

Where Digital Sovereignty Lives: The Micro-Foundations Problem in the Measurement of a Multi-Level Construct

Martin Kaloudis

martin@kaloudis.de

AKAD University

Research Article

Keywords: digital sovereignty, multi-level measurement, composite indicators, ecological fallacy, construct validity

Posted Date: October 16th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10848313/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Digital sovereignty is measured primarily at the state level, yet it is produced through procurement, contractual, and architectural decisions at the organisational and installation levels. This article treats the gap between these levels as a measurement problem. Transposing Coleman's macro-micro-macro schema shows that country indices capture the macro endpoints while much of the causal mechanism between them remains unmeasured. Three problems follow: ecological inference, composition error, and construct heterogeneity. The central claim is that the dependency core of digital sovereignty behaves as a compilation construct: critical minima and configurations constrain the whole, so compensatory aggregation tends to overstate sovereignty precisely where unresolved dependencies matter most. A calibrated illustration shows the stakes. Using published Digital Sovereignty Index data, the EU would rank 22nd under mean aggregation and 50th under minimum aggregation. In a simulation of 10,000 synthetic units, mean-based and minimum-based rankings correlate at only $\rho = .16$, while roughly one in five units in the top mean quartile contains a critical minimum in the bottom quartile. The article derives five requirements for level-consistent measurement: declared level ontology, compilation-faithful aggregation, tested isomorphism, observation of the meso level, and level-matched confounder discipline. It concludes with implications for index construction, the evaluation of sovereignty policy, and theories of digital dependency.

1. Introduction: A Construct Measured at One Level, Produced at Another

States are increasingly ranked according to their digital sovereignty. Country indices combine indicators of technology, infrastructure, industrial capacity, governance, and trade into national scores, and these scores have become part of policy debates on strategic autonomy and technological dependency (Kaloudis, 2022b; Madiaga, 2020; Schmitz & Seidl, 2023). Yet the indicators captured by such indices are not digital sovereignty itself. In the operational sense used here, digital sovereignty refers to the capacity for self-determined action in the procurement, use, and development of digital technologies under dependencies that remain manageable (Kaloudis, 2021, 2022a). This capacity is realised in concrete arrangements: a contract that does or does not provide a credible exit path, an application that can or cannot be re-hosted on another server, a cryptographic key that is held by the organisation itself or otherwise under its control. Two organisations operating under the same national framework conditions can therefore have very different dependency profiles. This gives rise to the article's central dilemma: digital sovereignty is measured predominantly at the level of states, while much of it is produced at the level of organisations and individual installations. The theoretical relationship between these levels remains largely unresolved.

This is more than a technical issue of index design. Without considering how macro-level framework conditions translate into lower-level decisions and how those decisions in turn aggregate into national outcomes, it remains unclear what a country-level sovereignty score can legitimately say beyond country-level description. The critique is deliberately field-specific: the index used for illustration in

Section 3.3 (Kaloudis, 2022b) is one example of the measurement approach examined here, and the argument does not deny the value of such instruments for relevant macro-level questions. The problem lies in the inferences drawn from them. Other areas of social measurement have developed tools for exactly this situation. Robinson (1950) showed that relationships observed in aggregate data need not hold for the underlying units. Coleman (1990) placed the micro-level mechanism at the centre of explanations linking macro conditions and macro outcomes. Kozlowski and Klein (2000) distinguished between conditions under which lower-level properties can be composed into higher-level constructs and those under which their configuration matters instead. Digital sovereignty measurement has so far made little systematic use of this multi-level apparatus, even though it produces precisely the kind of aggregate claims for which these distinctions matter.

This article develops a level-consistent account of digital sovereignty measurement that addresses this gap. Section 2 transposes Coleman's macro-micro-macro schema to the digital sovereignty context and identifies the unobserved organisational level between national conditions and national outcomes. Section 3 discusses three problems that arise when this middle level remains implicit: ecological inference, composition error, and construct heterogeneity. It then argues that the dependency core of digital sovereignty is better understood as a compilation construct, in which critical minima and configurations matter in ways that simple averaging cannot preserve. Section 4 converts these problems into five requirements for level-consistent measurement, and Section 5 derives implications for index construction, policy evaluation, and theories of digital dependency. Section 6 concludes. The argument is instrument-agnostic throughout: particular indices serve as examples, but the underlying problem applies to any measure that moves between states, sectors, organisations, and installations without making the bridging assumptions explicit.

2. The Coleman Schema for Digital Sovereignty

2.1 The Diagram and a Procurement Example

Coleman's diagram essentially shows how macro-level conditions lead to macro-level outcomes through micro-level decisions (Coleman, 1990). The descending arrow runs from general framework conditions to the concrete situation in which an actor must make a decision. For digital sovereignty, these conditions include regulation, industrial structure, skills, market conditions, and geopolitical exposure. They shape the choices faced by ministries, public agencies, firms, and operators of critical infrastructure. The horizontal arrow represents the actual procurement or architecture decision: which provider is selected, what the contract stipulates, whether a system can be migrated, and who controls critical elements such as encryption keys. The ascending arrow then shows how thousands of such structures aggregate into a broader national dependency profile. Figure 1 transposes this logic to the context of digital sovereignty.

Consider a simple example: a sovereign-cloud strategy in the public sector. At the macro level, a government may introduce procurement rules, certification requirements, or funding programmes

intended to strengthen sovereign solutions. A public agency must then translate these relatively abstract requirements into a concrete procurement decision. In practice, a tender may look highly sovereign on paper – hosting in Europe, compliance requirements, data-location clauses – while the decisive dependencies are buried deeper in the contract and architecture. The agency might, for example, choose a platform that technically permits exit but use proprietary services so extensively that later migration becomes costly and operationally risky; key custody, interfaces, and the actual exit procedure can make the difference between a formal and a realistic option. This single decision creates a concrete dependency structure that then sits alongside the decisions of hospitals, ministries, network operators, and thousands of other organisations. Only much later do parts of these decisions reappear in national statistics, which is why a country index captures mainly the macro-level starting point and fragments of the eventual outcome while the descending, horizontal, and ascending arrows in between remain largely unobserved. Comparing one national index with another may therefore confirm agreement between two macro-level measurements (Arndt, 2008), but it cannot show whether the assumed macro-to-micro-to-macro mechanism actually holds. This is less a weakness of any particular index than a structural limitation of the measurement genre itself.

2.2 Why the Diagonal Cannot Be Assumed

The diagonal in Fig. 1 would be easy to accept if national framework conditions translated more or less directly into organisational outcomes. In practice, they do not. Experience from the provision of IT services to the public sector reveals a recurring pattern: organisations operating under nearly identical national and even regional framework conditions arrive at very different dependency structures. Two public agencies in the same federal state may be subject to the same procurement law, the same certification requirements, and essentially the same provider market, yet one retains meaningful exit options, avoids unnecessary proprietary dependencies, and keeps critical control points in its own hands, while the other becomes deeply locked into a single provider. The difference arises from organisational decisions: how requirements are specified, which architectural trade-offs are accepted, how contracts are negotiated, and how much attention is paid to substitutability after implementation. This makes substantial within-country variation at least plausible and potentially as important as between-country variation. Whether it is in fact larger is an empirical question, but one that a country-level index cannot answer. The problem becomes even clearer because important channels of digital dependency are not contained within national borders in the first place. Firms, platforms, and network structures span jurisdictions, and control over strategically important nodes can create power regardless of the formal national framework in which an organisation operates (Farrell & Newman, 2019).

Sovereignty policy adds another layer of indirection. Governments usually intervene at the beginning of the chain: they change procurement rules, introduce certification schemes, support particular technologies, or pursue “sovereign cloud” strategies. In Coleman’s terms, these interventions act on the descending arrow by changing the situation in which organisations make decisions. What policymakers actually seek to change, however, lies on the horizontal arrow: the provider selected, the architecture built, the contractual exit rights secured, and the dependencies that remain after implementation. These

decisions are rarely observed systematically. Years later, policymakers return to national statistics and ask whether the country has become more sovereign. By then, the figures reflect not only the policy intervention but also market developments, investment cycles, technological change, organisational discretion, and many other influences. The field therefore intervenes at one end of the causal chain and measures at the other while observing surprisingly little of what happens in between. This is the “missing middle”: the organisational and sectoral level at which sovereignty policy either produces changed dependency structures or fails to do so.

3. Three Aggregation Errors

3.1 Ecological Inference

The first problem is a classical one. Robinson (1950) showed that a relationship observed in aggregate data need not hold for the individual units beneath it – either in strength or in direction. The same warning applies to digital sovereignty. If wealthier countries score highly on both a digital sovereignty index and a digitalisation index, this does not mean that the most digitalised organisations in those countries are also the most sovereign. Nor does a positive country-level relationship between digital sovereignty and cybersecurity show that an organisation becomes more secure because it has greater control over its own technology stack. Both claims move from a relationship between countries to a relationship between organisations without evidence for the intermediate step. Openshaw’s (1984) modifiable areal unit problem complicates matters further. Results depend partly on how units of analysis are drawn, and national borders are a particularly awkward boundary for technology infrastructures that may rely on cloud regions, platforms, providers, and shared services distributed across multiple jurisdictions. A country is a convenient statistical unit. It is not necessarily the unit in which the relevant dependency structure actually exists.

The more troubling possibility is that the relationship could reverse once we move down one level. Suppose that, across countries, wealth drives both digitalisation and sovereignty scores upward because richer countries possess more infrastructure, skills, and institutional capacity. Within those same countries, however, the organisations most deeply integrated into global platforms might also be the hardest to move away from them. They could therefore be highly digitalised while being comparatively weak in substitutability. The between-country relationship would then be positive while the within-country relationship would be negative. Whether this pattern exists in reality is an empirical question; the important point is that current country-level measurement would be unable to detect it. This calls for something like a “discipline of silence”: a country-level finding should remain a country-level finding unless separate evidence supports the transition to another level. In consulting and sales contexts, this discipline is remarkably easy to lose. One slide shows that Country X ranks twelfth on a given digital indicator; one slide later, the discussion has already shifted to “therefore your company should ...”. The transition feels natural because the topic is the same, but analytically an entire level has been crossed without new evidence. The ranking may provide useful context. On its own, however, it cannot justify a recommendation for the organisation.

3.2 Composition Error: Sovereignty as a Compilation Construct

Multi-level theory distinguishes between composition and compilation constructs (Kozlowski & Klein, 2000). In a composition model, lower-level observations are manifestations of essentially the same property and can therefore be aggregated into a higher-level measure. Organisational climate is a classic example: individual employees report how they perceive the climate, and the organisational construct can be represented by aggregating those perceptions if sufficient agreement exists. Compilation works differently. Here, the higher-level construct emerges from the configuration of lower-level elements that are neither identical nor interchangeable. Team performance, for example, depends on how different skills work together and whether particular roles are adequately filled. Strong performance in one part of the system does not automatically offset failure in another. The distinction therefore has direct measurement consequences. Composition permits averaging; compilation requires an aggregation rule that preserves the structure and interaction of the elements being combined. Digital sovereignty, particularly where dependencies are concerned, has much more in common with the second case.

The first reason is the dominance of critical components. Technology estates contain components whose importance is disproportionate to their numerical share of the overall estate. A public agency may have hundreds of well-managed, portable, competitively procured applications while a single irreplaceable specialist system remains tied to one provider. That system might handle a legally mandated administrative process, depend on proprietary interfaces, and offer no realistic migration path within an acceptable time frame. Viewed as an average, the infrastructure would make the organisation appear relatively independent even though the failure or withdrawal of one provider could halt a critical process. An organisation whose system landscape is excellent on average but whose irreplaceable core system constitutes an unmitigated dependency is not averagely sovereign; it is trapped at one point, and that point governs the whole. For this reason, the dependency core of sovereignty has minimum-like semantics. Critical weaknesses place an upper bound on what the system as a whole can credibly claim – a logic closely related to the chokepoints described by Farrell and Newman (2019).

Relationality raises a different problem. Many sovereignty-relevant properties cannot meaningfully be assigned to a single component in isolation because their value depends on the surrounding configuration. Redundancy is an intuitive example. Two workloads may run in separate data centres and therefore appear redundant when each is assessed individually. If both environments ultimately depend on the same hyperscaler, the same identity layer, or the same contractual control point, the apparent redundancy may disappear precisely in the scenario for which it was created. The same applies to vendor diversity. Diversity is a property of a portfolio, not of an individual server, contract, or application. Substitutability is similarly relational: whether a component can be replaced depends on available alternatives, interfaces, skills, data portability, and the dependencies of surrounding components. An assessment that scores each resource separately and then adds the scores cannot adequately capture these relationships because the relevant information lies in the patterns connecting the resources.

The resulting aggregation problem resembles Kremer’s (1993) O-ring production logic. In an O-ring setting, the value of the whole depends multiplicatively on the quality of several necessary tasks, so a severe failure in one task can undermine the contribution of all the others. Dependency structures often behave in this way: ninety-nine substitutable components do little to protect an organisation if the hundredth is both critical and practically irreplaceable. The relationship between mean-like and bottleneck-like aggregation can be expressed by the family of generalised means M_p in Eq. (1) (Hardy et al., 1934). At $p = 1$, the operator is the arithmetic mean; as p tends towards negative infinity, the result approaches the minimum. Choosing an aggregation operator therefore amounts to choosing a substantive theory of the extent to which strengths in one part of a system may compensate for weaknesses elsewhere. This is where the composition error becomes decisive. Weighted sums of normalised indicators implicitly assume compensability even when the underlying construct is governed by critical dependencies and configurations. The composite-indicator literature has long noted that compensatory and non-compensatory aggregation encode different assumptions (Munda, 2008; OECD/JRC, 2008; Mazziotta & Pareto, 2013; for a recent review of open issues, see Alaimo & Seri, 2024). For digital sovereignty, the problem goes further: if the dependency core is compilation-like, treating compensation as the default is a category error about the construct itself. Bottlenecks, minima, and configuration-sensitive operators therefore follow from the measurement problem rather than from a preference for one statistical method over another.

$$M_p(\mathbf{s}; \mathbf{w}) = \left(\sum_{i=1}^n w_i s_i^p \right)^{1/p}, \quad \text{with} \quad M_1 = \sum_{i=1}^n w_i s_i \quad \text{and} \quad \lim_{p \rightarrow -\infty} M_p = \min_i s_i \quad (1)$$

3.3 A Calibrated Illustration: How Much the Aggregation Choice Matters

Sensitivity to aggregation can be illustrated with field-specific data by moving one level upward. The European Union is itself a composite unit, yet it is often treated in debates on strategic autonomy as a single political and economic actor. This raises the same measurement question in especially visible form: how should the sovereignty of the whole be derived from the scores of its members? Using published Digital Sovereignty Index data for all 27 EU member states, the answer changes substantially with the aggregation rule. Under a composition logic using the mean of member-state scores, the EU obtains a score of 1,698 and would rank 22nd among 214 states. Under a criticality logic using the lowest member-state score, the EU obtains 692 – determined by Bulgaria – and would rank 50th. The difference is therefore 28 ranking positions, produced solely by the choice of aggregation procedure. The underlying heterogeneity is substantial: Ireland scores 3,305 and Bulgaria 692, a factor of 4.8 between the strongest and weakest member state. This exercise should not be read as an attempt to identify the EU’s “true” sovereignty score. Its purpose is narrower: to show how strongly any statement about a composite unit depends on an aggregation rule that often remains implicit. Table 1 and Fig. 2a present the comparison.

Table 1
One composite unit, two aggregation models: the EU-27 based on published DSI data.

Aggregation model	Construct logic	EU score (DSI points)	Implied world rank (of 214)	Governing element
Weighted mean (composition)	deficits compensable	1,698	22	all members equally
Minimum (compilation limit)	bounded by critical worst case	692	50	weakest member (Bulgaria)
Spread within unit	—	max/min = 4.8	28 rank positions apart	Ireland (3,305) vs. Bulgaria (692)
<i>Note: All 27 member states are covered by the underlying index data. The exercise is illustrative only: it shows the sensitivity of any statement about a composite unit to the unstated aggregation operator and does not claim to identify the EU's "true" score.</i>				

A calibrated simulation shows that the point is not specific to the EU example. The simulation generated 10,000 synthetic units with 25 components each from a Beta(0.51, 2.65) distribution fitted to the distribution of normalised DSI country scores. Each unit was then aggregated once by its mean and once by its minimum. The rank correlation between the two orderings is low: Spearman's $\rho(\text{mean}, \text{min.}) = .16$. More importantly, 19.8% of the units that fall into the top quartile under mean aggregation contain a critical minimum in the bottom quartile of the minimum distribution. In other words, almost one in five apparently high-performing units contains a weakness that the compensatory aggregate largely conceals. The result is not specific to the fitted Beta distribution. Across alternative specifications using a uniform distribution, Beta(2,5), and Beta(5,2), rank correlation ranges from .16 to .44, while the share of masked-fragility cases ranges from 10.6% to 19.8%. Actual digital infrastructures will correspond to neither pure operator because not every component is equally important; operational aggregation will therefore lie somewhere between the arithmetic mean and the minimum. The simulation is intended to bound the stakes of operator choice, not to prescribe a universal rule. Table 2 and Fig. 2b show the results; replication material is available at [anonymised repository link].

Table 2
Calibrated simulation: divergence between mean-based and minimum-based aggregation.

Design element / result	Value
Synthetic units × components per unit	10,000 × 25
Component score distribution	Beta(0.51, 2.65), fitted to normalised DSI country scores
Rank correlation ρ (mean aggregate, minimum aggregate)	.16
Pearson r (mean aggregate, minimum aggregate)	.19
Top-mean-quartile units with bottom-quartile critical minimum (masked fragility)	19.8%
<i>Note: Seed-fixed simulation; replication code available from the author. The pattern is robust to calibration: across alternative component distributions (uniform; Beta(2,5); Beta(5,2)), rank correlation ranges from .16 to .44 and the share of masked fragility from 10.6% to 19.8%. Real-world estates contain mixes of critical and non-critical components, so operational aggregation lies between the pure operators; the simulation bounds the stakes of that choice.</i>	

3.4 Construct Heterogeneity: Does Sovereignty Mean the Same Thing at Every Level?

A third difficulty goes deeper because it concerns the meaning of the construct itself. A multi-level approach must ask whether digital sovereignty at the state, organisational, and installation levels is genuinely the same construct, or whether the same term is being used for related but distinct concepts. There are good reasons for caution. At the state level, sovereignty carries meanings that organisations simply do not possess, including jurisdiction, political authority, and international recognition (Couture & Toupin, 2019; Pohle & Thiel, 2020). Mueller (2020) goes further and questions whether sovereignty remains an appropriate concept even for states in cyberspace, where infrastructure, platforms, and control frequently extend beyond territorial boundaries. At the same time, there is a recognisable functional core that extends across levels. The capacity to exit a dependency, replace a provider or technology, and enforce control over relevant resources matters to states, organisations, and individual installations alike. Digital sovereignty may therefore not be fully identical at every level, but the cross-level similarity is not merely linguistic either. The measurement problem is to determine how far that continuity actually extends.

Sartori's (1970) ladder of abstraction provides a useful framework for doing so. Digital sovereignty has become a concept that routinely moves between parliamentary debate, corporate strategy, and the server room, and each move carries a risk of concept stretching. The solution should not be to assume either complete equivalence or complete difference in advance. Instead, the degree of cross-level continuity must be specified conceptually and, where scores are compared, tested empirically using measurement-invariance approaches such as those developed by Meredith (1993). Table 3 sketches

what such a mapping might look like by distinguishing among components that are constitutive at a particular level, those that appear only distributionally across lower-level units, and those that exist as direct attributes. The middle rows of the table suggest a plausible translatable core: dependency exposure, substitution capability, and enforcement capacity have recognisable counterparts at the state, organisational, and installation levels. Other components do not translate in the same way. Jurisdictional authority is constitutive at the state level but does not descend to organisations or installations, while exit rights and key custody are direct properties at lower levels and appear at the state level only as distributions across many organisations and systems. Table 3 should therefore be read as an initial hypothesis about where the construct is isomorphic, not as proof that sovereignty means exactly the same thing everywhere.

Table 3

Sketch of a cross-level component mapping for digital sovereignty.

Construct component	State level	Organisational level	Installation level
Jurisdictional authority / recognition	constitutive	absent	absent
Dependency exposure	distributional (aggregate)	portfolio property	direct attribute
Substitution capability	industrial capacity	demonstrated migration ability	architectural property
Enforcement capacity	regulation and oversight	contract and audit rights	technical control (keys, configuration)
Exit rights / key custody	only as distributions	direct attribute	direct attribute
Note: “Constitutive” = defining at that level; “distributional” = exists only as an aggregate over lower-level units. The translatable core – rows two to four – is the domain in which cross-level comparison is meaningful in principle; rows one and five mark the non-isomorphic remainder.			

4. Five Requirements for Level-Consistent Measurement

The three pathologies outlined above translate directly into measurement requirements. These requirements are deliberately instrument-agnostic. They apply to any instrument that claims to measure digital sovereignty, whether it operates at the level of states, sectors, organisations, or individual installations. The aim is not to prescribe a particular metric but to provide a set of criteria against which any such instrument can be assessed.

R1 – Declare the level ontology. Every measurement instrument should state clearly the level at which its object is observed, the level at which its results are intended to support inferences, and the assumptions required to move from one level to the other. If a national index is used to support organisational decisions, the link between country conditions and organisational outcomes must not remain implicit. The same applies in the opposite direction when organisational observations are

aggregated into sectoral or national claims. The logical leap from “Country X ranks twelfth” to “therefore this organisation should ...” is exactly the kind of undeclared level crossing this requirement is intended to prevent.

R2 – Use compilation-faithful aggregation. Whenever lower-level elements are combined into a higher-level score, the aggregation rule must reflect how the construct is actually formed. For the dependency core of digital sovereignty, this means taking two aspects into account at the same time: critical minima and configuration. An unresolved dependency in a critical component must be able to constrain the aggregate, while relational properties such as vendor diversity or common-mode failure sources must be represented at the level of the overall configuration rather than disappearing into the scores of individual components. A compensatory weighted mean meets neither requirement. It allows strengths elsewhere to offset critical weaknesses and treats configuration as though it could be reduced to independent attributes.

R3 – Test rather than assume isomorphism. Cross-level measurement is defensible only if the construct shows sufficient continuity across the levels being compared. The first step is conceptual: the components of digital sovereignty should be mapped explicitly to the levels at which they exist and to the form they take there. The second step is empirical. When scores are compared across levels or between units that may interpret the construct differently, measurement invariance should be tested following approaches such as Meredith (1993). If invariance does not hold, the appropriate conclusion is not to preserve comparability by assumption. It is to report that the scores, as currently measured, are not comparable.

R4 – Observe the middle. The most important empirical gap lies neither at the national level nor at the level of individual technical components, but between the two. Sectoral and organisational dependency structures form the meso level through which macro conditions are translated into concrete installations and through which those installations condense into broader patterns. This level should therefore be a priority for future data collection. Emerging installation-level assessments make such observation increasingly feasible, but the requirement is architectural rather than tied to a particular method: lower-level observations must be aggregable according to R2 and comparable under the conditions of R3. Otherwise, more data may be collected without actually closing the “missing middle”.

R5 – Match confounder discipline to the level of analysis. Different levels bring different sources of spurious correlation. At the macro level, national prosperity is an obvious candidate because it influences many of the indicators that feed simultaneously into measures of sovereignty, digitalisation, and institutional capacity. At the meso level, organisational size and sector may play a similar role. Validity claims must therefore survive conditioning on the dominant confounders of the level at which they are made. Experience with the measurement of governance and sovereignty at the macro level provides a warning: correlations that look convincing in aggregate form can weaken or disappear once common background factors are taken into account (Arndt, 2008; Kaloudis, 2024). This lesson should not have to be rediscovered at every level below.

5. Implications

5.1 For Index Construction

National sovereignty indices remain useful for the questions they were actually designed to answer: cross-country comparisons, description of macro-level framework conditions, tracking change over time, and directing policy attention to weaknesses. What changes under the multi-level perspective is the claim attached to the score. A single sentence in the documentation would already remove much of the ambiguity; wording such as the following would suffice: “This index measures the macro-level conditions of digital sovereignty and does not support direct inferences about the sovereignty of individual organisations or installations.” Such a limitation costs nothing analytically but prevents the most common misuse across levels. The same discipline is required within the index itself. Where indicators or pillars represent dependency structures with bottleneck characteristics, compensatory aggregation should not simply be assumed. The weighted mean should bear the burden of justification rather than enjoy a methodological default (Munda, 2008). The EU example in Section 3.3 shows why this matters: changing the aggregation operator alone moved the same composite unit by 28 positions in the implied world ranking. Formative measurement provides an appropriate framework for thinking about the problem because the indicators form the construct rather than merely reflect an underlying latent variable (Diamantopoulos & Winklhofer, 2001). Once indicators are constitutive, the rule by which they are combined becomes part of the substantive theory of sovereignty itself.

5.2 For the Evaluation of Sovereignty Policy

The Coleman schema also changes what counts as evidence that sovereignty policy has worked. Governments intervene primarily at the macro level through procurement rules, funding programmes, certification requirements, industrial policy, or cloud strategies. Their intended effect, however, appears further down the chain when organisations make different procurement and architecture decisions. Evaluating a policy solely by observing a later change in a national index therefore amounts to testing a joint hypothesis: that the intervention changed the relevant organisational decisions, that those decisions changed dependency structures, and that their aggregation was strong enough to affect the macro indicator. Everything else that changes over the same period – economic growth, investment cycles, provider strategies, geopolitical events, or general technological development – enters the observed macro trend as additional noise. A rising index score may therefore coincide with an effective policy, but on its own it does not identify the mechanism through which that policy worked.

A more convincing evaluation would observe the mechanism directly. This means moving measurement to the meso level and asking whether the decisions targeted by an intervention actually changed. Difference-in-differences designs could compare sectoral dependency indicators around a specific policy measure, with affected and unaffected sectors serving as treatment and comparison groups. Cohorts of procurement decisions could be tracked before and after a change in procurement rules to examine whether provider selection, exit provisions, or dependency profiles changed. Policy evaluations could also pre-specify the meso-level outcomes that a particular intervention is expected to affect rather than

selecting favourable macro indicators after the fact. This is especially important because programme evaluation may otherwise drift toward measuring activity – funds committed, projects launched, certifications issued – rather than the dependency structures the programme was meant to change. The decisive test therefore lies in implementation rather than ranking: a sovereignty policy that cannot show changed implementations has not demonstrated its intended effect, however positively the national score develops; a policy that can show such changes may have worked even if it takes years for them to appear in macro statistics.

5.3 For the Theory of Digital Dependency

The multi-level perspective also connects measurement more closely to theories of digital dependency. Farrell and Newman's (2019) account of weaponized interdependence is formulated at the level of international political economy, but its mechanism is inherently concrete. A chokepoint is always located somewhere: in a platform, a clearing system, a network node, an infrastructure component, or another arrangement whose position makes substitution difficult and control valuable. The compilation perspective translates this insight into the language of measurement. If coercive potential resides in particular configurations and critical points, then measures based mainly on averages risk smoothing away exactly the structures that the theory treats as decisive. At the same time, the level discipline of R1 protects the theory from stretching its own vocabulary too far. States are not simply large organisations, just as organisations are not small states. Similar dependency mechanisms may recur across levels without the units themselves becoming equivalent. For a concept such as sovereignty, which moves so easily between international politics, organisational strategy, and technical architecture, Sartori's (1970) ladder of abstraction therefore remains an ongoing rather than a one-off check.

5.4 Limitations of the Argument

Three limitations matter. First, the argument remains primarily conceptual. The EU comparison and calibrated simulation show that the aggregation choice matters, but they do not establish that any particular country or organisation is currently misclassified. Likewise, the proposed importance of within-country variation and the possibility of sign reversal remain empirical hypotheses. Testing them would require exactly the meso-level data whose absence motivates the argument in the first place. There is an unavoidable circularity here: the article identifies a missing observation level and then argues that some of its central claims can only be resolved once that level is observed. Second, the compilation argument is strongest for the dependency core of digital sovereignty. Other components, such as skills or general digital capacity, may behave more like composition constructs and may legitimately permit averaging. A complete measurement model may therefore need to be mixed, with the boundary between composition and compilation specified component by component rather than imposed on the construct as a whole. Third, the argument rests on a largely liberal European understanding of digital sovereignty as self-determined action under manageable dependencies. Alternative conceptions that place territorial control or exclusive jurisdiction at the centre would define both the construct and its cross-level relationships differently. The multi-level framework could still be applied to such conceptions, but its

substantive mapping and aggregation rules would need to be re-derived rather than carried over unchanged.

6. Conclusion

Where, then, does digital sovereignty live? It cannot be located in national rankings alone, nor can it be reduced to isolated technical components in individual server rooms. It emerges in the largely unmeasured space between them: in the way national framework conditions shape organisational decisions, in the dependency structures created by those decisions, and in the way those structures condense into broader sectoral and national patterns. Taking that space seriously requires five changes in measurement practice: clearer statements about the level at which an instrument observes and supports inferences; aggregation rules that preserve critical minima and relevant configuration where the construct is compilation-like; empirical testing of construct continuity across levels rather than simply assuming it; systematic observation of the meso level connecting policy frameworks to installations; and confounder control matched to the level at which claims are made. These requirements constrain some of the claims that existing measurement instruments can support, but they also make those claims more defensible. A concept important enough to shape European technology policy deserves to be measured at the levels where its dependencies actually arise, with equal clarity about the levels on which the available evidence does not support claims.

Declarations

Author Contribution

M.K. conceived the study, developed the theoretical framework, performed the analyses, prepared all tables and figures, and wrote and reviewed the manuscript.

References

1. Alaimo, L.S., Seri, E.: Measuring human development through composite indicators: Open issues and new methodological tools. *Qual. Quant.* **58**(6), 5275–5307 (2024)
2. Arndt, C.: The politics of governance ratings. *Int. Public. Manage. J.* **11**(3), 275–297 (2008)
3. Coleman, J.S.: *Foundations of social theory*. Belknap Press of Harvard University, Cambridge, MA (1990)
4. Couture, S., Toupin, S.: What does the notion of 'sovereignty' mean when referring to the digital? *New. Media Soc.* **21**(10), 2305–2322 (2019)
5. Diamantopoulos, A., Winklhofer, H.M.: Index construction with formative indicators: An alternative to scale development. *J. Mark. Res.* **38**(2), 269–277 (2001)
6. Farrell, H., Newman, A.L.: Weaponized interdependence: How global economic networks shape state coercion. *Int. Secur.* **44**(1), 42–79 (2019)

7. Kaloudis, M.: Digital sovereignty – European Union's action plan needs a common understanding to succeed. *History Compass*. **19**(12), e12698 (2021)
8. Kaloudis, M.: Sovereignty in the digital age – How can we measure digital sovereignty and support the EU's action. plan? *New. Global Stud.* **16**(3), 275–299 (2022a)
9. Kaloudis, M.: From quality to quantity: How can digital sovereignty be parsed into measurable components? *Eur. J. Bus. Sci. Technol.* **8**(2), 172–189 (2022b)
10. Kaloudis, M.: Digital sovereignty as a weapon of diplomacy in cyber warfare in democracies. In: *National security in the digital and information age*. IntechOpen, London (2024)
11. Hardy, G.H., Littlewood, J.E., Pólya, G.: *Inequalities*. Cambridge University Press, Cambridge (1934)
12. Kozlowski, S.W.J., Klein, K.J.: A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In: Klein, K.J., Kozlowski, S.W.J. (eds.) *Multilevel theory, research, and methods in organizations*, pp. 3–90. Jossey-Bass, San Francisco, CA (2000)
13. Kremer, M.: The O-ring theory of economic development. *Quart. J. Econ.* **108**(3), 551–575 (1993)
14. Madiaga, T.: *Digital sovereignty for Europe*. EPRS Ideas Paper. European Parliamentary Research Service, Brussels (2020)
15. Mazziotta, M., Pareto, A.: Methods for constructing composite indices: One for all or all for one? *Rivista Italiana di Economia. Demografia e Statistica*. **67**(2), 67–80 (2013)
16. Meredith, W.: Measurement invariance, factor analysis and factorial invariance. *Psychometrika*. **58**(4), 525–543 (1993)
17. Mueller, M.L.: Against sovereignty in cyberspace. *Int. Stud. Rev.* **22**(4), 779–801 (2020)
18. Munda, G.: *Social multi-criteria evaluation for a sustainable economy*. Springer, Berlin (2008)
19. OECD/JRC: *Handbook on constructing composite indicators: Methodology and user guide*. OECD Publishing, Paris (2008)
20. Openshaw, S.: The modifiable areal unit problem. *Concepts and Techniques in Modern Geography* No. 38. Geo Books, Norwich (1984)
21. Pohle, J., Thiel, T.: Digital sovereignty. *Internet Policy Rev.*, **9**(4). (2020)
22. Robinson, W.S.: Ecological correlations and the behavior of individuals. *Am. Sociol. Rev.* **15**(3), 351–357 (1950)
23. Sartori, G.: Concept misformation in comparative politics. *Am. Polit. Sci. Rev.* **64**(4), 1033–1053 (1970)
24. Schmitz, L., Seidl, T.: As open as possible, as autonomous as necessary: Understanding the rise of open strategic autonomy in EU trade policy. *J. Common Mark. Stud.* **61**(3), 834–852 (2023)

Figures

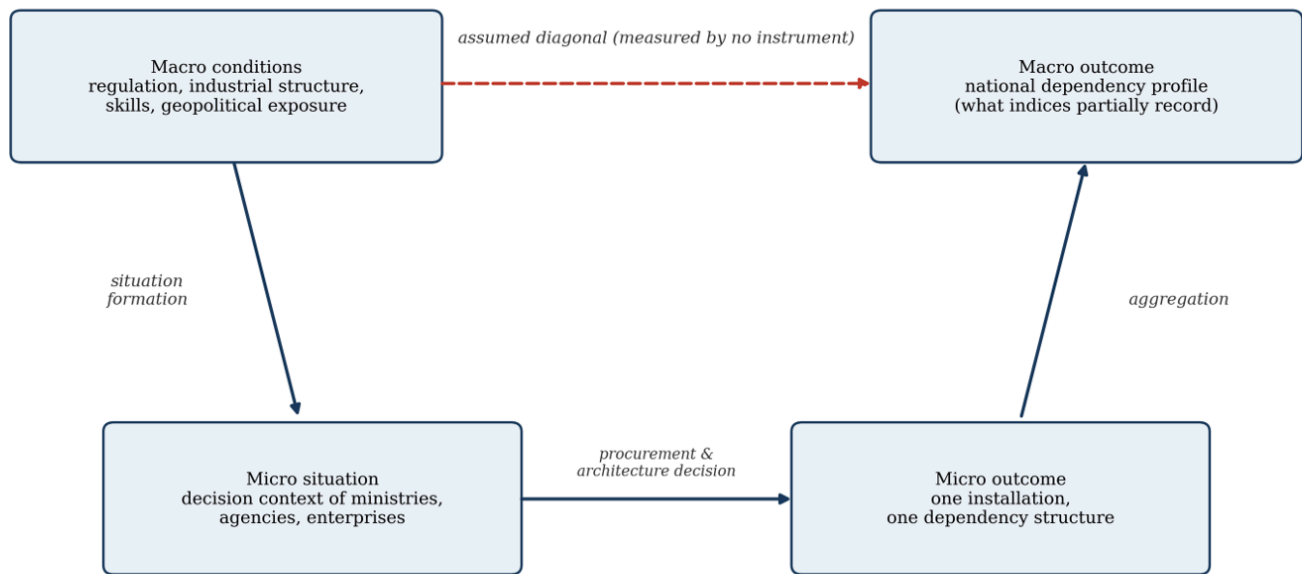


Figure 1. The Coleman schema applied to digital sovereignty. Country indices capture the two macro boxes; the descending, horizontal, and ascending arrows – where the construct is substantively produced – are not observed, and the macro-to-macro diagonal (dashed) is assumed rather than measured.

Figure 1

See image above for figure legend.

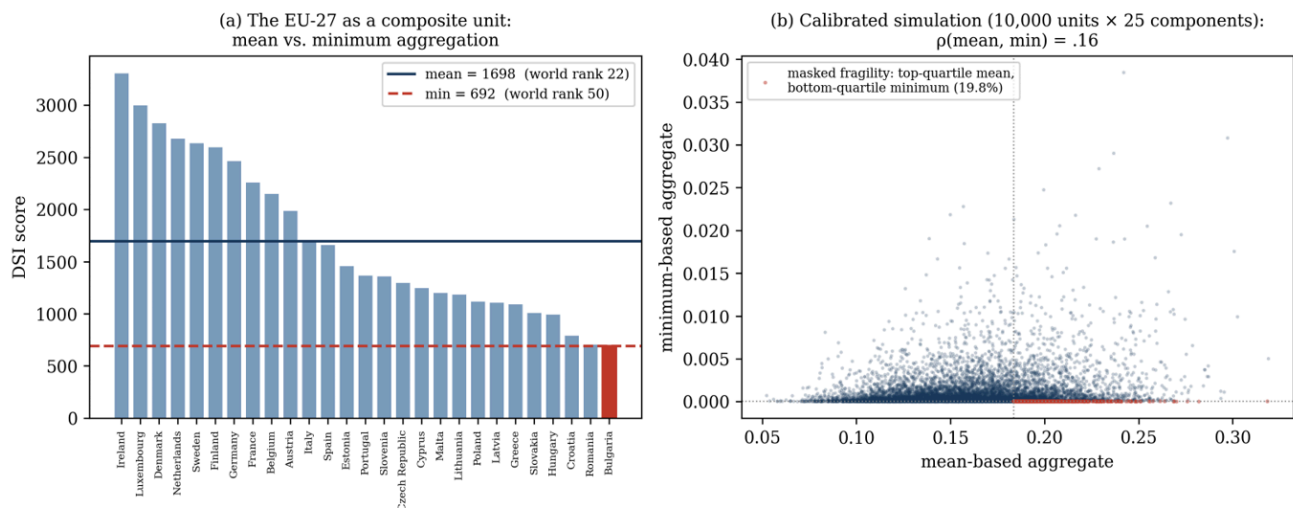


Figure 2. The aggregation choice quantified. (a) The EU-27 as a composite unit based on published DSI data: member scores (bars; weakest member highlighted), with mean- and minimum-based unit scores and their implied world ranks. (b) Calibrated simulation of 10,000 synthetic units: mean-based versus minimum-based aggregates; highlighted points are cases of masked fragility – top quartile by mean, bottom quartile by critical minimum.

Figure 2

See image above for figure legend.