

SUPPLY CHAIN RESEARCH · EDITORIAL FEATURE

Compliance Theatre

The Sustainability-Reporting Software Scramble and the Deadlines That Moved

A wave of regulation drove enterprises to buy carbon-accounting and ESG-reporting software under deadline pressure. Then the deadlines moved, the scope narrowed, and some rules collapsed entirely, leaving buyers with tools for cancelled requirements and data too weak to decide on but strong enough to be sued over.

Independent, vendor-neutral research for supply chain and technology leaders

Supply Chain Research | supplychainresearch.com | 2026

Contents

The Sustainability-Reporting Software Scramble and the Deadlines That Moved

01	Executive summary	01
02	The compliance-theatre purchase	02
03	The forcing function that moved.....	03
04	The scope that narrowed	04
05	The rules that collapsed	05
06	The Scope 3 data problem	06
07	Spend-based estimation and its perverse incentive	07
08	The assurance gap	08
09	Greenwashing enforcement as the downside	09
10	The vendor market and its unmeasurable size	10
11	The fairness case: measurement must start somewhere	11
12	Buy for decision-usefulness, not compliance.....	12
13	A procurement protocol, and a scoring rubric.....	13
14	Governing who owns the numbers.....	14
15	Conclusion: the tool is not the disclosure	15
16	Methodology, caveats, and sources.....	16

01 Executive summary

Between 2022 and 2024 a wave of enterprises bought sustainability-reporting and carbon-accounting software, driven by a set of regulations that appeared to make detailed disclosure mandatory on a fixed timeline. Then the timeline moved. The European Union postponed the second and third waves of its flagship directive by two years and narrowed its scope so far that an estimated four in five previously-covered companies fell out of it. The United States securities regulator adopted a climate-disclosure rule, stayed it, abandoned its defense, and proposed to rescind it. A European directive against misleading green claims was withdrawn in all but formality. Buyers who had purchased tools for a 2026 first report found the requirement deferred to 2028, narrowed, or gone.

This article is about what that scramble reveals. Much sustainability-software purchasing was compliance theatre: buying a tool to produce a report rather than to reduce emissions, on a deadline that then proved unstable. The problem runs deeper than wasted licenses, because the data underneath the reports is frequently too weak to support any decision. The majority of most companies' emissions sit in Scope 3, the value chain, which is dominated by spend-based estimation, dollars multiplied by an industry-average emission factor, where the factors vary by three to five times across databases and the method rewards spending less rather than emitting less. We say honestly that the regulation is legitimate and that measurement must start somewhere. But a buyer should understand that they are frequently purchasing a tool to generate a number that cannot guide a decision, on a deadline that may move again, in a domain where a weak disclosure is not merely useless but a live legal exposure, because greenwashing enforcement is real, recent, and expensive.

~80%

of previously-covered companies estimated to fall out of EU CSRD scope after the Omnibus

11.4x

the average company's Scope 3 emissions relative to its combined Scope 1 and 2 (CDP 2024)

3-5x

the variation in emission factors for similar categories across databases

The argument in brief

- **The forcing function moved after the software was bought.** The EU delayed waves two and three by two years and narrowed scope; the SEC rule was adopted then abandoned; the Green Claims Directive was effectively dropped. Tools were bought for requirements that changed.
- **The data cannot support a decision.** Scope 3 dominates the footprint and is mostly spend-based estimation. Factors vary three to five fold, and improving data quality usually raises reported emissions rather than lowering them.
- **A weak disclosure is a legal exposure.** Recent greenwashing enforcement against Keurig, WisdomTree, Invesco, and DWS shows that inaccurate sustainability claims draw fines under ordinary anti-fraud rules, even after the SEC disbanded its ESG task force.

- **Assurance is thinner than buyers assume.** Most sustainability assurance is limited, a negative-form opinion, not the reasonable assurance investors imagine, and reasonable assurance is rarely offered on Scope 3 estimates at all.
- **Buy for decision-usefulness, not for a document.** Demand methodology transparency, align to the regime that survived the Omnibus, know the assurance level, and govern who owns the numbers.

02 The compliance-theatre purchase

A compliance-theatre purchase is one made to satisfy a rule rather than to accomplish the thing the rule was meant to encourage. In sustainability software the distinction is unusually sharp. The purpose of emissions regulation is to cause companies to reduce emissions. The purpose of much emissions software, as actually bought and used, is to produce a report that satisfies a disclosure requirement. Those two purposes overlap in principle and diverge sharply in practice, because a tool optimized to generate a compliant document is not necessarily a tool that helps anyone decide what to change, and a company under deadline pressure to disclose will buy the former and call it the latter.

The tell is what happens to the output. A sustainability report that drives decisions changes procurement, supplier selection, product design, or logistics, and the organization can point to the decisions it changed. A compliance-theatre report is produced, filed, published on the corporate website, and acted upon by no one, because its numbers are too uncertain to justify a decision and its purpose was never to inform one. The software did its job, which was to generate the document. The document did its job, which was to exist. And the emissions the whole exercise was meant to address were neither measured accurately enough to target nor reduced as a result, because targeting and reducing were not what the tool was bought to do.

This is not a claim that the buyers are cynical or that the vendors are fraudulent. Both are usually responding rationally to a real legal requirement on a real deadline. A company facing a disclosure mandate must produce a disclosure, and buying software to help is sensible. The theatre arises from the gap between what the regulation appears to demand, which is meaningful measurement and reduction, and what a company under deadline can actually deliver, which is a document assembled from estimates. The tool bridges that gap by producing something that looks like the former while being the latter, and everyone involved can plainly say they have done what the rule required, because the rule, in the end, required a report.

The scramble of 2025 and 2026 exposed the theatre because it removed the deadline that had justified it. When the forcing function moved, the companies that had bought tools to reduce emissions kept using them, because the tools were doing useful work regardless of the regulation. The companies that had bought tools to satisfy a deadline found themselves with software for a requirement that had been deferred or repealed, and no independent reason to keep using it, which revealed that the deadline, not any internal purpose, had been the reason for the purchase all along. Figure 1 shows how sharply the forcing function moved.

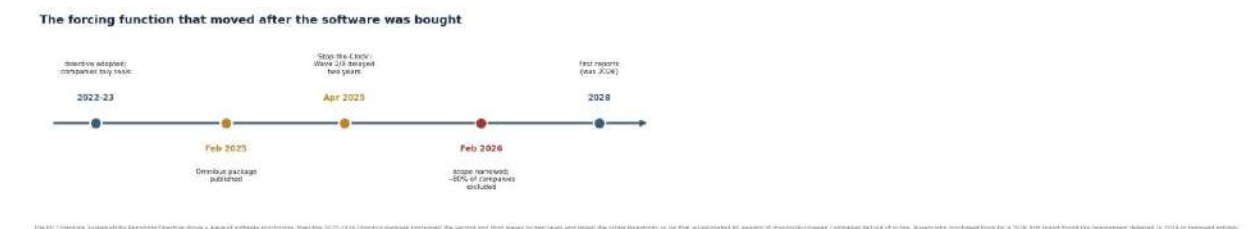


Figure 1. The EU forcing function drove a software-buying wave, then moved: waves two and three delayed two years, scope narrowed, first reports pushed from 2026 to 2028.

It is worth being precise about why the theatre is so hard to detect from inside an organization, because the people involved are usually acting in good faith and would be surprised to hear their work described this way. The sustainability team producing the report is truly trying to comply, the finance team funding the software is responding to a real regulatory signal, and the executives approving it are managing a real risk. No one in the chain has decided to produce a document that will be ignored. The theatre is emergent rather than intended: it arises from each participant optimizing for the deadline in front of them, and the deadline rewards producing a compliant document rather than producing a decision-grade understanding. The result is theatre even though no one set out to stage it, which is exactly why it persists, because there is no villain to remove and no single decision to reverse.

The financial scale of the theatre is easy to underestimate, because the license fee is only the visible part. A sustainability-reporting deployment consumes internal effort across finance, procurement, operations, and sustainability functions; it requires data collection from across the business and often from suppliers; it demands ongoing maintenance as the regime and the standards evolve; and it frequently pulls in consultants to configure and validate it. The fully-loaded cost of producing a sustainability disclosure runs to many multiples of the software price, and when the disclosure is theatre, all of that cost is spent producing a document that changes nothing. A buyer weighing the cost of a sustainability tool against its benefit should weigh the fully-loaded cost of the capability, not the license, and should ask whether the benefit, in decisions actually changed, justifies it, because for a theatre purchase the honest answer is that it does not.

The phenomenon has a precise analogue in an older domain that clarifies what is happening. When financial reporting became mandatory, some companies produced accounts to satisfy the requirement rather than to understand their own finances, and the accounts were technically compliant while conveying little that management acted upon. The difference is that financial accounting matured over a century into a discipline with settled standards, rigorous assurance, and numbers reliable enough to run a business on, whereas sustainability accounting is at an early and unsettled stage, with contested standards, thin assurance, and numbers that cannot yet bear the weight placed on them. A buyer of sustainability software is buying into a discipline at the stage financial accounting occupied long ago, before its numbers were trustworthy, and should calibrate their expectations accordingly rather than assuming that a sustainability figure carries the reliability that a mature financial figure does.

This immaturity is not a permanent condition, and it is worth saying that sustainability measurement will likely improve, as primary data replaces estimates, as standards settle, and as assurance deepens. A buyer today is operating in the early, unreliable phase of a discipline that may eventually become as rigorous as financial accounting, and the compliance-theatre problem is partly a symptom of that immaturity, of a requirement to report having arrived before the capacity to measure reliably. But a buyer must operate in the present, where the immaturity is real and the numbers are weak, and the fact that measurement may improve in the future does not make today's spend-based estimates decision-grade today. The prudent posture is to invest in the capability to measure and improve, which will compound as the discipline matures, while treating the current outputs with the skepticism their current reliability warrants, rather than either dismissing the enterprise as permanently hopeless or crediting today's numbers with a reliability they do not yet possess.

03 The forcing function that moved

The central regulatory event is the European Union's treatment of the Corporate Sustainability Reporting Directive, which drove more sustainability-software purchasing than any other single rule and which was then substantially rewritten under a simplification package known as the Omnibus. The sequence matters, because a buyer's tool was frequently selected against one version of the regime and must now serve a materially different one.

The directive, as originally enacted, phased in reporting obligations across waves of companies on a fixed schedule, with large public-interest entities reporting first and progressively smaller companies following. Enterprises across and beyond the European Union bought software to prepare, because the obligations were detailed, the standards were extensive, and the timelines appeared firm. Then, in February 2025, the European Commission published the Omnibus simplification package, and in April 2025 the co-legislators adopted a directive, commonly called Stop-the-Clock, that postponed the reporting requirements for the second and third waves by two years and delayed the first phase of the related due-diligence directive by a year. The substantive content changes were agreed later in 2025 and formally adopted in early 2026, applying to financial years beginning on or after the first of January 2027, with first reports in 2028.

Consider the position this creates for a buyer. A company in the second wave that had purchased and configured a reporting tool for a first report due in 2026 now faces that report in 2028, if at all, because the same package that delayed the timeline also narrowed the scope in ways described in the next section. The tool was bought against a deadline that has moved two years and a scope that may no longer include the company. The software is not defective, and the vendor did nothing wrong, but the purchase was justified by a regulatory certainty that has evaporated, and the company must now decide whether the tool serves any purpose absent the deadline that prompted it.

A further complication is that the first wave was treated differently from the later ones. The companies already reporting were not covered by the Stop-the-Clock delay, and the Commission issued separate quick-fix relief for them. So the regime fragmented not only across jurisdictions but across waves within the same directive, with first-wave companies proceeding under one set of adjustments, second and third-wave companies deferred and rescope, and the whole edifice subject to a simplification whose final contours were still settling as national governments worked to transpose it. A buyer trying to align a tool to the applicable requirement faced not a stable target but a moving and fragmenting one, which is precisely the condition under which compliance-theatre purchasing flourishes, because the deadline pressure is real even as the requirement itself is unstable.

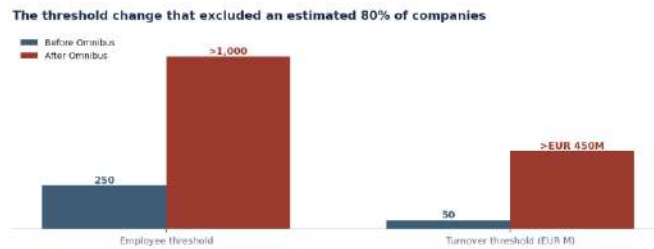
The transposition of the directive into national law added a further layer of instability that a buyer had to navigate, because a European directive is not directly binding until member states enact it into their own law, and the enactment was uneven and delayed. A substantial number of member states missed the original transposition deadline, and as the Omnibus rewrote the directive at the European level, national governments were left transposing a moving target, with some having already enacted the original version, others waiting, and others enacting partial or divergent versions. A company operating

across several European jurisdictions therefore faced not one regime but several, at different stages of enactment, all subject to a simplification whose final form was still settling. The tool bought to satisfy the directive had to satisfy whichever national implementation actually applied, and those implementations diverged.

This national fragmentation compounds the compliance-theatre problem, because it multiplies the ways in which a tool bought against an expected requirement can miss the requirement that actually applies. A company that configured its software against the directive as originally drafted, expecting a particular national implementation, might find that its jurisdiction enacted a delayed or modified version, or had not yet enacted at all, so that the obligation the tool was configured for did not exist in the form anticipated. The buyer is left maintaining a capability tuned to a requirement that the relevant national law may never impose in that form, which is the compliance-theatre problem refracted through the complexity of European legislative implementation, and it is a reason to be especially wary of tooling tightly to a specific expected regime before that regime is actually enacted where the company operates.

04 The scope that narrowed

The Omnibus did not merely delay the directive; it narrowed who the directive applies to, and the narrowing was dramatic enough to remove a large majority of previously-covered companies from the regime entirely. This is the single most consequential fact for a buyer, because a tool bought to satisfy an obligation is worthless if the obligation no longer applies. Figure 2 shows the change.



THE NARROWING OF SCOPE FROM A TWO-OF-THREE TEST TO MORE THAN 1,000 EMPLOYEES AND MORE THAN 450 MILLION EURO IN TURNOVER, ESTIMATED TO EXCLUDE 80 PERCENT OF PREVIOUSLY COVERED COMPANIES

Figure 2. The Omnibus raised the scope thresholds from a two-of-three test to more than 1,000 employees and more than 450 million euro turnover, estimated to exclude 80 percent of previously-covered companies.

Under the original regime, a company fell within scope if it met two of three criteria: more than 250 employees, more than 50 million euro in turnover, or more than 25 million euro on the balance sheet. The Omnibus replaced that test with a far higher bar: more than 1,000 employees and more than 450 million euro in turnover, applied together rather than as a two-of-three test. The effect of raising the thresholds this far, according to estimates cited in the legislative process and by law firms tracking it, is to exclude roughly 80 percent of the companies that the original directive would have covered. Four in five previously-in-scope companies, in other words, are no longer required to report at all.

For the buyers among that 80 percent, the software they purchased was bought to satisfy an obligation they no longer have. This is the starkest possible illustration of the compliance-theatre problem, because it removes not just the deadline but the requirement. A mid-sized company that bought a reporting suite to prepare for an obligation it expected to face now faces no such obligation, and the only reasons to continue with the tool are internal ones, that the measurement is useful for the company's own decisions, or external ones unrelated to this directive, such as customer demands or other jurisdictions' rules. If neither applies, the purchase was pure theatre, justified entirely by a requirement that has been legislated away.

The companies that remain in scope, the largest enterprises above the new thresholds, face a different but related problem. They are still required to report, but the standards themselves are being simplified, and the precise content of what must be disclosed was still being settled as this was written. A tool configured against the original, more extensive standards may now be over-specified, gathering and processing data that the simplified regime no longer requires, which is a milder form of the same mismatch. The requirement did not disappear for these companies, but it changed shape, and a tool bought against its original shape must be reconfigured or partly abandoned. In both cases, the buyer is managing the consequences of having purchased against a regime that was rewritten after the purchase.

The rescoping also created a quieter problem for the companies that remained within the directive, concerning the standards rather than the thresholds. The European Sustainability Reporting Standards that specify what in-scope companies must actually disclose were themselves subject to simplification under the Omnibus, with the number of mandatory data points reduced and certain requirements softened. A company that had configured its reporting tool against the original, more extensive standards, gathering and processing every data point they specified, found that some of that data collection was no longer required, which is a milder mismatch than falling out of scope but a mismatch nonetheless. The tool was over-specified relative to the simplified requirement, and the company had built data-collection processes for disclosures it no longer had to make, which is wasted effort even though the company remained in scope for the directive as a whole.

This interaction between the threshold changes and the standards changes illustrates a general feature of tooling to a specific regulatory requirement, which is that the tool must track not one parameter but several, any of which can move independently. The scope thresholds determine whether the company must report at all; the standards determine what it must report; the timeline determines when; and the assurance requirements determine to what level of reliability. The Omnibus moved several of these at once, and a tool configured against the original settings of all of them was misaligned on multiple dimensions simultaneously. A buyer tooling to a regulatory requirement is therefore exposed to movement on every one of these parameters, which is a larger attack surface for regulatory instability than the buyer usually appreciates, and another reason to prefer tooling to the company's durable decisions rather than to the regulation's movable parameters.

05 The rules that collapsed

Beyond the European directive, several other forcing functions that had driven sustainability-related purchasing collapsed, retreated, or were abandoned over the same period, and the pattern across them reinforces the central point: a company that standardized on a tool for a specific regime frequently found that regime gone. Figure 3 surveys the landscape.

A patchwork of forcing functions, several of which collapsed

EU CSRD / ESRS	named and delayed by Omnibus; first reports 2028
EU Green Claims Directive	withdrawn in intent; effectively dead; not formally rescinded
EU EmpCo Directive	in force; bans generic green claims; up to 4% of turnover
US SEC climate rule	adopted 2026, stayed, defense dropped 2025, rescission proposed
California SB 253	survives; first Scope 1+2 report due Aug 2026
California SB 261	enforcement enjoined pending First Amendment challenge

The regulatory landscape that drove sustainability-related purchasing is now fragmented and, in several jurisdictions, retreating. A few who standardized on a tool for one regime frequently find that the specific requirement it was bought to comply has rescinded, collapsed, or unenforced, while a different and binding rule, such as the EU ban on generic green claims, remains in force and under-served.

Figure 3. A fragmented and partly retreating landscape. Several forcing functions that drove purchasing collapsed, while a binding rule against generic green claims survives and is under-served.

In the United States, the Securities and Exchange Commission proposed a climate-disclosure rule in 2022 and adopted it in March 2024, prompting companies to prepare and, in some cases, to buy tools. The Commission then stayed the rule in April 2024 pending litigation, voted in March 2025 to end its defense of the rule, and in 2026 proposed to rescind it. The rule never took effect. A company that bought software to comply with a federal climate-disclosure mandate discovered that the mandate was adopted, frozen, and then abandoned in the space of two years, and that the operative baseline reverted to much older and lighter guidance. This is the same whipsaw as the European one, executed by a different mechanism: not legislative simplification but a change of administration and a decision to stop defending a rule in court.

In Europe, a proposed directive against misleading environmental claims, the Green Claims Directive, was withdrawn in intent in mid-2025 when the Commission signaled its abandonment and the final negotiation was cancelled. Its status is truly ambiguous, and it was not formally rescinded through the ordinary process, so a precise account must say that it is effectively dead rather than definitively buried. A company that anticipated a substantiation regime for green marketing claims, and bought or planned tooling accordingly, found that regime collapse before it took effect.

The instructive counterpoint is what survived, because it is binding and comparatively under-served. A separate European directive, on empowering consumers for the green transition, entered into force and remains on track, banning generic environmental claims such as climate-friendly or carbon-neutral without on-medium substantiation and threatening penalties up to a share of turnover. And in the United States, while the federal rule collapsed, California's disclosure laws survived, with the first reports on operational emissions due in 2026 for the largest companies, one of the two laws having its enforcement temporarily enjoined pending a constitutional challenge. So the landscape did not simply empty out; it rearranged, and the rules that survived are frequently not the ones companies tooled up for. A buyer's task is now to serve the regime that actually binds them, which for many is a different

regime from the one that prompted their purchase, and for some is a rule they have not yet addressed because the collapsed rules absorbed all the attention.

The pattern across these collapses is worth abstracting, because it reveals something about the reliability of regulation as a foundation for purchasing that extends beyond sustainability. In each case, a regulation that appeared firm enough to justify significant expenditure proved subject to reversal by a mechanism outside the buyer's view: legislative simplification in Europe, a change of administration and litigation posture in the United States, an abandoned negotiation for the green-claims rule. None of these reversals was predictable from the text of the regulation itself, which in each case appeared to impose a firm requirement, and a buyer reading the regulation at the time of purchase would have had no way to anticipate the collapse. The lesson is that the apparent firmness of a regulation is a poor guide to its durability, because the forces that reverse regulations operate independently of the regulatory text and frequently without warning.

This has a direct implication for how a buyer should weight a regulatory requirement in a purchasing decision, which is to discount it for the possibility of reversal even when it appears firm. A requirement that has just been enacted, or that is newly in force, carries more reversal risk than its firm appearance suggests, because it has not yet survived the political and legal contestation that tests whether a regulation endures. A buyer who treats a newly-enacted requirement as a permanent fixture, and tools heavily to it, is exposed to exactly the reversal risk that these cases illustrate. A more prudent buyer treats a new regulatory requirement as provisional until it has proven durable, tools lightly to it in the interim, and invests in capabilities that serve the company's own decisions and therefore retain value regardless of whether the regulation survives. The collapses of 2025 and 2026 are an expensive demonstration of why that prudence is warranted.

06 The Scope 3 data problem

Set the regulatory instability aside and suppose a company truly wants to measure its emissions to reduce them. It confronts a second problem, independent of the first and arguably deeper: the majority of its emissions are the part it can measure least well. This is the Scope 3 problem, and it is where the claim that the data cannot support a decision is grounded. Figure 4 shows the dominance of Scope 3.

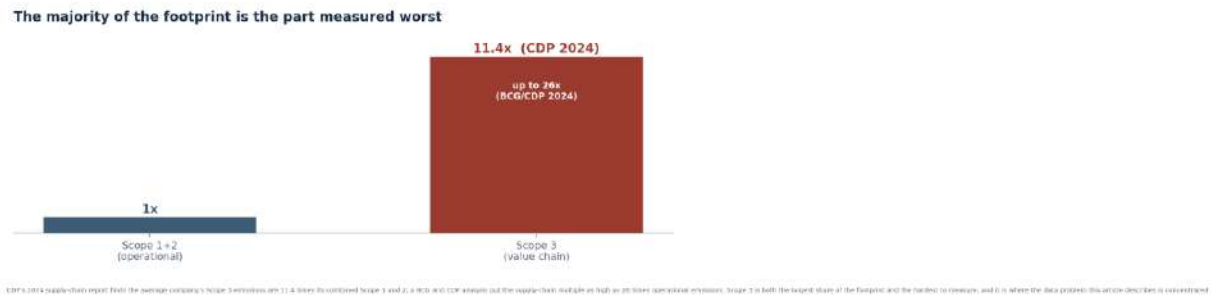


Figure 4. Scope 3 dominates the footprint. CDP's 2024 data puts it at 11.4 times operational emissions on average, and a BCG and CDP analysis as high as 26 times.

Emissions are conventionally divided into three scopes. Scope 1 is direct emissions from sources the company owns or controls. Scope 2 is indirect emissions from purchased energy. Scope 3 is everything else in the value chain, upstream and downstream, from purchased goods and services through the use and disposal of sold products. For most companies Scope 3 is by far the largest category. CDP's 2024 supply-chain report finds the average reporting company's Scope 3 emissions to be 11.4 times its combined Scope 1 and 2, and a BCG and CDP analysis put the supply-chain multiple as high as 26 times operational emissions. Whatever the exact multiple, the conclusion is the same: the operational emissions a company can measure directly are a small fraction of its total, and the total is dominated by a category the company cannot measure directly at all.

The reason Scope 3 cannot be measured directly is that it occurs in other organizations. A company's upstream Scope 3 emissions are the operational emissions of its suppliers, and its downstream emissions arise from how customers use and dispose of its products. The company does not control these activities and usually cannot observe them, so it must either obtain data from its suppliers and customers, which most cannot do at scale, or estimate them. Only about a third of companies obtain primary data from their suppliers; the rest estimate, and the dominant estimation method, examined in the next section, is weak enough that the resulting Scope 3 figure is frequently uncertain by a wide margin.

This creates a structural absurdity at the heart of much sustainability reporting. The number that dominates the footprint, and therefore the number that any serious reduction effort must target, is the number measured worst. A company that reports its emissions with apparent precision is reporting a small, well-measured operational figure alongside a large, poorly-estimated value-chain figure, and presenting their sum as though it carried uniform reliability. A decision based on that sum, or on changes in it, is a decision based mostly on the poorly-estimated part, because the poorly-estimated part is most of the total. The precision of the operational measurement lends a borrowed and

undeserved credibility to the value-chain estimate, and the combined figure looks far more decision-grade than the underlying data can support.

The three-scope framework itself deserves a moment of scrutiny, because its familiarity conceals how much of the difficulty it creates. The framework is a sensible accounting construct, dividing emissions by the degree of a company's control, but it has the effect of placing the largest and most important category, Scope 3, in the position of greatest measurement difficulty, which means that the accounting framework and the measurement difficulty are aligned in the worst possible way. If the categories that were easiest to measure were also the largest, sustainability accounting would be far more tractable. Instead, the framework guarantees that the bulk of the footprint sits in the category the company controls least and can observe least, which is not a flaw in the framework so much as a reflection of the underlying reality that most of a company's emissions truly do occur outside its own operations.

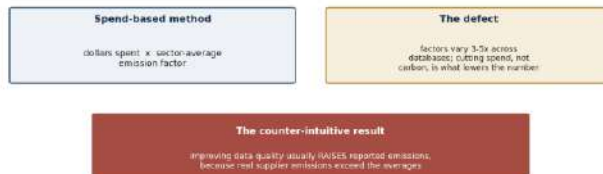
This alignment of size and difficulty has a strategic implication that companies frequently miss. Because Scope 3 dominates and is measured worst, the greatest opportunities for actual emissions reduction usually lie in Scope 3, in the value chain, precisely where the company has the least visibility and the least control. A company that focuses its reduction efforts on its well-measured operational emissions, because those are the ones it can see and act on directly, is optimizing the small part of the footprint while the large part remains both unmeasured and unaddressed. Serious decarbonization requires engaging the value chain, which requires better Scope 3 data than spend-based estimation provides, which is why the data problem is not merely an accounting inconvenience but the central obstacle to the reduction the whole exercise is meant to achieve.

A concrete illustration makes the dominance of Scope 3 vivid. Consider a consumer-goods company whose own factories and offices emit a certain amount, its Scope 1 and 2. The emissions embedded in the raw materials it purchases, the packaging it buys, the transport of goods to and from its facilities, the retail that sells its products, and the use and disposal of those products by consumers, all of which are Scope 3, will typically dwarf its operational emissions many times over. The company could eliminate its operational emissions entirely, running its factories on renewable energy and its offices carbon-free, and still have addressed only a small fraction of its total footprint, because the bulk of that footprint was never in its operations to begin with. This is not unusual; it is the normal structure of a modern company's emissions, and it is why any serious engagement with corporate emissions must engage the value chain, which is exactly the part the data cannot measure well.

07 Spend-based estimation and its perverse incentive

The method by which most companies estimate the Scope 3 emissions they cannot measure deserves specific attention, because understanding it dissolves most of the credibility of the resulting numbers and reveals a perverse incentive at the center of the exercise. The method is spend-based estimation, and it is exactly what its name suggests. Figure 5 sets out the mechanism and its defect.

Carbon accounting that is financial data times a generic coefficient



Because most companies cannot obtain primary supplier data, Scope 3 is typically estimated by multiplying spend by an industry-average emission factor. Those factors vary by three to five times across databases, with the method rewards spending less rather than spending less. It is a reference starting from sanctioned by the Greenhouse Gas Protocol, and an emissions factor for a database.

Figure 5. Spend-based estimation multiplies dollars by a sector-average factor. Factors vary three to five fold, the method rewards spending less rather than emitting less, and better data usually raises reported emissions.

Spend-based estimation works by taking the amount of money a company spends in a category, multiplying it by an emission factor that represents the average emissions per unit of spending in that sector, and treating the product as the emissions associated with that spending. The factors come from environmentally-extended input-output databases that estimate, for an economy or a sector, how much greenhouse gas is emitted per unit of economic activity. The method has the great advantage of being possible: a company knows what it spends, and factors exist for most categories, so an estimate can always be produced. It is sanctioned by the Greenhouse Gas Protocol as a legitimate approach, particularly as a starting point where primary data is unavailable.

It has two defects that together render its outputs close to useless for decisions. The first is that the factors are averages of enormous ranges and vary substantially across the databases that produce them, by three to five times for similar categories according to academic analysis. Two companies with identical spending, or the same company using two different databases, can arrive at Scope 3 figures that differ by a multiple, purely from the choice of factor. A number that can vary several fold based on which reference database an analyst happens to use is not a measurement in any meaningful sense; it is an order-of-magnitude indication dressed as a figure.

The second defect is more insidious, because it inverts the incentive the whole exercise is meant to create. Under a spend-based method, a company's reported emissions in a category are proportional to its spending in that category. This means the way to reduce reported emissions is to spend less, not to emit less. A supplier that truly halves its emissions intensity, through renewable energy or process improvement, produces no change whatsoever in its customer's spend-based Scope 3 figure, because the customer is still applying a sector average to the same spending. Conversely, a company can reduce its reported emissions by switching to a cheaper supplier that happens to be dirtier, because the lower spending lowers the calculated figure regardless of the actual emissions. The method rewards the wrong behavior and is blind to the right one, which means that reported progress under a spend-based method can be entirely disconnected from, or even opposed to, actual decarbonization.

There is a final consequence that a buyer must be prepared for, because it routinely surprises organizations and can be mistaken for failure. When a company improves its data by replacing spend-based estimates with actual supplier data, its reported emissions usually rise, because real supplier emissions typically exceed the sector averages the estimates assumed. An organization that invests in better measurement and is rewarded with a higher reported footprint has not gotten worse; it has gotten more accurate, and the increase reflects emissions that were always there but were understated by the averages. A buyer who does not anticipate this will interpret the improvement as a regression, and may even retreat to the flattering estimates, which is the opposite of the intended effect. The perverse incentive and the counter-intuitive data dynamic together mean that a spend-based Scope 3 number tells a company remarkably little about its actual emissions or its actual progress, which is a serious problem for a figure that dominates the reported footprint.

A defender of spend-based estimation might reasonably object that the method is not meant to be precise and that criticizing its imprecision misses its purpose, which is to provide a first, rough baseline that a company then improves. This is fair as far as it goes, and the trajectory from spend-based estimates toward primary data is exactly the right one. But the objection does not rescue the numbers as reported, because companies do not disclose their spend-based estimates as rough first baselines to be improved; they disclose them as their emissions, in reports that present them with the same apparent authority as any other figure, and readers act on them as though they were measurements. The problem is not that spend-based estimation exists or that companies use it as a starting point; the problem is the gap between what the method can support, a rough directional indication, and how its outputs are presented and used, as decision-grade figures.

The perverse incentive deserves one further illustration, because its implications for corporate behavior are deeply troubling. Consider a company that wishes to show emissions reductions and is measuring Scope 3 on a spend-based method. It can achieve a reported reduction by consolidating suppliers to negotiate lower prices, which reduces spending and therefore reduces calculated emissions, even if the physical emissions are unchanged or higher. It can achieve a reported reduction by outsourcing a high-spend activity to a lower-cost provider, moving the spending off its books even if the emissions continue. In each case the company reports progress while nothing physical has improved, and in some cases while emissions have actually risen. The method does not merely fail to capture real reductions; it actively rewards financial engineering that has no environmental benefit and can have environmental harm, which means that reported progress under a spend-based method cannot be trusted to reflect real progress even directionally, and a company truly trying to reduce emissions cannot rely on its own spend-based figures to tell whether it is succeeding.

08 The assurance gap

A buyer might reasonably suppose that whatever the weaknesses of the underlying data, the assurance process, the external audit of the sustainability report, provides a backstop of reliability. It provides far less than most buyers assume, and understanding the gap between the assurance that is given and the assurance that is imagined is essential to understanding how much weight a sustainability disclosure can bear. Figure 6 sets out the distinction.

What the auditor will actually attest to

LIMITED ASSURANCE <i>(what you get today)</i>	REASONABLE ASSURANCE <i>(what a decision needs)</i>
Nothing came to our attention!	In our opinion, the figures
Negative form of opinion	Positive form of opinion
Proxies and compensation	The numbers themselves
Much lower cost and effort	Far higher cost: rarely given on Scope 3

Most sustainability assurance today is limited assurance, a negative-form opinion that nothing came to the auditor's attention suggesting the report is materially misstated. Reasonable assurance, the more extensive positive opinion, is far more expensive and is rarely offered on Scope 3 estimates. It is not about how much the software and the audit actually deliver, but about the gap between them in large.

Figure 6. Most sustainability assurance is limited assurance, a negative-form opinion, not the reasonable assurance investors imagine. Reasonable assurance is rarely offered on Scope 3 estimates.

Assurance comes in two levels, and the difference between them is fundamental. Reasonable assurance, the level provided in a conventional financial-statement audit, is a positive-form opinion: the auditor states that, in their opinion, the figures are fairly presented. Limited assurance, the level provided for most sustainability reports today, is a negative-form opinion: the auditor states that nothing came to their attention to suggest the figures are materially misstated. The difference is not a technicality. A positive opinion asserts that the numbers are right, to a reasonable standard, based on substantial testing. A negative opinion asserts only that the auditor, doing a limited amount of work, did not happen to notice anything obviously wrong. The evidentiary weight of the two is entirely different, and a reader who treats a limited-assurance sustainability figure as though it carried the reliability of an audited financial figure has badly misread what the assurance provides.

The gap matters most precisely where the data is weakest, which is Scope 3. Auditors are far more willing to provide assurance, even limited assurance, over operational emissions that can be measured, and far more reluctant to attest to spend-based Scope 3 estimates whose uncertainty they understand well. Reasonable assurance over Scope 3 is rarely offered at all, because no auditor wishes to state positively that a figure derived from sector-average factors varying several fold is fairly presented. So the assurance that exists is concentrated on the small, well-measured part of the footprint and is thin or absent on the large, poorly-estimated part, which mirrors and compounds the underlying data problem. The report as a whole carries an assurance label, but the label attaches most firmly to the part that matters least.

A buyer evaluating sustainability software and the reporting it enables should therefore ask precisely what level of assurance the output is intended to support and on which figures, because the answer determines how much the disclosure can be relied upon and how much legal protection it offers. A tool that produces a report destined for limited assurance over operational figures and no meaningful assurance over Scope 3 is producing a document whose reliability is far lower than its polished presentation suggests. This is not the software's fault, and better software cannot close the gap,

because the gap originates in the data and in what auditors are willing to attest to. But a buyer who does not understand the assurance gap will overestimate the reliability of the reports the tool produces, and will be surprised when those reports prove to be a weaker foundation, both for decisions and for legal defense, than their appearance implied.

The assurance gap has a consequence for the reliability of comparisons between companies that is worth drawing out, because comparison is one of the main purposes for which sustainability disclosures are used. Investors and analysts compare the emissions of companies within a sector to identify leaders and laggards, and those comparisons assume a degree of consistency and reliability across the disclosures that the assurance gap undermines. If two companies in the same sector use different emission-factor databases, apply spend-based estimation to different extents, and obtain different levels of assurance on different portions of their footprints, their reported figures are not comparable in the way the comparison assumes, and a ranking derived from them may reflect differences in measurement methodology rather than differences in actual emissions. The apparent comparability of the figures, all expressed in the same units with the same apparent precision, conceals a deep underlying incomparability.

This matters because decisions are made on these comparisons, by investors allocating capital, by customers selecting suppliers, and by the companies themselves benchmarking against peers. A comparison that reflects methodology rather than performance can drive capital toward a company that measures flatteringly rather than one that performs well, and can penalize a company that measures accurately, including one that has improved its data and seen its reported emissions rise as a result. The assurance gap and the underlying data inconsistency thus do not merely weaken individual disclosures; they corrupt the comparisons built on them, which are among the most consequential uses of the disclosures. A user of sustainability data who compares companies without understanding the methodological differences beneath their figures is drawing conclusions the data cannot support, and a company that measures accurately may find itself disadvantaged in exactly the comparisons the disclosures were meant to inform.

09 Greenwashing enforcement as the downside

If a weak sustainability disclosure were merely useless, the cost of compliance-theatre purchasing would be limited to the wasted software and effort. But a weak disclosure is worse than useless, because it creates a legal exposure that a company would not otherwise have, and the enforcement of that exposure is real, recent, and expensive. Figure 7 collects several recent cases.

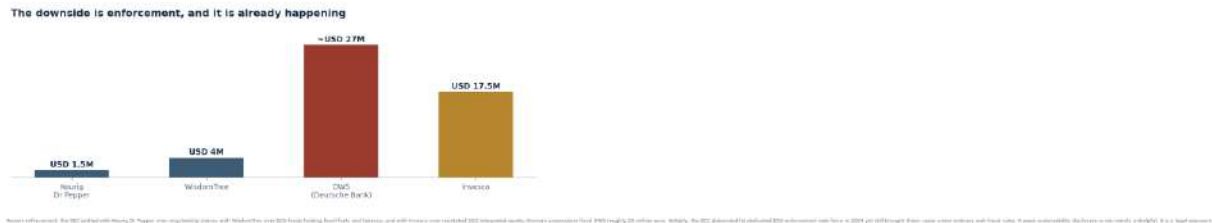


Figure 7. Recent greenwashing enforcement. Inaccurate sustainability claims drew fines under ordinary anti-fraud rules, even after the SEC disbanded its dedicated ESG task force.

The mechanism of the exposure is straightforward. When a company makes a public claim about its environmental performance, whether in a regulatory filing, a marketing statement, or a fund prospectus, that claim can be tested against the facts, and if it proves inaccurate or unsubstantiated, it becomes actionable as a misrepresentation. The sustainability report that a compliance-theatre purchase produces is a collection of exactly such claims, made with the apparent precision that software confers, and if the claims are wrong, whether because the underlying data was weak or because the presentation overstated the company's performance, the company has created a documented, published set of statements that a regulator or litigant can challenge.

The enforcement is not hypothetical. The United States securities regulator settled with a beverage company over inaccurate statements about the recyclability of its coffee pods, with an asset manager whose funds marketed as environmentally focused held fossil-fuel and tobacco investments, and with an investment adviser over misstatements about the share of its assets that were environmentally integrated, imposing penalties measured in millions and, in one case, over seventeen million dollars. German prosecutors fined a major asset manager roughly twenty-five million euro over greenwashing. And the most telling detail is that the securities regulator disbanded its dedicated environmental-enforcement task force in 2024 and continued to bring these cases anyway, under ordinary anti-fraud and disclosure rules, which means the exposure does not depend on any special sustainability-enforcement apparatus and survives the political retreat from climate regulation. The anti-fraud rules that catch a misleading environmental claim are the same ones that catch any other material misstatement, and they are not going anywhere.

The implication for a buyer is a genuine and underappreciated risk. A company that buys software to produce a sustainability disclosure it does not need, from data too weak to be reliable, has not merely wasted money; it has manufactured a set of published claims that expose it to enforcement and litigation if those claims prove inaccurate. The safest disclosure, from a pure liability standpoint, is frequently a more modest and more heavily caveated one, or in some cases no voluntary disclosure at all, precisely because every specific claim is a potential liability. A tool that encourages a company to

make more claims, in more detail, with more apparent precision, may be increasing the company's legal exposure faster than it is increasing the company's insight, and a buyer should weigh the disclosure it enables not only as a compliance asset but as a litigation risk. This is the sharpest reason that compliance-theatre purchasing is actively harmful rather than merely wasteful: it generates exposure without generating the understanding that would justify accepting it.

There is a particular trap in the interaction between voluntary disclosure and enforcement that catches well-intentioned companies, and it is worth naming because it runs counter to the instinct that more disclosure is always safer. A company that discloses more, in more detail, with more specific quantified claims, is frequently increasing rather than decreasing its legal exposure, because each specific claim is a statement that can be tested and found wanting. A vague, heavily-caveated disclosure offers little for a regulator or litigant to attack, while a detailed, confident, specific one offers many targets. The company that invests in sophisticated software to produce a detailed, precise-looking disclosure may be manufacturing exactly the specific, testable claims that create the most exposure, and doing so from data too weak to reliably support them, which is the worst combination: high confidence in presentation, low reliability in substance.

This does not mean companies should avoid disclosure, which is often required and which serves legitimate purposes, but it means the relationship between disclosure and risk is not monotonic, and more is not always safer. The safest posture is disclosure that is accurate and defensible, which frequently means more modest and more caveated than a company's marketing function would prefer, with every specific claim substantiated before it is made. A tool that encourages maximal, confident, detailed claims is pushing the company toward the high-exposure end of this tradeoff, and a buyer should be wary of the instinct, which the software often encourages, to disclose as much as possible as precisely as possible. The goal is accurate disclosure, not maximal disclosure, and the two are frequently in tension, particularly where the data cannot support the precision the tool makes it easy to present.

10 The vendor market and its unmeasurable size

The market that grew to serve this demand is large, crowded, and, tellingly, impossible to measure with any precision, and the impossibility is itself informative about how much of the market's growth is real signal and how much is hype. Figure 8 illustrates the point.

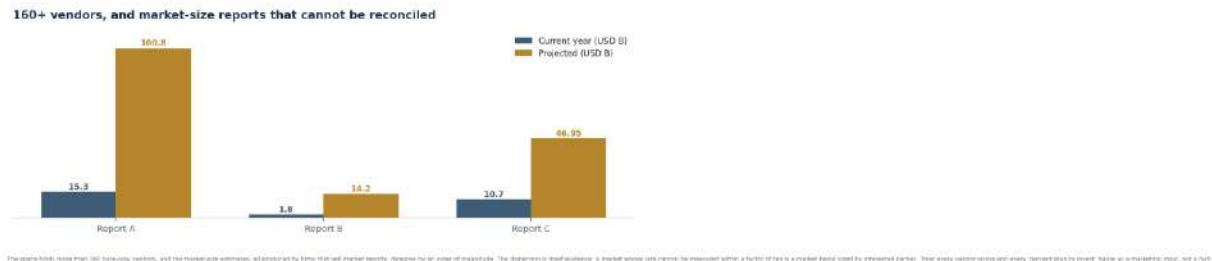


Figure 8. More than 160 pure-play vendors, and market-size reports that disagree by an order of magnitude. The dispersion is evidence that the market is being sized by interested parties.

The sustainability-software space holds more than 160 pure-play vendors by one careful count, alongside modules from the large enterprise-software firms. The vendors span carbon accounting, disclosure management, supply-chain data collection, and adjacent functions, and they are backed by substantial venture funding, with several having raised well over a hundred million dollars. The space is real and the funding is real. But the estimates of the market's size, all produced by firms that sell market-research reports, disagree wildly: one report puts the current market at fifteen billion dollars growing to a hundred billion, another at under two billion growing to fourteen, another at ten billion growing to forty-seven. These are not estimates that differ at the margin; they differ by an order of magnitude, and they cannot all be even approximately correct.

The dispersion should be read as evidence rather than as noise. A market whose size cannot be estimated within a factor of ten is a market being sized by parties with an interest in the number, using inconsistent definitions of what counts, and a buyer should treat every such figure, and every related statistic about adoption or investment intentions, as a marketing input rather than a fact. When a vendor or a sponsored report states that some large percentage of companies plan to invest in sustainability software, or that the market is growing at some impressive rate, the buyer should ask who produced the figure, how they defined the market, and what interest they have in its appearing large. The answers usually reveal that the figure is an artifact of the seller's incentives rather than a reliable measurement, and the wild disagreement among the figures confirms it.

This is not a reason to avoid the vendors, many of which offer materially useful products, but a reason to discount the market's self-description heavily and to evaluate any specific tool on what it actually does for the buyer's actual regulatory situation and decisions, rather than on the momentum the market claims for itself. The size of the market, the growth rate of the market, and the share of peers said to be investing are all irrelevant to whether a particular tool will help a particular company make better decisions or satisfy the regime that actually binds it. A buyer swept up in the market's narrative of inevitable growth is being sold the category rather than the product, and the category's own inability to measure itself is the clearest sign that the narrative should be weighted lightly.

11 The fairness case: measurement must start somewhere

This article has been sharply critical, and fairness requires a serious statement of the opposing case, which is strong. A reader who concludes that sustainability measurement is worthless and that all the software is a scam would be making a mistake as large as the compliance-theatre buyers this article criticizes, and in the opposite direction.

Begin with the regulation. The purpose behind emissions disclosure is legitimate and serious. Climate change is real, corporate emissions are a major contributor, and the premise that companies should measure and disclose their emissions so that investors, customers, and regulators can act on the information is entirely defensible. The instability of the specific regimes reflects ordinary political contestation, not a verdict that the underlying goal was wrong, and a company that measures its emissions well is doing something materially valuable regardless of whether any particular rule requires it. The critique in this article is of compliance-theatre purchasing, buying a tool to produce a document nobody acts on, not of emissions measurement as such, which is worth doing for its own sake.

On the data, the defense is that measurement must start somewhere, and that imperfect measurement is better than none. Spend-based estimation, for all its defects, is a legitimate starting point sanctioned by the governing protocol, and it is how a company with no supplier data can produce a first estimate and begin the process of improving it. The trajectory matters: a company that starts with spend-based estimates and progressively replaces them with primary supplier data is doing exactly what it should, and the fact that the early estimates are weak does not condemn the enterprise, because the estimates are meant to improve. Criticizing spend-based estimation as though it were meant to be the final answer misunderstands its role as a floor to be built upon. And the software materially helps with this: organizing the data, tracking the improvement, managing the disclosure, and maintaining the audit trail are real tasks that tools perform well, and a company doing this seriously benefits from good software.

The honest synthesis is that the tools and the measurement are valuable when bought for the right reason and used in the right way, and wasteful or harmful when bought as theatre. A company that buys sustainability software to truly understand and reduce its emissions, that treats spend-based estimates as a starting point to be improved, that aligns its disclosure to the regime that actually binds it, and that understands the limits of its data and its assurance, is making a sound investment. A company that buys the same software to produce a document that satisfies a deadline, from data it does not understand, making claims it cannot support, is engaged in theatre that wastes money and creates exposure. The difference is not in the software, which is the same, but in the purpose and the understanding the buyer brings, and the entire argument of this article is that a buyer should bring the former rather than the latter, not that the buyer should avoid the tools.

There is a version of the opposing case that deserves particular respect, which holds that the regulatory instability this article treats as a problem is actually the system working as intended. Regulation should be contested, refined, and adjusted as its costs and benefits become clear, and the Omnibus simplification can be read not as a collapse but as a correction, narrowing an over-broad directive to focus on the largest companies where the reporting burden is most justified and the data most

obtainable. On this view, the companies that fell out of scope were companies for which the original requirement was disproportionate, and their release is a feature rather than a failure. A buyer who tooled up prematurely for an over-broad regime made a forgivable error, but the regime's correction was healthy, and the lesson is about timing rather than about the legitimacy of the underlying enterprise.

This is a serious argument and it has real force, particularly for the smaller companies that the Omnibus released, for whom detailed emissions reporting truly was a disproportionate burden relative to their footprint and their capacity. But it does not fully dissolve the compliance-theatre critique, for two reasons. First, the correction came after the purchasing, so the companies that responded to the original regime bore the cost regardless of the correction's wisdom, and the instability imposed real waste even if the eventual settling point is better. Second, and more fundamentally, the critique of compliance-theatre purchasing does not depend on the regulation being wrong; it depends on the purchase being made for the deadline rather than for the decision. Even under a perfectly stable, perfectly calibrated regime, a company that buys software to produce a document nobody acts on, from data that cannot support a decision, is engaged in theatre. The regulatory instability made the theatre visible and expensive, but the theatre would exist even if the regulation were beyond criticism, because its source is the buyer's purpose, not the regulation's quality.

12 Buy for decision-usefulness, not compliance

The governing principle for a buyer follows from everything above and can be stated simply: buy sustainability software for its usefulness to decisions, not for its ability to produce a compliant document, because the document is unstable in its requirement and weak in its data, while decision-usefulness is durable and real. Figure 9 sets out the resulting discipline.

Buying sustainability software without buying compliance theatre

Pre-brief that better data raises emissions	So the improvement is not mistakes for a failure
Know the assurance level	Limited or reasonable, and on which figures
Align to the post-Omnibus regime	The rule that actually applies to you now, not the deferred one
Demand methodology transparency	Activity-based vs spend-based, stated uncertainty ranges
Buy for decisions	Not just to produce a report nobody acts on

None of this argues against ensuring emissions, which is important and increasingly required. It argues for buying a tool that supports decisions rather than one that produces a document, for demanding to know how the numbers are made, and for aligning the purchase to the regime that survives the omnibus rather than the one that was deferred or intended.

Figure 9. Buying sustainability software without buying compliance theatre: for decisions, with methodology transparency, aligned to the surviving regime, with clear assurance.

The principle inverts the usual purchasing logic, which starts from the compliance requirement and asks what tool satisfies it. That logic is exactly what the regulatory instability has undermined, because the requirement it starts from has proven to move, narrow, and collapse. A more durable logic starts from the company's own decisions and asks what measurement would improve them: which suppliers to prefer, which products to redesign, which logistics to change, which investments to prioritize. Measurement that serves those decisions is worth having regardless of what any regulation requires, because the decisions exist independently of the regulation, and a tool that supports them retains its value even when the compliance requirement moves.

This does not mean ignoring compliance, which remains a real obligation for the companies still in scope and for those bound by the rules that survived. It means treating compliance as one output of a measurement capability built for decisions, rather than as the purpose of the whole exercise. A company that has built a genuine capability to measure and reduce its emissions can produce whatever disclosure the applicable regime requires as a byproduct, and can adapt as the regime changes, because the underlying capability is regime-independent. A company that has built only a compliance-reporting capability, tuned to a specific rule, must rebuild every time the rule changes, and has nothing of independent value when the rule collapses. The decision-first approach is more robust precisely because it does not depend on the stability of any particular regulation, which recent history has shown cannot be relied upon.

Applying the principle in practice means, at the point of purchase, evaluating a tool against the company's decisions rather than against a compliance checklist. Does the tool produce data at a granularity and reliability that could actually inform a supplier or product decision, or only an aggregate figure suitable for a disclosure? Does it distinguish activity-based from spend-based data and expose the uncertainty, so that decisions rest on the more reliable numbers? Does it help the company improve its data over time, moving from estimates toward primary data? A tool that scores well on these questions

is worth buying whatever the regulation does, and a tool that scores well only as a compliance-document generator is a theatre purchase waiting to be stranded by the next regulatory change. The buyer who evaluates on decision-usefulness is buying something durable; the buyer who evaluates on compliance is buying something the legislature can devalue overnight.

The decision-first approach also changes the internal conversation about sustainability software in a way that improves the organization's relationship to the whole enterprise. When a tool is bought for compliance, it belongs to whoever owns compliance, is funded as a cost of regulation, and is resented as overhead that produces nothing the business values. When a tool is bought for decisions, it belongs to the functions whose decisions it improves, is funded as an investment in those decisions, and is valued for the insight it provides. The same software, bought for a different reason, occupies an entirely different place in the organization, and the decision-first framing tends to produce better outcomes not only because it is more robust to regulatory change but because it embeds the tool in the parts of the business that will actually use its outputs, rather than isolating it in a compliance function that produces reports the business ignores. The framing shapes the usage, and decision-first framing produces usage that creates value.

13 A procurement protocol, and a scoring rubric

The practices in this article combine into a procurement protocol that treats sustainability software as a decision tool with a compliance byproduct, and that protects the buyer against both the regulatory instability and the data weakness. The protocol has several components, applied before purchase and during use.

Before purchase, the buyer establishes which regime actually binds the company after the Omnibus and the other regulatory changes, rather than the regime that prompted the initial interest, and scopes the requirement to that. The buyer then evaluates candidate tools against the company's own decisions, requiring the tool to expose the distinction between activity-based and spend-based data and to represent the uncertainty in its estimates, and establishes what level of assurance the output is intended to support and on which figures. During use, the buyer governs who owns the numbers and their methodology, plans for the reality that improving data will raise reported emissions, and treats every published claim as a potential liability to be substantiated rather than a marketing opportunity to be maximized.

A scoring rubric

The dimensions below let a buying committee assess whether it is purchasing a decision tool or a theatre prop.

Dimension	What good looks like	Red flag
Purpose	Bought to inform specific decisions	Bought to satisfy a deadline
Regime alignment	Scoped to the rule that binds you now	Scoped to a deferred or repealed rule
Data transparency	Distinguishes activity- from spend-based	Presents one blended figure
Uncertainty	Shows ranges and confidence	Reports point estimates as precise
Assurance	Level and scope explicitly understood	Assurance assumed, not checked
Improvement path	Helps move from estimates to primary data	Locks in spend-based estimation
Legal posture	Claims substantiated before publication	Maximizes claims for marketing

A purchase scoring in the left-hand column is a decision tool that will retain its value as regulations shift and will not manufacture legal exposure. A purchase scoring in the right-hand column is compliance theatre, tuned to an unstable requirement, built on data that cannot support a decision, generating claims that create risk. The rubric does not make sustainability measurement easy, which the data problem prevents. It ensures the buyer knows which kind of purchase they are making, which is the distinction the deadline pressure is designed to obscure.

The protocol repays a specific discipline at the moment of vendor selection, which is to run the tool against a real decision the company actually faces, rather than against a compliance checklist. Give the candidate tool a genuine question the business needs answered, which supplier in a category is the

higher-emissions choice, or where in the product line the largest reduction opportunity lies, and see whether the tool can produce an answer reliable enough to act on. A tool built for decisions will engage the question and expose the uncertainty candidly; a tool built for compliance will produce an aggregate figure that does not resolve the decision, or will produce a confident-looking answer that dissolves under scrutiny of its data. The decision test reveals, better than any feature comparison, whether the tool is a decision instrument or a document generator, and it is available before purchase for the cost of posing the question.

A second discipline worth building into the protocol is an explicit sunset review tied to the regulatory situation. Because the regime that justified a purchase can move, the buyer should schedule a review, at a defined interval, of whether the requirement the tool was bought to satisfy still exists in the form anticipated, and whether the tool still serves the company's decisions independent of that requirement. A tool that was bought for a deadline that has since moved, and that serves no independent decision purpose, should be discontinued rather than maintained out of inertia, and the sunset review forces that question to be asked rather than allowing a stranded compliance tool to persist as a line item nobody re-examines. The review institutionalizes the recognition that a compliance purchase is only as durable as the compliance requirement, and it prevents the organization from continuing to pay for a capability whose justification has quietly disappeared.

14 Governing who owns the numbers

A dimension that purchasing decisions routinely neglect, and that determines whether a sustainability-measurement capability is trustworthy, is governance: who inside the organization owns the numbers, their methodology, and the claims made from them. Software does not answer this question, and a tool deployed without clear ownership produces figures that no one is accountable for, which is precisely the condition under which weak data and overstated claims flourish.

The ownership question has several parts. Someone must own the methodology, deciding which emission factors and databases are used, when spend-based estimates are replaced with primary data, and how uncertainty is represented, because these choices determine the numbers and must be made deliberately and consistently rather than defaulting to whatever the tool does out of the box. Someone must own the data quality, ensuring that the inputs are accurate and that the improvement from estimates toward primary data actually happens rather than stalling once the first compliant report is produced. And someone must own the claims, reviewing every public statement derived from the numbers to ensure it is substantiated and defensible before publication, given the enforcement exposure that unsubstantiated claims create.

In many organizations these responsibilities are diffuse or unassigned, with the sustainability report produced by a small team under deadline pressure, using whatever the software generates, reviewed by no one with the authority or the incentive to challenge the numbers or moderate the claims. This is the governance vacuum in which compliance theatre thrives, because there is no one whose job is to ask whether the numbers are reliable enough to publish or whether the claims are defensible enough to make. The report is produced because a deadline requires it, using data no one owns, making claims no one has substantiated, and the organization discovers the weakness only when a regulator or litigant tests a claim and finds it wanting.

The remedy is to assign the ownership explicitly, with real authority and accountability, before the reporting capability is built, and to give the owner the standing to moderate the claims that the organization publishes. This is uncomfortable, because it means someone must be accountable for numbers that are inherently uncertain and for claims that marketing would prefer to maximize, and it means the owner must sometimes insist on a more modest disclosure than the organization wants to make. But the alternative, numbers and claims that no one owns, is exactly what produces both the useless reports and the legal exposure this article has described. A sustainability-measurement capability is only as trustworthy as its governance, and a buyer who invests in software without investing in the governance to own its outputs has bought a tool that will generate unowned numbers and unsubstantiated claims, which is a liability rather than an asset. The software is the easy part; the accountable ownership of what it produces is the part that actually determines whether the capability is worth having.

The governance vacuum has a specific and dangerous manifestation in the handling of the uncertainty ranges that ought to accompany every emissions figure. Someone must decide how uncertainty is represented in the report, and in the absence of clear ownership, the default is to represent it not at all,

reporting point estimates without ranges because ranges look less authoritative and invite questions. This default is precisely wrong, because it presents figures that are uncertain by wide margins as though they were precise, which is both misleading to readers and legally dangerous to the company, since a point estimate that proves wrong is a clearer misrepresentation than a range that encompassed the truth. An owner accountable for the methodology would insist on representing the uncertainty, accepting that the resulting figures look less impressive in exchange for their being defensible, but in the governance vacuum no one makes that choice and the misleadingly precise point estimate becomes the default.

Ownership of the claims, as distinct from ownership of the numbers, deserves particular emphasis because it is the point at which the enforcement exposure is actually created or avoided. The numbers are internal until they are published as claims, and it is the claims, the public statements derived from the numbers, that a regulator or litigant tests. An owner accountable for the claims would review each public statement against the data that supports it, would insist that unsupported or weakly-supported claims be moderated or dropped, and would resist the pressure from marketing and leadership to make the strongest possible claims. This is an uncomfortable and frequently unpopular role, because it means being the person who insists on saying less than the organization wants to say, but it is the role that prevents the company from publishing the specific, testable, weakly-supported claims that greenwashing enforcement targets. A company that assigns clear ownership of its claims, with the authority to moderate them, has a defense against enforcement that a company with unowned claims does not.

15 Conclusion: the tool is not the disclosure

The scramble of 2025 and 2026 taught a lesson that extends well beyond sustainability software, which is that a purchase justified entirely by a regulatory deadline is only as stable as the deadline, and regulatory deadlines are far less stable than buyers assume. The companies that bought sustainability tools to satisfy a mandate discovered that mandates can be delayed by two years, narrowed to exclude most of the companies they covered, or abandoned outright, and that a tool bought against a mandate is stranded when the mandate moves. The companies that bought the same tools to truly understand and reduce their emissions were unaffected by the regulatory instability, because their purpose was independent of the regulation, and their tools kept doing useful work regardless.

The deeper lesson concerns the relationship between the tool and the disclosure. A sustainability disclosure is a set of claims about a company's environmental performance, and those claims are only as good as the data beneath them and only as safe as their substantiation. The software produces the disclosure but cannot make the underlying data reliable, cannot close the assurance gap, and cannot substantiate the claims; those depend on the quality of the measurement and the rigor of the governance, which are the company's responsibility and not the tool's. A buyer who believes that purchasing the software has produced a reliable disclosure has confused the instrument with the result, and will discover the difference when the disclosure is tested, either by a decision that the data cannot actually support or by a regulator who finds a claim unsubstantiated.

What a serious buyer does is straightforward, if more demanding than buying a compliance-document generator. They buy for decision-usefulness rather than for a deadline, so the purchase survives regulatory change. They demand to understand how the numbers are made, so they know which figures can bear weight. They align to the regime that actually binds them after the Omnibus, rather than the one that prompted their interest. They understand the assurance gap and the enforcement exposure, so they neither overstate the reliability of their disclosures nor manufacture unnecessary legal risk. And they govern who owns the numbers and the claims, so that the outputs are accountable. The regulation may move again, as it has moved before, and the data will remain difficult, as it is inherently difficult. A buyer who has built a genuine capability to measure and reduce emissions is well positioned for whatever comes; a buyer who has bought compliance theatre is one legislative session away from owning a tool for a requirement that no longer exists, and one published claim away from an enforcement action they created themselves.

There is a broader reflection worth offering, about what the sustainability-software episode reveals concerning the relationship between regulation, technology, and corporate behavior more generally. The premise behind mandatory disclosure is that requiring companies to measure and reveal something will cause them to improve it, and the premise is sound in principle. But the episode shows how the premise can fail in practice, when the measurement is too weak to guide improvement and the requirement produces a document rather than a change. The regulation succeeded in causing companies to buy software and produce reports; it frequently failed to cause them to reduce emissions, because the reports rested on data that could not guide reduction and were produced to satisfy a deadline rather than to inform a decision. The gap between producing a disclosure and producing a

change is the gap in which compliance theatre lives, and it is a gap that mandatory-disclosure regimes must close if they are to achieve their purpose rather than merely their letter.

16 Methodology, caveats, and sources

Methodology

- This article synthesizes regulatory texts and timelines, emissions-accounting standards and academic analysis of emission factors, enforcement records, assurance-standard materials, and vendor-market data, current to mid-2026. Supply Chain Research is independent and accepts no payment from the software vendors, consultancies, or auditors discussed.
- Where a statistic originates with an interested party, such as a vendor market-size estimate or an investment-intention survey, it is identified as such and treated as directional rather than authoritative.

Caveats

- Regulatory regimes described here were changing as this was written. The EU Omnibus package, the status of the Green Claims Directive, the US SEC rule, and the California laws were all subject to ongoing legislative, administrative, or judicial processes, and readers should verify the current state of any regime that applies to them.
- The estimate that roughly 80 percent of previously-covered companies fall out of CSRD scope is drawn from figures cited in the legislative process and by law firms tracking the Omnibus, and reflects the raised thresholds; precise counts depend on definitions and data not settled here.
- Scope 3 multiples, emission-factor variation, and the share of companies using primary versus estimated data are drawn from CDP, BCG, academic, and standards-body sources and represent averages across varied populations. Individual companies vary widely.
- The enforcement cases are described from public regulatory records. Several involved settlements without admission of wrongdoing, and the fairness of any particular action is not assessed here; the cases are cited to establish that the exposure is real, not to characterize any company.
- Figures 5 and 8 are illustrative constructions used to convey a mechanism and a dispersion, not published benchmarks, and are labelled as such. Market-size figures in Figure 8 are drawn from differing third-party reports to show their disagreement.

Sources

- [1] European Parliament. [Legislative train: the Omnibus simplification package and the Stop-the-Clock directive \(Directive \(EU\) 2025/794\)](#).
- [2] European Commission. [Corporate Sustainability Reporting Directive: scope, timeline, and simplification](#).
- [3] US Securities and Exchange Commission. [SEC charges Keurig for inaccurate statements regarding recyclability \(Release 2024-122\)](#).
- [4] US Securities and Exchange Commission. [SEC charges WisdomTree and Invesco over ESG-related misstatements \(enforcement releases\)](#).
- [5] CDP. [Global Supply Chain Report 2024: Scope 3 emissions relative to operational emissions](#).
- [6] Greenhouse Gas Protocol. [Technical Guidance for Calculating Scope 3 Emissions \(spend-based and activity-based methods\)](#).
- [7] Greenhouse Gas Management Institute. [On the limits of spend-based emission factors and decarbonization incentives](#).
- [8] Latham & Watkins. [The EU Green Claims Directive: intended withdrawal and uncertain status](#).

Additional context drawn from regulatory filings and enforcement records, from CDP and BCG emissions research, from academic analysis of emission-factor variation, from assurance-standard materials, and from third-party market reports whose disagreement is itself cited. Interested-party statistics are identified as such. This article is analysis, not legal, accounting, or investment advice, and applicable regimes should be verified against current sources.

Supply Chain Research is an independent, vendor-neutral research platform for supply chain and technology leaders. We accept no payment from the vendors, consultancies, or auditors discussed. This article is analysis, not legal, accounting, procurement, or investment advice, and its conclusions should be validated against your own circumstances and current regulation before any decision.