; Entity B is a systemically important institution with €500 billion in derivatives exposures whose identifiers underpin contagion analysis and macroprudential surveillance. A raw error-rate metric would treat both deficiencies as equivalent. An exposure-weighted metric, by contrast, would assign Entity B a materially higher supervisory priority score, reflecting the fact that its data-quality failures directly impair aggregate visibility into interconnectedness, concentration risk, and potential contagion channels. This asymmetry is not a distortion of the compliance framework — it is a deliberate calibration of supervisory effort to systemic stakes, consistent with the principle that vulnerabilities must be assessed at the level of systemically important institutions (ESRB, 2026; Adrian and Brunnermeier, 2011).

Formal specification. Building on the materiality-weighted non-compliance rate introduced in Section A.4.3, define the exposure-weighted non-compliance rate for entity ee over period T as:

$$MWNCR(e,T) = \frac{\sum_{r \in R_e(T)} \left[1_{non-compliant} \cdot w(r) \right]}{\sum_{r \in R_e(T)} w(r)} . \quad (14)$$

where $w(r)$ denotes a materiality weight assigned to report r, capturing one or more dimensions of economic significance: transaction notional value, derivatives exposure, position valuation, or a composite systemic importance indicator. The numerator aggregates the weighted mass of non-compliant reports, while the denominator normalises by total weighted reporting activity. This formulation ensures that the indicator remains bounded on [0,1] and is directly comparable across entities of heterogeneous size, while preserving sensitivity to the economic scale of detected defects.

The choice of weight $w(r)$ is a supervisory design parameter. A notional-based weight prioritises large transaction flows; a systemic-importance indicator — derived, for instance, from network centrality scores or marginal systemic risk contributions of the type formalised in Adrian and Brunnermeier (2011) and Brunnermeier and Cheridito (2019) — redirects attention toward contagion-relevant nodes regardless of transaction size. In practice, a composite weight combining both dimensions may be warranted, particularly where the supervisory objective spans both microprudential adequacy and macroprudential resilience. The ECB's treatment of the materiality threshold under Article 178 CRR illustrates this dual sensitivity, combining an absolute threshold with a relative threshold expressed as a percentage of total credit obligations, precisely to ensure that both large and proportionally significant exposures attract supervisory attention (ECB-SSM, 2015).

Institutional grounding. This materiality-aware design is consistent with established supervisory practice at the European level. The ECB's SREP methodology explicitly constructs overall risk scores by weighing individual risk assessments according to size, interconnectedness, and systemic importance, embedding the principle that supervisory intensity should be proportional to systemic stakes rather than entity count (ECB Banking Supervision, 2024b). The 2026 SREP update further refines this through a Flexible Risk Assessment System that calibrates the depth of assessment to entity-specific risk profiles, reinforcing the proportionality logic underlying exposure-weighted scoring (ECB Banking Supervision, 2025c). Similarly, the ESRB's risk dashboard structures its indicators around interlinkages, composite measures of systemic stress, concentration, and liquidity transformation, with explicit analytical focus on systemically important institutions as the primary loci of contagion risk (ESRB, 2026). The exposure-weighted scoring framework developed here operationalises these institutional principles within a reproducible, data-driven supervisory architecture, ensuring that the NCA's analytical and inspection resources are systematically directed toward defects that matter — large in scale, systemic in reach, or directly consumer-facing — while routine compliance issues in peripheral activities are managed through proportionately less intensive monitoring channels.

Appendix B: Economics of Data-Driven Supervision: Cost Structure, Scale Economies, and Long-Run Efficiency

The burden-reduction potential of DDS is fundamentally rooted in its economic structure. Unlike traditional supervisory approaches that exhibit near-linear cost scaling with supervised populations, DDS exhibits **high fixed costs but low marginal costs**, generating substantial scale economies and declining average costs per entity over time. This section formalises the cost structure of DDS, compares it to traditional supervision, and demonstrates why the economic logic of DDS aligns with the European

Commission's efficient regulation agenda and the broader simplification objectives articulated in the Draghi and Letta reports[80].

B.1 Cost model: formalising the economics of supervision

Let total supervisory cost over planning horizon T be modelled as:

$$C = F + \sum_{t=1}^{T} (c_{run} + c_{case} \cdot n_t) \tag{15}$$

where:

- F: **Fixed setup costs**, including data infrastructure (servers, databases, integration platforms), analytics development (indicator design, validation, dashboards), governance frameworks (model risk management, documentation), and staff training.
- c_{run}: **Baseline operating costs** per period, covering infrastructure maintenance, software licences, ongoing data-quality management, and core analytics team salaries.
- c_{case}: **Marginal cost per supervisory case**, representing the incremental resource cost of handling an escalated entity (desk-based review, engagement, remediation oversight).
- n_t: **Number of escalated cases** in period t, determined by the triage model and entity risk profiles.

The **average cost per supervised entity** is:

$$\bar{C}(N, T) = \frac{F}{N \cdot T} + \frac{c_{run}}{N} + \frac{c_{case}}{N} \cdot \frac{\sum_{t=1}^{T} n_t}{N \cdot T} \tag{16}$$

where N is the total supervised population. As N and T increase, the first term (amortised fixed cost) declines toward zero, yielding substantial **scale economies**.

B.1.1 Traditional supervision: linear cost scaling

Traditional supervision, reliant on periodic inspections, manual data requests, and qualitative reviews, has a cost structure approximated by:

$$C_{trad} = F_{trad} + \sum_{t=1}^{T} (c_{inspect} \cdot m_t) \tag{17}$$

where F_{trad} is low (basic IT, office infrastructure), but $c_{inspect}$ (cost per inspection) is high and m_t (number of inspections) grows roughly proportionally with N to maintain coverage. The average cost per entity is:

$$\bar{C}_{trad}(N, T) \approx \frac{F_{trad}}{N \cdot T} + c_{inspect} \cdot \frac{\alpha N}{N} = \frac{F_{trad}}{N \cdot T} + \alpha c_{inspect} \tag{18}$$

where a is the inspection rate (e.g. 10% of entities inspected per year). Because $c_{inspect}$ is large and a is typically fixed, traditional supervision exhibits **near-linear cost scaling** with N.

B.1.2 DDS: sublinear scaling and declining average costs

DDS fundamentally changes this cost structure:

---

[80] *See FSB (2025) for a recent illustration of costs of supervision.*

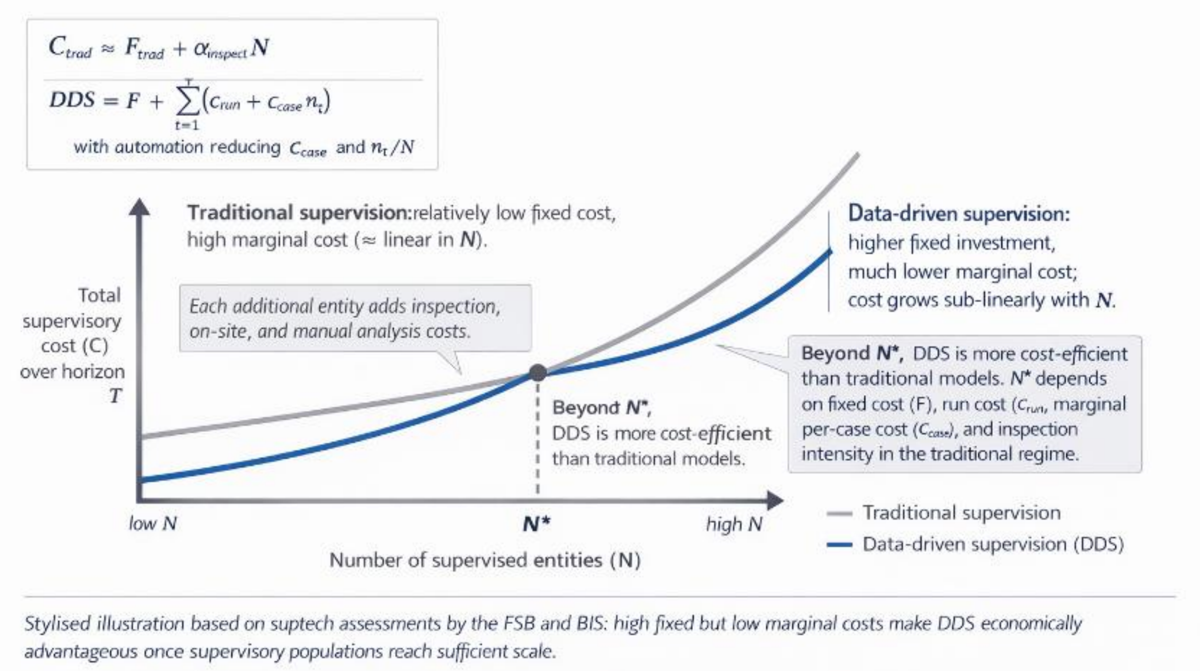
1. **Higher *F***, reflecting substantial upfront investment in suptech platforms, analytics, and governance.
2. **Lower c_{case}**, because evidence compilation, benchmarking, and routine follow-up are automated, reducing per-case staff time.
3. **Lower n_t** relative to N, because improved triage and early intervention reduce the fraction of entities requiring intensive engagement.

As a result:

$$\lim_{N,T\to\infty} \bar{C}_{DDS}(N,T) \ll \bar{C}_{trad}(N,T) \tag{19}$$

and the **breakeven population** N^* (where total DDS cost equals traditional supervision cost) depends on F, c_{case}, and $a\ c_{inspect}$.

Figure 1 – Cost structure of traditional supervision and DDS



Stylised illustration based on suptech assessments by the FSB and BIS: high fixed but low marginal costs make DDS economically advantageous once supervisory populations reach sufficient scale.

B.2 Mechanisms of cost reduction in DDS

DDS reduces supervisory costs through four interlocking mechanisms:

Mechanism 1: Automation of routine tasks

Data ingestion, validation, reconciliation, and metric computation are fully automated. For example, ESMA's data-quality reports document how automated validation pipelines process millions of EMIR and MiFIR reports continuously, flagging errors and generating feedback to entities without manual intervention (ESMA, 2025b, 2025c).

Mechanism 2: Reuse and scalability

Once indicator pipelines are built and validated, they can be applied to additional entities, time periods, and datasets at near-zero marginal cost. Multi-mission reuse (applying the same infrastructure to conduct, prudential, and operational resilience supervision) further amortises fixed costs.

Mechanism 3: Improved triage and prioritization

Risk-based scoring and decision trees concentrate supervisory resources on high-priority cases, reducing the total number of intensive engagements n_t relative to the population N. Early intervention enabled by continuous monitoring shortens case durations and reduces escalation to enforcement, lowering per-case costs.

Mechanism 4: Evidence automation

Standardised evidence packs are generated automatically, reducing the staff time required for case preparation, supervisory dialogue, and enforcement documentation from weeks to hours. This acceleration enables supervisors to handle larger caseloads without quality degradation.

B.3 Alignment with efficient regulation and simplification agendas

The economic logic of DDS aligns directly with the European Commission's **efficient regulation** and **simplification** agendas, which frame burden reduction not as deregulation but as achieving the same (or better) supervisory outcomes at lower administrative and compliance cost.

DDS achieves this by:

1. **Maximising reuse of existing regulatory reporting**: rather than imposing new reporting obligations, DDS extracts more supervisory value from data firms already submit, reducing external compliance burden[81].
2. **Reducing ad hoc supervisory requests**: systematic use of common indicators and evidence packs limits the need for bespoke data requests, lowering firm-side response costs and supervisory processing costs.
3. **Accelerating time-to-resolution**: early detection and automated evidence compilation shorten supervisory cycles, reducing the duration and intensity of supervisory interactions for both parties.

This dual burden reduction—internal (for NCAs) and external (for supervised entities)—is central to the **"Less is More"** logic articulated in banking-sector discussions: complexity and volume do not automatically translate into better outcomes, and rationalisation through data-driven approaches can improve effectiveness while lowering costs[82].

The Draghi and Letta reports further emphasise that **regulatory fragmentation and administrative complexity** are structural impediments to EU competitiveness, investment, and scale formation. To the extent that supervisory fragmentation (heterogeneous metrics, divergent evidence standards) contributes to this complexity, DDS's convergence mechanism—shared indicators and analytical infrastructures—directly supports the competitiveness and Single Market integration agendas by reducing multiplicative compliance burdens on cross-border entities.

---

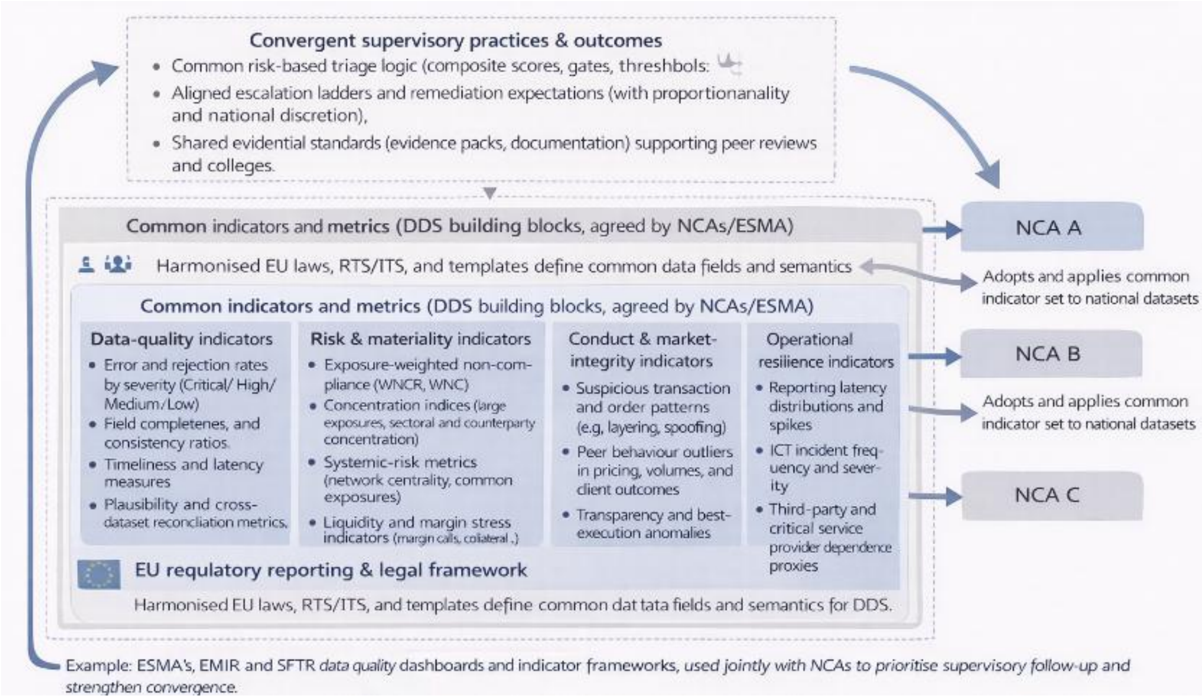
[81] See CSSF-CAA (2024) as well as CSSF Communiqué on Monitoring the quality of transaction reports received under Article 9 of EMIR (2022), CSSF Communiqué on Monitoring the quality of transaction reports received under Article 26 of MiFIR (2026a) and CSSF Communiqué on the Outcomes of the 2025 SFTR Data Quality indicators review (2026b).

[82] See European Commission's Fitness Check (2019).

Appendix C: Supervisory convergence in practice: governance, coordination, and peer consistency

Supervisory convergence is an **analytical property** of DDS (common indicators applied to harmonised datasets yield convergent signals). Convergence must be **operationalised through institutional arrangements**—governance frameworks, coordination mechanisms, and peer-review processes that translate technical capabilities into consistent supervisory practice.

Figure 2 – Common indicators as building blocks of DDS and supervisory convergence



C.1 Three layers of convergence

Effective convergence through DDS requires alignment across three interrelated layers:

Layer 1: Semantic convergence: NCAs must share common interpretations of regulatory requirements, data fields, and reporting logic. For example, what constitutes a "late report" under EMIR? How should "transaction timestamp" be interpreted in cross-border, multi-venue trades? Without semantic agreement, identical data can yield divergent indicators.

Layer 2: Methodological convergence: NCAs must agree on indicator definitions, computation methods, thresholds, and persistence criteria. For example, should data-quality error rates be computed as simple averages or exposure-weighted? Should escalation thresholds be absolute (e.g. rejection rate >5%) or relative (e.g. top decile of peer distribution)?

Layer 3: Procedural convergence: NCAs must align escalation pathways, remediation expectations, and evidential standards. For example, at what indicator level does desk-based engagement escalate to on-site inspection? What documentation is required in evidence packs?

DDS supports all three layers by **embedding supervisory interpretation into measurable, reproducible functions** applied to common datasets, making differences in approach **visible and actionable** rather than hidden in qualitative judgements.

C.2 Peer coordination mechanisms and DDS

ESMA's supervisory convergence toolkit—peer reviews, supervisory colleges, Common Supervisory Actions (CSAs), Q&As, and working groups—can leverage DDS outputs to enhance coordination:

- **Peer reviews of DDS implementations**: ESMA can assess how NCAs design, validate, and govern indicators, identifying best practices and divergences, and promoting methodological alignment.
- **Common indicator libraries**: ESMA working groups can develop and maintain reference indicator definitions (e.g. for EMIR data quality, MiFIR market integrity), which NCAs adopt or adapt, reducing fragmentation.
- **Benchmarking and outlier analysis**: NCAs can compare indicator distributions, escalation rates, and remediation outcomes across jurisdictions, identifying cases where divergence reflects legitimate differences (e.g. market structure) versus inconsistent interpretation.
- **Evidence-based peer dialogue**: supervisory colleges and coordination fora shift from abstract legal discussions to **concrete evidence** (e.g. "Entity X triggers high data-quality alerts in Country A but not Country B—why?"), enabling faster resolution of interpretive differences.

Galán and Svetiev (2025) analyses ESMA's peer-review evolution and documents a shift from compliance checklists toward **qualitative assessments of supervisory practices and root-cause analysis**, with increasing emphasis on **common tools and shared methodologies** as convergence mechanisms. DDS provides the technical substrate for this evolution by making supervisory practices observable, comparable, and improvable.

C.3 Convergence without uniformity: preserving proportionality

Critically, DDS-enabled convergence operates at the level of **measurement and prioritisation**, not enforcement outcomes. NCAs applying the same indicators to the same data will generate the same **supervisory signals** (e.g. Entity X has elevated data-quality issues), but may legitimately choose different **supervisory actions** based on:

- **Entity-specific context**: track record, cooperation, governance quality, systemic importance;
- **National supervisory priorities**: thematic focuses, resource constraints, enforcement strategies;
- **Proportionality and discretion**: tailoring intensity to risk profiles and local market conditions.

This **convergence without uniformity** preserves subsidiarity and supervisory flexibility while ensuring that divergence is **deliberate, documented, and justifiable**, not accidental or driven by inconsistent evidence standards. The result is a more efficient Single Market for financial services, where cross-border entities face **predictable, harmonised supervisory expectations** at the measurement stage, reducing compliance fragmentation and uncertainty.

Appendix D – Data Architecture and Validation Layers

D.1 Methodological architecture: the DDS pipeline as a closed-loop control system

From a methodological perspective, DDS can be conceptualised as a **closed-loop supervisory control system** in which regulatory reporting data are continuously observed, transformed into risk signals, used to inform supervisory actions, and fed back into entities' behaviour and future reporting quality[83].

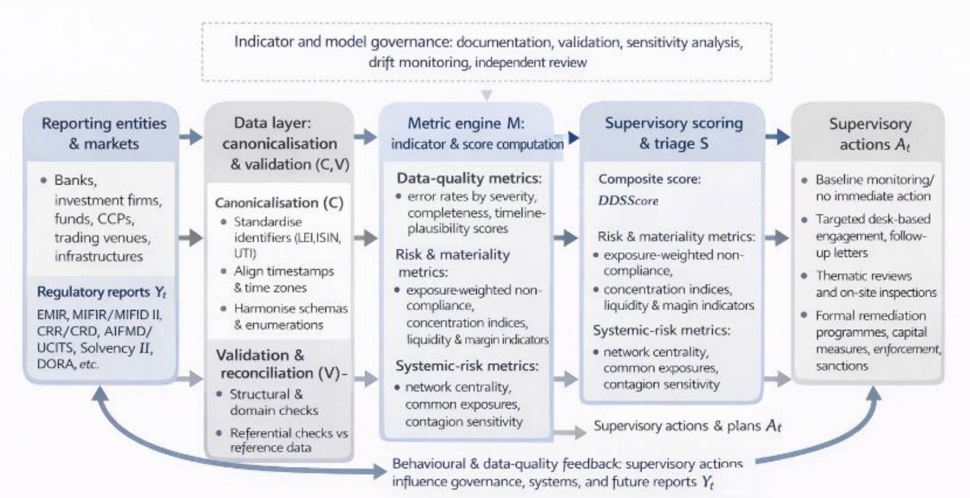
Let Y_t denote the regulatory reporting dataset at time t. The DDS pipeline applies a sequence of transformations:

$$Y_t \xrightarrow{C} \widetilde{Y}_t \xrightarrow{V} Z_t \xrightarrow{M} M_t \xrightarrow{S} A_t$$

where:

- C: **Canonicalisation** (standardisation of identifiers, timestamps, taxonomies, and schemas to enable cross-dataset integration and reproducibility).
- V: **Validation and reconciliation** (application of rule-based and cross-source checks to detect errors, inconsistencies, and anomalies).
- Z_t: **Validated supervisory dataset** (clean, reconciled data suitable for analytical use).
- M: **Metric construction** (transformation of validated data into supervisory indicators capturing compliance, risk, materiality, and uncertainty).
- M_t: **Vector of supervisory metrics** (quantitative signals used for triage and prioritisation).
- S: **Scoring and selection** (application of composite scoring models and decision rules to determine supervisory actions).
- A_t: **Supervisory actions** (ranging from continued monitoring to engagement, remediation, or enforcement).

Figure 3 – Data-driven supervision pipeline as a closed-loop control system



The DDS pipeline mirrors a closed-loop control system: data are transformed into metrics, metrics into actions, and actions alter future data *and behaviour; under explicit model* and data governance.

[83] *See FSB (2020).*

This modular decomposition is essential for **burden reduction and governance**:

- **Automation**: Early stages (*C, V, M*) are largely automated, reducing manual effort and enabling scalability.
- **Reproducibility**: Each transformation is versioned and documented, ensuring that outputs are deterministic and auditable.
- **Human oversight**: Supervisory judgement is concentrated at the final stage (*S*), where contextual knowledge, proportionality, and policy priorities are applied to model outputs.

D.2 Canonicalisation Requirements

Canonicalisation is the process of transforming heterogeneous raw data into a **common, standardised representation** that enables consistent analysis and cross-jurisdictional comparison. Without semantic stability, identical legal requirements can generate divergent supervisory interpretations purely due to technical differences in data formats, identifier conventions, or lifecycle representations.

Key canonicalisation tasks include:

- **Identifier normalisation**: Mapping entity identifiers (LEI), transaction identifiers (UTI), instrument identifiers (ISIN), and venue identifiers to canonical reference data, resolving synonyms and correcting errors.
- **Temporal alignment**: Standardising timestamps to common time zones and reporting cut-offs, enabling accurate sequencing and latency analysis.
- **Schema harmonisation**: Aligning field definitions, enumerations, and units across reporting templates and versions, supporting longitudinal and cross-dataset analysis.
- **Lifecycle reconstruction**: Transforming event-based reports (trade inception, valuation updates, terminations) into state-based representations (current positions, net exposures), enabling risk aggregation and network analysis.

The ECB's work on **additional supervisory data-quality checks** provides a concrete example: these checks are designed to enhance reporting quality and are applied systematically across significant institutions, with strong encouragement for firms to implement checks within their own systems **before submission**, thereby reducing downstream validation burden for NCAs. This approach aligns canonicalisation with data governance, embedding quality assurance at the source rather than treating it as a post-hoc correction[84].

From a convergence perspective, canonicalisation ensures that **NCAs apply the same verification functions to semantically equivalent data**, reducing interpretive variance and supporting horizontal consistency.

---

[84] *In January 2026 the ECB's Banking Supervision Data Division, in cooperation with National Competent Authorities within the Single Supervisory Mechanism (SSM), published a set of extra data quality checks in addition to those already included in the EBA's supervisory reporting frameworks (COREP and FINREP). These extra checks are aimed at enhancing the quality of supervisory reporting data. The ECB applies these checks in conjunction with other controls as part of the quality assessment of supervisory reporting data. In the communication, the ECB strongly encouraged significant institutions (SIs) within the SSM to implement the extra checks in their own reporting systems prior to the submission of data to NCAs. SIs are also encouraged to ensure compliance with the checks before reporting to the NCAs under the supervisory reporting frameworks.*

D.3 Validation, reconciliation, and uncertainty management

Validation operationalises legal reporting requirements into **measurable, executable rules** that systematically detect non-compliance, errors, and anomalies. Validation frameworks typically encompass multiple layers:

1. **Structural validation**: Schema compliance (data types, mandatory fields, cardinality constraints).
2. **Domain validation**: Business-rule checks (e.g. valuations must be positive, maturity dates must be in the future, leverage ratios must satisfy regulatory thresholds).
3. **Referential validation**: Cross-referencing against external reference data (LEI registries, ISIN databases, venue codes) to detect invalid identifiers.
4. **Temporal validation**: Consistency over time (e.g. position continuity, absence of unexplained jumps).
5. **Cross-dataset reconciliation**: Matching reports from multiple sources (e.g. both counterparties in EMIR, transaction reports vs. transparency data in MiFIR) to detect discrepancies.

However, **no reporting system is error-free**. Residual measurement error ε_t persists even after validation, arising from incomplete data, model uncertainty, or genuinely ambiguous cases. DDS must therefore **explicitly manage uncertainty** rather than assuming perfect observability:

- **Uncertainty qualifiers**: Metrics should be accompanied by data-quality scores or confidence intervals reflecting the reliability of underlying inputs.
- **Persistent errors as supervisory signals**: Repeated data-quality failures are not merely technical nuisances but indicators of weak governance, inadequate systems, or deliberate non-compliance, warranting supervisory escalation.
- **Uncertainty penalties in triage**: Entities with high data-quality uncertainty should receive higher priority scores or more conservative treatment to reflect elevated supervisory risk.

This approach aligns with supervisory guidance from ESMA and the ECB, which treat data quality as a **first-order risk indicator** and governance signal, not a secondary administrative issue.

Appendix E: Limitations of data-driven supervision and related mitigation strategies

E.1 Dependence on data availability and quality

DDS is fundamentally constrained by the **availability, coverage, completeness, and accuracy of regulatory reporting data**. If critical data are missing, inconsistent, or systematically biased, supervisory indicators will be unreliable, potentially misleading supervisors or creating blind spots.

Measurement error cannot be fully eliminated, even with extensive validation and reconciliation. Residual errors propagate through indicators, and non-linear transformations (aggregations, network metrics) can amplify small input errors into large output distortions. For example, a single mis-reported counterparty identifier can create spurious network links, biasing systemic-importance rankings and misdirecting supervisory attention.

Mitigation strategies:

- **Data-quality as first-order priority**: NCAs must invest in robust validation pipelines, reconciliation, and feedback loops to entities, treating data quality as a supervisory outcome, not a technical nuisance[85].
- **Reporting handbook:** There is currently some ambiguity as to how certain data need to be reported, and a common understanding of certain reporting fields needs to be established. A so-called reporting handbook would provide more clarity on the reporting specifics[86].
- **Uncertainty quantification**: indicators should be accompanied by confidence intervals or data-quality flags, enabling supervisors to interpret signals in light of underlying data reliability.
- **Hybrid approaches**: combining data-driven indicators with traditional supervisory tools (interviews, on-site inspections, whistleblower reports) to cross-validate findings and fill gaps.

E.2 Model risk: mis-specification, overfitting, and concept drift

Model risk—the risk that indicators, scoring functions, or early-warning models are mis-specified, biased, or unstable—is intrinsic to DDS. Sources of model risk include:

- **Mis-specification**: indicators may fail to capture relevant risk dimensions or may conflate compliance with risk (e.g. high rejection rates may reflect either weak controls or complex, high-quality activity with strict validation).
- **Overfitting**: models trained on historical data may capture noise or regime-specific patterns that do not generalise to new conditions, producing false positives or negatives.
- **Concept drift**: financial markets and entity behaviours evolve; indicators that performed well historically may degrade over time as underlying relationships change (e.g. risk drivers, reporting practices, business models).
- **Gaming and manipulation**: sophisticated entities may alter behaviour to "optimise" indicators without improving underlying compliance or risk, undermining supervisory effectiveness.

Mitigation strategies:

- **Champion–challenger frameworks**: maintaining multiple indicator variants and periodically comparing performance to select the most stable and predictive.
- **Back-testing and out-of-sample validation**: assessing indicator performance on historical data not used in calibration, and monitoring performance continuously in production.
- **Drift detection and recalibration**: implementing statistical process control to flag when indicator distributions or prediction errors shift significantly, triggering review and recalibration.
- **Explainability and human oversight**: ensuring indicators link clearly to risk drivers and regulatory requirements, enabling supervisors to challenge and contextualise model outputs.

E.3 Reconciliation and common-mode errors

While reconciliation (e.g. matching EMIR counterparty reports) is a powerful data-quality tool, it **cannot detect common-mode errors**—systematic mistakes made identically by both reporting

[85] For instance, EMIR already requires that all the information reported is correct, whether it is information reported under Article 9 or any other article of that Regulation. Therefore, what is commonly considered being a data quality issue appear to be a non-compliance with the requirement to report correct information. The same approach shall be applied to any other regulatory reporting.

[86] Mazzaferro. F. (12 July 2022), letter to the Commissioner McGuinness on ESRB's view regarding data quality issues and risks for financial stability.

parties. For example, if both counterparties mis-classify a transaction's asset class or apply the same incorrect valuation model, reconciliation will show agreement but the data will be wrong.

This limitation is particularly relevant for **valuation consistency monitoring**: if market participants use similar (but flawed) models, cross-entity valuation dispersion may be low, masking systematic mispricings that only become visible during stress events. DDS must therefore complement reconciliation with **cross-dataset checks** (e.g. comparing reported valuations to independent market data) and **stress-testing of valuation models**.

E.4 Segmentation and false positives

Poorly designed indicators can generate **false positives** when applied to heterogeneous populations without appropriate segmentation. For example, valuation dispersion may be legitimately high for illiquid, bespoke instruments but indicative of problems for liquid, standardised products. Applying a single threshold to both creates noise.

Mitigation: indicators should be **segmented by business model, product type, and market conditions**, with thresholds calibrated separately for each segment. Peer-group benchmarking should compare entities to **relevant peers**, not the entire population.

E.5 Legal and confidentiality constraints

Data integration across supervisory missions or jurisdictions may face **legal barriers**: data collected under one legal mandate (e.g. EMIR derivatives reporting) may not be usable for other purposes (e.g. conduct supervision) without explicit legal basis. Cross-border data sharing faces additional constraints under data protection and confidentiality regimes.

Mitigation: NCAs and ESMA must ensure that legal frameworks explicitly authorise multi-purpose use of regulatory data for supervisory and macroprudential purposes, consistent with data minimisation and proportionality principles.

- **Bibliography**

Adrian, T. and Brunnermeier, M.K. (2016). 'CoVaR', American Economic Review, vol. 106, no. 7, July 2016, (pp. 1705–41)

Aldasoro, I., Delli Gatti, D., & Faia, E. (2016). Bank networks: Contagion, systemic risk and prudential policy (BIS Working Papers No. 597). BIS.

Angelova, V., W. S. Dobbie and C. S. Yang, (2023). Algorithmic Recommendations and Human Discretion. NBER Working Paper No. 31747.

Atlan (2025). A Guide to Regulatory Reporting Automation in 2025.

A2ii, Financial Stability Institute, and IAIS (2022). Joint note on SupTech in insurance supervision. BIS.

Banca d'Italia (2025). EMIR Data Analytics for Research, Financial Stability and Supervision.

Basel Committee on Banking Supervision, (2025), Working Paper 45 Lesson on supervisory effectiveness – a literature review, BCBS-BIS.

Broeders, D., & Prenio, J. (2018). Innovative technology in financial supervision (suptech): The experience of early users (FSI Insights No. 9). BIS.

Brunnermeier, M.K. and Cheridito, P. (2019). Measuring and Allocating Systemic Risk, (April 1, 2019).

Central Bank of Ireland (2025a) Transforming regulation and supervision.

Central Bank of Ireland (2025b) Our Approach to Supervision.

Central Bank of Ireland (2025c) Regulating & Supervising well – a more effective and efficient framework.

Chen, N., Liu, X., and Yao, D. D. (2014). Systemic risk in financial systems [Working paper]. The Chinese University of Hong Kong.

Choy, A., et al..(2013). Using micro data to assess systemic risk (IFC Bulletin No. 64, paper 12). BIS

CSSF (2022). Communiqué on Monitoring the quality of transaction reports received under Article 9 of EMIR.

CSSF (2023). Communiqué EMIR Reporting – Prepare for the Next Challenges.

CSSF (2026a). Communiqué on Monitoring the quality of transaction reports received under Article 26 of MiFIR.

CSSF (2026b). Communiqué on the Outcomes of the 2025 SFTR Data Quality indicators review.

CSSF-CAA (2024). Derivative Market in Luxembourg.

De Nederlandsche Bank (2015). Supervision of Behaviour and Culture. Foundations, practices & future developments.

Buckley, R.P., Arner, D.W., Dirk A. Zetzsche, D.A. (2023). FinTech - Finance, Technology and Regulation, Cambridge University Press.

Draghi, M. (2024). The Future of European Competitiveness—A Competitiveness Strategy for Europe.

European Banking Authority (2020). Decision of the European Banking Authority of 28.04.2020 adopting a methodology for the conduct of peer reviews. EBA.

European Banking Authority (2025). SupTech Can Make Anti-Money Laundering and Countering the Financing of Terrorism Supervision More Effective. EBA.

European Central Bank (2010). Modelling systemic risk. ECB.

European Central Bank (2015). Systemic risk, contagion and financial networks, in Financial Stability Review, November 2015. ECB.

European Central Bank - SSM (2015). Policy Options for the Definition of the Materiality Threshold for Credit Obligations, SSM consultation paper. ECB Banking Supervision.

European Central Bank (2024a). SREP supervisory methodology. ECB Banking Supervision.

European Central Bank (2024b). Operational & ICT risk SREP methodology. ECB Banking Supervision.

European Central Bank (2024c). Supervisory Methodology 2024. ECB Banking Supervision.

European Central Bank (2025a). Simplification of the European Prudential Regulatory and Supervisory Reporting Framework. 2025. ECB.

European Central Bank (2025b). Supervisory Methodology 2025. ECB Banking Supervision.

European Central Bank (2025c). ECB Annual Report on supervisory activities 2024. ECB Banking Supervision.

European Central Bank (2026). Additional supervisory data quality checks. ECB Banking Supervision. (public portal material).

European Central Bank (2025e). Aggregated Data Quality Tables, 2025 Q1. ECB.

European Fund and Asset Management Association (2025). EFAMA response to ESMA discussion paper on the integrated collection of funds data. EFAMA.

European Securities and Markets Authority (2022). ESMA Strategy 2023–2028. ESMA.

European Securities and Markets Authority (2024a). Report on the Quality and Use of Data 2023. ESMA.

European Securities and Markets Authority (2024b). ESMA Identifies Measures to Further Enhance Depositary Supervision. ESMA.

European Securities and Markets Authority (2025a). ESMA 2024 Annual Report. ESMA.

European Securities and Markets Authority (2025b). Report on the Quality and Use of Data 2024. ESMA.

European Securities and Markets Authority (2025c). MiFIR Reporting. (public portal material). ESMA.

European Securities and Markets Authority (2025d). Supervision and Convergence Portal. ESMA.

European Securities and Markets Authority (2025e). Principles on third-party risks supervision. ESMA.

European Securities and Markets Authority (2025f). Peer Review Methodology and Peer Reviews (portal and reports). ESMA.

European Securities and Markets Authority (2026). Principles on risk-based supervision. ESMA.

European Systemic Risk Board (2014). ESRB Risk Dashboard – Description of the Indicators (Annex II). ESRB.

European Systemic Risk Board (2018). The ESRB Handbook on Operationalising Macroprudential Policy. ESRB.

European Systemic Risk Board (2025). Systemic Liquidity Risk: A Monitoring Framework. ESRB.

European Systemic Risk Board (2026). Risk Dashboard — Annex II: Description of Indicators, March 2026 . ESRB.

European Commission (2019). Fitness Check of EU Supervisory Reporting Requirements (Staff Working Document). European Commission.

European Commission (2023). Better Regulation Portal and Annual Burden Survey 2023. 2023. European Commission.

European System of Central Banks / European Central Bank (2010). Recent Advances in Modelling Systemic Risk Using Network Analysis. ECB.

Financial Stability Board (2020). The Use of Supervisory and Regulatory Technology by Authorities and Regulated Institutions. FSB.

Financial Stability Board (2024), The Financial Stability Implications of Artificial Intelligence. FSB.

Financial Stability Board (2025). Thematic Review on FSB Global Regulatory Framework for Crypto-asset Activities. FSB.

Financial Stability Institute (2019). The suptech generations (FSI Insights on policy implementation No. 19). BIS.

Galán, A., & Svetiev, Y. (2025). Peer review in networked financial supervision: the evolution of ESMA's collegial panopticon. *Yearbook of European Law*.

Galytix (2022). Banks Must Act on their Early Warning Systems or Risk ROE Downturn.

Hogan Lovells (2016). MiFID II Market Data Reporting.

International Monetary Fund (2000). Toward a Framework for Assessing Data Quality (Working Paper WP/00/212). IMF.

International Monetary Fund (2003). Data Quality Assessment Framework (DQAF).(and related factsheets). IMF.

International Monetary Fund (2025). AI Projects in Financial Supervisory Authorities (Working Paper WP/25/199). IMF.

International Organization of Securities Commissions (2009). Guidelines to Emerging Market Regulators Regarding Requirements for the Oversight of Markets (Final Report). IOSCO.

International Organization of Securities Commissions (2010). Principles Regarding Cross-Border Supervisory Cooperation. IOSCO.

International Organization of Securities Commissions (2017). Objectives and Principles of Securities Regulation. IOSCO.

International Organization of Securities Commissions (2023). Principles for the Regulation and Supervision of Commodity Derivatives Markets. IOSCO.

International Organization of Securities Commissions (2025). Thematic Review on the Implementation of its Crypto and Digital Asset Recommendations. IOSCO.

International Organization of Securities Commissions, Financial Stability Board (2025)., Progress towards implementing comprehensive regulatory frameworks for crypto-asset activities. IOSCO.

Letta, E. (2024). "Much more than a Market", Report, European Council.

Mazzaferro. F. (2022), Letter of 12 July 2022 to the Commissioner McGuinness on ESRB's view regarding data quality issues and risks for financial stability. ESRB

Toronto Centre. (2018). Risk-based supervision. Toronto Centre Global Leadership in Financial Supervision.

Toronto Centre (2021). SupTech: Leveraging technology for better supervision (updated)

United Nations Capital Development Fund (2023). Empowering data-driven supervision in developing countries. UNCDF.

UK Finance (2019). UK Finance response to EC fitness check on supervisory reporting.

Yanquen, E., Livan, G., Montañez-Enriquez, R., Martinez-Jaramillo, S. (2022). Measuring systemic risk for bank credit networks: A multilayer approach, *Latin American Journal of Central Banking*, Volume 3, Issue 2, 2022.

Yıldırım, Y., & Sanyal, A. (2022). Evaluating the effectiveness of early warning indicators: an application of receiver operating characteristic curve approach to panel data. *Scientific Annals of Economics and Business*, *69*(4), 557-597.

World Bank (2028). SupTech: Technology Solutions for Market Conduct Supervision. World Bank.

World Bank (2025). *AI for risk-based supervision.* World Bank.