SUPPLY CHAIN RESEARCH · EDITORIAL FEATURE

The RFP

Industrial Complex

How Enterprise Software Selection Is Decided Before the Competition Begins

Requirements copied from a vendor capability list. Weights set before a single response arrives. Demonstrations the vendor scripts, on the vendor data, in the vendor environment. The process produces the appearance of rigorous selection. The decision was made earlier.

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

Supply Chain Research | supplychainresearch.com | 2026

Contents

How Enterprise Software Selection Is Decided Before the Competition Begins

01 Executive summary

An enterprise software selection looks like a competition. Requirements are published, vendors respond, responses are scored against weighted criteria, finalists demonstrate their products, references are checked, and an award is made to the highest-scoring proposal. The apparatus is elaborate, it consumes months of effort on both sides, and it produces a documented trail showing that the decision was reached objectively. What it frequently does not produce is a decision, because the decision was already substantially constrained by the time the first vendor was invited. The requirements were written from one vendor's capability description, the weights were set to favour the criteria that vendor leads on, and the demonstrations were run by the vendors on their own data in their own environments following their own scripts.

This article argues that enterprise software selection has drifted into a form of procedural theatre in which the visible competition documents a choice that was effectively made earlier and elsewhere. The evidence comes from an unusual place: public procurement, where the practices this article describes are unlawful and where an external forum publishes what happens when buyers are challenged. Federal bid-protest data show that roughly half of challenges result in some relief for the protester, and that the most common grounds for sustaining a protest concern unreasonable evaluation and flawed selection decisions. Those are the regulated cases, subject to statutory bans on restrictive specifications, mandatory equivalence language, adviser conflict rules, and external review. Private enterprise buying has none of that, and this article says honestly that the absence of a rule is not evidence of a problem. But the mechanisms that public law bothers to prohibit are the mechanisms that operate freely everywhere else, and a buyer who has never examined the authorship of its own requirements has no basis for believing its process is deciding anything.

~52%
of US federal bid protests result in some relief for the protester, in a regime that bans wired specifications

3
the most common grounds for sustaining a protest: unreasonable technical evaluation, flawed selection, unreasonable cost evaluation

0
independent sampled studies of private enterprise RFP outcomes located; the published figures come from vendors selling response software

The argument in brief

- **The requirement is the decision.** Whoever writes the requirements sets the field before it is assembled. Everything after that stage documents a choice already constrained, which is why the authorship of requirements deserves the scrutiny usually spent on the evaluation.
- **The weights decide, and nobody audits them.** Identical raw scores produce opposite winners under different weightings. The weighting is set before any response arrives, is rarely published, and is almost never challenged.
- **Public law prohibits what private buying permits.** Restrictive specifications, unjustified brand-name requirements, and adviser conflicts are regulated in public procurement and merely customary in enterprise buying. The rules exist; they do not apply.

- **A vendor-controlled demonstration proves nothing.** A script run on curated data in a tuned environment establishes that the product can be made to perform. Only a buyer-controlled test on the buyer's own data and exceptions establishes whether it performs here.
- **Move the rigour, do not abandon the process.** The remedy is not to stop running competitions but to relocate the discipline to the requirement and the weighting, separate the adviser from the implementer, and test on real data.

02 The decision is made before the competition

The structural claim of this article can be stated before any evidence is offered, because once it is stated most experienced buyers recognise it immediately from their own practice. In an enterprise software selection, the decisive act is the authorship of the requirements, and that act occurs before any competitor is invited to participate. Figure 1 locates it.

Figure 1. The competition is decided before it is announced. Requirements written from one vendor's capability description narrow the field before the field is assembled, and every later stage documents a choice already constrained.

Consider what a requirements document does. It specifies what the software must do, in what manner, with what interfaces, at what scale, and to what standards. Every specification narrows the set of products that can satisfy it, which is the point of writing specifications, and the narrowing is legitimate when it reflects what the organization actually needs. The narrowing becomes something else when the specifications describe not what the organization needs but what a particular product happens to do, expressed at a level of detail that only that product satisfies. A requirement that the system support a named workflow pattern, or a specific data structure, or a particular integration approach, may be a genuine operational necessity or may be a transcription of one vendor's design decisions, and from the outside the two are indistinguishable.

The consequence is that the field of viable respondents is determined at the requirements stage, and everything afterward operates on that field. If three products can satisfy the requirements and eight cannot, then eight vendors are eliminated before they are invited, without any evaluation of whether they might have served the organization better under a different specification. The scoring that follows is a comparison among survivors, and it is conducted with real care, which is precisely what makes the process persuasive. The scores are honest, the evaluators are diligent, the arithmetic is correct, and the conclusion is sound given the field. The field was the decision.

This is not usually a conspiracy and it is rarely dishonest. The most common path to a constrained requirements document is entirely innocent: the team charged with writing requirements has limited time and no template, it has recently seen a vendor demonstration or read a vendor capability guide, and it uses that material as the scaffold for its own document because doing so is faster than starting from a blank page. The resulting requirements reflect that vendor's structure because the structure came from there, and nobody involved intended to prejudice the outcome. The effect on the competition is identical whether the constraint was deliberate or accidental, which is why the diagnosis matters more than the motive.

The purpose of this article is not to argue that competitive selection should be abandoned, which would be a considerable overcorrection, but to relocate the scrutiny to where the decision actually occurs. An organization that spends four months evaluating responses and four days writing requirements has allocated its diligence in inverse proportion to the influence of each stage. The sections that follow set out the evidence for how the mechanisms work, drawing on the one domain where they are regulated and documented, and then describe what a buyer can do about them.

There is a diagnostic test a buyer can apply to its own past selections that settles the question quickly. Take the last three completed software selections and, for each, ask two questions of the surviving documents: where did the requirements document originate, and could any respondent other than the eventual winner have satisfied the mandatory requirements as written. The first question is usually answerable from file properties, email history, or the memory of whoever assembled the document. The second requires reading the mandatory requirements against what the losing bidders actually offered, which the responses themselves record. Organizations that run this test are frequently surprised, not because anyone behaved improperly but because the constraint had never been examined from that angle.

The test matters because the alternative is to reason about selection quality from outcomes, which is unreliable. A selection that produced a satisfactory system is assumed to have been a good process, and a selection that produced a disappointing one is assumed to have been a poor process, when in fact a constrained field can produce a satisfactory result by luck and an open field can produce a disappointment through ordinary risk. Process quality and outcome quality are related but distinct, and only the process can be examined directly. An organization that judges its selection discipline solely by whether it likes the resulting systems has no feedback loop at all on the part of the process this article addresses.

03 Who actually writes the requirements

If the requirements determine the outcome, the question of who writes them becomes the central question of the whole process, and the answer is frequently that they were written, at one remove, by a party with an interest in the result. There are three common routes by which this happens, and none of them requires anyone to behave improperly.

The first route is the analyst template. Major research firms sell packaged toolkits containing pre-built requirements and vendor-evaluation scoring models for specific software categories, including warehouse management, and market them as configurable templates that a buyer can use to define requirements, solicit proposals, and evaluate responses. These products are plainly useful, they save substantial time, and they are produced by firms that also sell research and advisory services to the vendors being evaluated and that publish comparative rankings of those vendors. A buyer using such a template is adopting a definition of the category, and of what matters within it, that was authored by a party embedded in the vendor ecosystem. That is not the same as bias, and it is not independence either.

The second route is the specialist template vendor. Several firms sell downloadable requirements and proposal templates across enterprise resource planning, supply chain, and customer relationship categories, typically as spreadsheets containing hundreds or thousands of individual requirement lines that a buyer marks as mandatory, desirable, or not required. The commercial model rewards comprehensiveness, because a template with more lines appears more thorough, and the result is requirement sets far longer than any organization needs, in which the lines that actually matter are diluted among hundreds that do not. A buyer working through such a list will mark many things mandatory that are merely conventional, and each of those marks narrows the field for no operational reason.

The third and most direct route is the vendor itself. It is a routine and openly acknowledged part of enterprise software sales to help a prospective buyer articulate its requirements, and vendors provide requirement checklists, discovery workshops, and capability documents for exactly this purpose. From the vendor's perspective this is helpful pre-sales support and it frequently is helpful, since the vendor knows the category better than a buyer purchasing in it for the first time. From the buyer's perspective it means that the document defining the competition may originate with a competitor. One advisory practitioner has described the endpoint plainly, observing that it is common for a preferred vendor to shape the requirements before the request for proposal is issued, sometimes to the point that the specifications read like one competitor's capability description. That observation is a practitioner assertion rather than a sampled finding, and it is presented here as such, but it describes a mechanism that anyone who has run these processes will recognise.

A fourth route deserves separate mention because it is the least visible and operates even where no external document is involved, which is the influence of prior exposure. A team that has spent two years watching demonstrations, attending vendor conferences, reading category marketing, and speaking with peers who use a particular product will have absorbed a way of thinking about the category that

reflects how the dominant vendors describe it. When that team writes requirements from scratch, believing itself uninfluenced, it reproduces the conceptual structure it has absorbed: the module boundaries, the terminology, the assumed workflow sequence, the implied division between what the system does and what surrounds it. No document was copied and the constraint is real.

This is why the remedy proposed later starts from the organization's own processes rather than from any account of what the software category contains. Walking an actual workflow, recording what the organization does and where it struggles, produces a description of need expressed in the organization's own terms, which is the only reliable defence against inherited category framing. It is slower than adapting a template and it is the step that separates a requirements document that describes a business from one that describes a product class. Buyers who have done both report that the process-derived version is shorter, more specific about what matters, and considerably less specific about how the software should work, which is exactly the profile that keeps a field open.

04 What the regulated case reveals

Private enterprise buying generates almost no public record, which is why the argument so far has rested on mechanism rather than measurement. There is one domain where the same practices occur, are regulated, are challenged, and are documented in an annual public report, and that domain provides the closest available evidence. Figure 2 presents it.

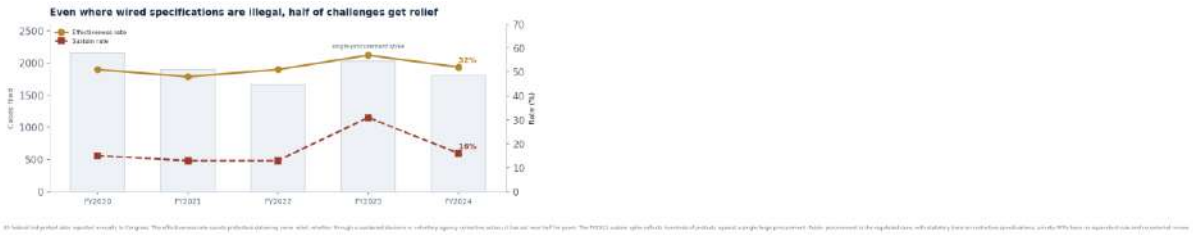

Figure 2. US federal bid-protest data reported annually to Congress. The effectiveness rate, counting protesters who obtain some relief through a sustained decision or voluntary agency corrective action, has sat near half for years.

The United States federal government publishes an annual accounting of bid protests, the mechanism by which a disappointed bidder challenges a procurement decision before an independent forum. In the most recent full year shown, roughly eighteen hundred cases were filed, of which sixteen percent were formally sustained. The more revealing figure is the effectiveness rate of approximately fifty-two percent, which counts protesters who obtained some relief either through a sustained decision or because the agency voluntarily took corrective action once challenged. Over a ten-year span the sustain rate averages about seventeen percent and the effectiveness rate about forty-nine, so the recent figures are typical rather than anomalous. The apparent spike in the sustain rate two years earlier is an artefact of several hundred protests filed against a single large procurement and should not be read as a trend.

The most useful detail in the report is not the rates but the grounds. The three most prevalent bases on which protests are sustained, unchanged across consecutive years, are unreasonable technical evaluation, flawed selection decision, and unreasonable cost or price evaluation. Every one of these concerns how the buyer conducted the assessment rather than what the vendors offered. The independent forum, reviewing the documented record of federal procurements conducted under statutory constraints, finds most often that the evaluation itself was defective. These are the procurements with published criteria, mandated documentation, conflict rules, and the knowledge that any bidder may demand external review.

The inference to draw requires care, and it is not that private buying must be worse by some measurable factor, which the data cannot support. It is narrower and still substantial. Evaluation defects sufficient to warrant relief occur at a material rate under the most constrained conditions available, in processes designed from the outset to withstand external challenge. Private enterprise selections operate without published criteria, without mandated documentation, without conflict rules, and with no external forum at all. There is no reason to expect the underlying human and organizational tendencies that produce evaluation defects to be absent, and there is no mechanism by which they would be detected. The public data do not measure private practice; they establish that the failure modes are real and common in the environment least conducive to them.

A second feature of the protest record repays attention, which is the distribution of outcomes between formal decisions and voluntary corrective action. The sustain rate counts only cases the forum decided in the protester's favour, while the effectiveness rate additionally counts cases where the agency, having seen the challenge, chose to reopen or amend the procurement rather than defend it. The gap between the two figures, roughly sixteen percent against fifty-two, is largely composed of procurements the buyer decided not to defend once someone examined them. That is a striking proportion, and it suggests that a substantial share of defects are visible enough to the buyer, on inspection, that defending them was judged unwise.

The implication for private buying follows directly. Voluntary corrective action requires two conditions: a party with standing and incentive to examine the decision, and a buyer who anticipates consequences from being wrong. Private enterprise selection supplies neither. A losing vendor has commercial reasons to accept the outcome gracefully and preserve the relationship for the next opportunity, and there is no forum in which a challenge could be heard even if one were made. The defects that public buyers correct when challenged would, in a private process, simply persist unexamined, and the resulting system would be implemented and lived with for a decade without anyone establishing whether the selection reasoning held together.

05 The line that public buyers cannot cross

The regulated domain is instructive in a second way, which is that the specific practices the law prohibits identify precisely which mechanisms legislators concluded were capable of subverting a competition. Reading the prohibitions as a diagnosis of the risks is more useful than reading them as compliance obligations. Figure 3 sets the two regimes side by side.

Figure 3. Public procurement law prohibits restrictive specifications, requires equivalence language, regulates adviser conflicts, and provides external review. None of this applies to a private enterprise buying planning or warehouse software.

United States federal acquisition regulation states that agency requirements shall not be written so as to require a particular brand name, product, or feature peculiar to one manufacturer, thereby precluding consideration of a product made by another company, unless the particular feature is essential and market research establishes that other products will not meet the need. Where a brand-name or equal description is used, the regulation requires the buyer to identify the salient characteristics that the equivalent product must meet, which prevents the brand name from doing the work of a specification. Sole-source awards require documented justification and approval, and for defence acquisitions the use of brand-name descriptions carries its own justification requirement.

European public procurement law reaches the same place by a different route. The directive governing technical specifications provides that specifications shall not refer to a specific make, source, process, trademark, patent, or production which would have the effect of favouring or eliminating certain undertakings, unless justified by the subject matter of the contract, and where such a reference is permitted it must be accompanied by words indicating that equivalents are acceptable. The Court of Justice has applied this in recent litigation concerning material specifications that had the practical effect of excluding a class of competing products. The principle is that a buyer must specify what it needs the thing to do, not which thing it has already chosen.

Both regimes also regulate the adviser problem examined later in this article, through organizational conflict-of-interest rules that restrict a contractor who helps prepare requirements from competing for the resulting work. And both provide the external forum whose output was examined in the previous section. The composite picture is that legislators, examining how competitive procurement can be subverted, identified requirement authorship, brand-name specification, adviser conflicts, and the absence of review as the four principal mechanisms, and prohibited or constrained each. A private enterprise buying enterprise software is subject to none of these constraints, which is not a scandal, since private buyers spend their own money and answer to their own governance. It does mean that the four mechanisms identified as dangerous enough to legislate against operate without any restriction in

the setting where most enterprise software is actually bought, and that a buyer who wants a real competition has to supply the discipline privately, because nothing external will supply it.

06 The weights decide, and nobody audits them

The second lever that determines outcomes, after requirement authorship, is the weighting applied to evaluation criteria. It receives less attention than either the requirements or the scores, and it is capable of reversing a result without anyone changing a single assessment. Figure 4 demonstrates the arithmetic.

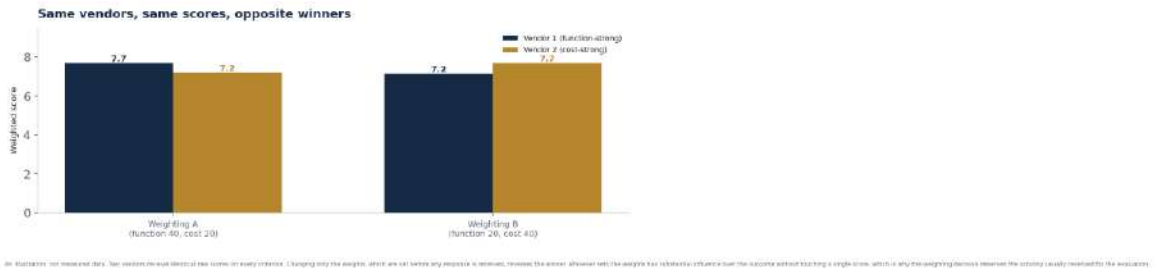

Figure 4. An illustration, not measured data. Two vendors receive identical raw scores on every criterion. Changing only the weights, set before any response is received, reverses the winner.

The demonstration is deliberately simple because the point does not require complexity. Two vendors are evaluated against five criteria and receive identical raw scores under both scenarios: one is stronger on functional fit, the other on total cost, and they are close elsewhere. Under a weighting that assigns forty percent to functional fit and twenty to cost, the first vendor wins. Under a weighting that reverses those two figures, the second vendor wins. No score changed, no evaluator changed an assessment, and no vendor performed differently. The outcome was determined by a decision taken before any response arrived, by whoever set the weights.

This would be unremarkable if the weighting were treated as the consequential decision it is, subjected to explicit deliberation, documented reasoning, and senior approval. In most selections it is not. The weighting is frequently assembled quickly, sometimes by the same person who drafted the requirements, sometimes carried forward from a previous exercise, and sometimes adopted from a template. It is rarely published to respondents, rarely revisited once scoring begins, and almost never challenged, because challenging it would require someone to notice that it is doing the work. An organization that would insist on executive sign-off for a shortlist decision will frequently allow a mid-level analyst to set the weighting that determines the shortlist.

There is a further and subtler problem, which is that weights are sometimes adjusted after responses arrive. The stated reason is usually reasonable: the responses revealed that a criterion mattered more or less than anticipated, and the weighting should reflect what was learned. The difficulty is that once the scores are visible, any adjustment to the weights has a knowable effect on the ranking, and the person making the adjustment cannot unknow it. Whether or not the adjustment is motivated by the effect, it is not distinguishable from an adjustment that was, and the process loses the property that made it credible. The discipline that solves this is trivial to state and uncomfortable to follow: fix the weights before any response is opened, publish them, and do not change them afterward.

The weighting problem has a compounding feature that deserves attention, which is the interaction between weights and criterion granularity. A criterion that is decomposed into many sub-criteria tends

to accumulate influence beyond its nominal weight, because each sub-criterion carries its own share and the aggregate frequently exceeds what a single-line criterion of the same stated weight would carry. A selection in which functional fit is broken into forty detailed sub-requirements while total cost of ownership appears as one line will, in practice, weight function far more heavily than the headline percentages suggest, regardless of what the summary weighting states. Buyers examining a weighting scheme should therefore look at the structure as well as the percentages, because the structure is doing work that the percentages conceal.

A related distortion arises from scoring scales that are not comparable across criteria. If one criterion is scored on demonstrated capability, where most serious vendors cluster near the top, and another is scored on price, where the spread is wide, then the criterion with the wider spread contributes more variance to the total and therefore more influence over the ranking, again irrespective of the nominal weights. This is a well-understood property of composite indices and it is routinely ignored in procurement scorecards. The practical remedy is to examine, before finalising the scheme, how much each criterion is likely to differentiate the field, and to set weights with that differentiation in mind rather than treating the percentages as though they map directly onto influence.

07 The contested method underneath the scorecard

Weighted scoring is so familiar in procurement that it is treated as a neutral arithmetic convenience rather than as a methodological choice. It is a methodological choice, it rests on a formal apparatus that has been contested in the peer-reviewed literature for decades, and buyers relying on it should know that. Figure 5 summarises the dispute.

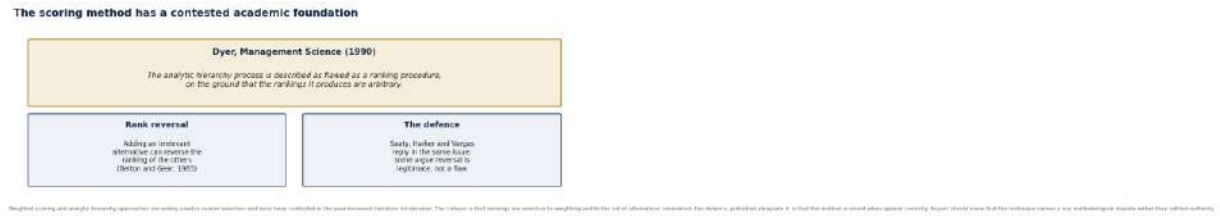

Figure 5. Weighted-scoring and analytic hierarchy approaches are widely used in vendor selection and have been contested in the literature for decades. The critique concerns sensitivity to weighting and to the set of alternatives; the defence was published alongside it.

The formal method underlying most structured vendor-selection scorecards is the analytic hierarchy process, introduced in the nineteen eighties as a way of decomposing a complex decision into a hierarchy of criteria, eliciting pairwise comparisons, and deriving weights and rankings from them. It is elegant and it made structured multi-criteria decision-making accessible to practitioners. In nineteen ninety, a paper in the leading operations research journal argued that the method is flawed as a procedure for ranking alternatives on the ground that the rankings it produces are arbitrary, and the associated literature had already identified the phenomenon of rank reversal, in which introducing an additional alternative can reverse the relative ranking of two others that were unaffected by the addition.

Rank reversal is worth dwelling on because its practical implication is directly relevant to procurement. If adding a competitor to the evaluation can change which of the existing competitors ranks higher, then the composition of the shortlist influences the result independently of the merits of the shortlisted products. A buyer who invites a fourth vendor to make the process look more competitive, or who includes a vendor it has no intention of selecting in order to satisfy a policy requiring three quotes, may thereby alter which of the serious candidates wins. That is a disquieting property for a method whose appeal is its objectivity.

Fairness requires noting that the method's originators and their colleagues published a vigorous defence in the same journal issue, and that a substantial literature since has argued both that the criticisms can be addressed through careful application and that rank reversal is in some circumstances a legitimate reflection of how preferences actually work rather than a defect. The dispute is live and this article takes no position on its resolution. The relevant point for a buyer is narrower: the scorecard that appears to deliver an objective ranking rests on a technique whose properties are debated among the people who study it, and it should be treated as an aid to structured judgment rather than as an oracle that produces the answer. A selection team that understands this will use the scorecard to organise its reasoning and will notice when the arithmetic and the judgment diverge, which is exactly when the arithmetic most needs examining.

08 The demonstration the vendor controls

After the responses are scored, the shortlisted vendors demonstrate their products, and buyers frequently treat this as the stage where paper claims meet reality. It is not, because in the standard format every variable that determines what the buyer sees is controlled by the party being evaluated. Figure 6 sets out the distinction that matters.

Figure 6. The distinction that matters is control. A vendor-run demonstration on vendor data in a vendor environment following a vendor script establishes that the product can be made to perform. A buyer-controlled proof of concept establishes whether it performs here.

In a conventional vendor demonstration, the vendor selects the dataset, which is curated to be clean, complete, and dimensioned to flatter the product. The vendor writes the script, which traverses the workflows the product handles well and does not traverse the ones it handles badly. The vendor supplies the environment, which is configured and tuned by people who build such environments professionally. And the vendor provides the operator, a specialist who has performed this demonstration many times and who navigates the product with a fluency no new user will possess for months. What the buyer observes under these conditions is that the product can be made to perform impressively by an expert on favourable data, which was never in question and tells the buyer nothing about its own situation.

The alternative is a proof of concept in which the buyer controls the variables. The buyer supplies an extract of its own production data, including the parts that are messy, because the messy parts are where systems fail. The buyer writes the scenarios, and includes the exceptions, the awkward cases, the high-volume periods, and the workflows that the current system handles poorly, because those are the reasons the buyer is replacing it. The buyer's own users operate the product, so that the observed difficulty of use is the difficulty the organization will actually experience. And the test runs against the buyer's integration points wherever practical. This is more effort than watching a demonstration, and it is the only version of the exercise that produces information the buyer did not already have.

A practitioner formulation of this point is worth recording because it is sharper than most: a demonstration shows what the product does under the vendor's control using the vendor's data, whereas a proof of concept tests whether it works in the buyer's environment, and a proof of concept designed with the vendor's data in the vendor's sequence is simply a demonstration that the buyer paid for. The same source observes that a vendor who declines a buyer-controlled test has communicated something useful about its confidence. That inference should be drawn carefully, since there are legitimate reasons a vendor may resist, including cost, data protection, and the risk of being judged on an unrepresentative extract. But the resistance is itself information, and a buyer should notice which vendors welcome the test and which negotiate to narrow it.

Fairness requires a qualification that is frequently lost in the enthusiasm for proofs of concept. A buyer-scripted demonstration, in which the buyer specifies the scenarios and supplies representative data but the vendor still operates the product, captures much of the value at a fraction of the cost, and for many selections it is the proportionate choice. The failure mode this section identifies is not the demonstration format but the surrender of control over data, script, environment, and operator to the party being assessed. A buyer who retains control of the script and the data has fixed most of the problem even without a full proof of concept.

An adjacent practice deserves scrutiny alongside the demonstration, which is the reference check. Reference customers are supplied by the vendor, which means they are selected from the subset of the installed base that is willing to speak positively, and the conversation typically covers whether the reference is satisfied rather than what specifically went wrong and how it was resolved. A more informative approach is to ask the vendor for a customer that had a difficult implementation and to ask what happened, which vendors answer more often than buyers expect and which produces materially better information than a curated success story. Buyers should also ask references about the parts of the workflow the buyer knows are hardest in its own operation, rather than accepting a general account of satisfaction.

The deeper point is that every source of evidence in a selection is supplied by the party being evaluated unless the buyer arranges otherwise. The proposal is written by the vendor, the demonstration is run by the vendor, the references are chosen by the vendor, and frequently the requirements originated in vendor material. A buyer that has not deliberately introduced at least one independent source of evidence has conducted an assessment in which the assessed party controlled every input. Independent sources are available: the buyer's own data in a controlled test, customers found through the buyer's own network rather than the vendor's list, and the buyer's own users operating the product. Each is more effort than accepting what is offered, and each is the only kind of evidence that can disconfirm what the vendor has said.

09 The adviser who writes and then delivers

A structural conflict runs through much enterprise selection that public procurement regulates explicitly and private buying does not address at all: the party advising on the choice is frequently positioned to profit from it. Figure 7 sets out the shape.

Figure 7. An adviser who writes the requirements, designs the weighting, and runs the evaluation, and whose firm then implements whichever product is selected, has an interest in the outcome that the buyer should price.

The arrangement is common and its appeal is obvious. A buyer facing an unfamiliar category engages a consultancy with deep experience in it. That consultancy helps articulate requirements, designs the evaluation approach, facilitates the scoring, and supports the negotiation. It also has an implementation practice, frequently with certified capability in particular products, and it expects to bid for the implementation programme once the selection is complete. The buyer gets continuity, since the team that understands the requirements carries them into delivery, and it gets a single accountable partner. These are real benefits and they are why the arrangement persists.

The conflict is equally clear once stated. An adviser with an implementation practice concentrated in a particular product has a commercial interest in that product being selected, because its delivery capability, partner status, trained staff, and accumulated assets are worth more if the buyer chooses it. An adviser whose implementation revenue scales with programme complexity has an interest in a more complex selection than the buyer may need. And an adviser expecting to bid for delivery has an interest in a requirements set that its own delivery practice is well positioned to satisfy. None of this requires dishonesty; it requires only that the adviser, making judgment calls under uncertainty across dozens of small decisions, tends to resolve them in directions that happen to suit its position, which is how conflicts of interest operate in every professional field.

Public procurement addresses this through organizational conflict-of-interest rules that constrain a contractor who has helped prepare requirements from competing for the resulting work, on the reasoning that the party who defines the need should not also supply it. That reasoning applies with equal force to private buying, where it is almost never implemented. The practical remedy available to a private buyer is straightforward and is stated in the protocol later in this article: whoever advises on the selection does not bid for the implementation, and if the buyer wants the continuity benefit badly enough to accept the conflict, it should at least require the adviser to disclose its partner relationships, its certification concentrations, and its expected delivery interest in each candidate product, in writing, before requirements are drafted. A buyer that has this disclosure can weigh the advice appropriately. A buyer that has never asked has accepted advice of unknown provenance on the most consequential decision in the process.

A less obvious variant of the adviser conflict operates through the research and advisory firms whose comparative rankings buyers use to construct shortlists. These firms sell advisory subscriptions and related services both to buyers and to the vendors they assess, and vendors invest considerable effort in the briefing and submission processes that feed the rankings. This is disclosed, it is well known in the industry, and it does not establish that any particular assessment is compromised. It does mean that a buyer using a comparative ranking to determine which vendors are worth inviting is relying on an assessment produced by a firm with commercial relationships on both sides of the transaction, and that the ranking's inclusion criteria, which determine who appears at all, are set by that firm.

The practical guidance is not to ignore such rankings, which contain real analytical work and are frequently the most efficient starting point available, but to treat them as one input rather than as the definition of the field. A buyer whose shortlist is simply the upper region of a published ranking has outsourced the most consequential filtering decision in the process to a third party whose criteria it has not examined and whose commercial position it has not weighed. Reading the inclusion criteria, noticing which credible vendors are absent and why, and adding candidates the ranking excluded on criteria irrelevant to this buyer are all straightforward correctives that most buyers never apply.

10 The economics, and who publishes them

A reader who has followed the argument will reasonably want to know how large the phenomenon is, and here the evidence base becomes thin in a way that is itself part of the story. Figure 8 shows what exists.

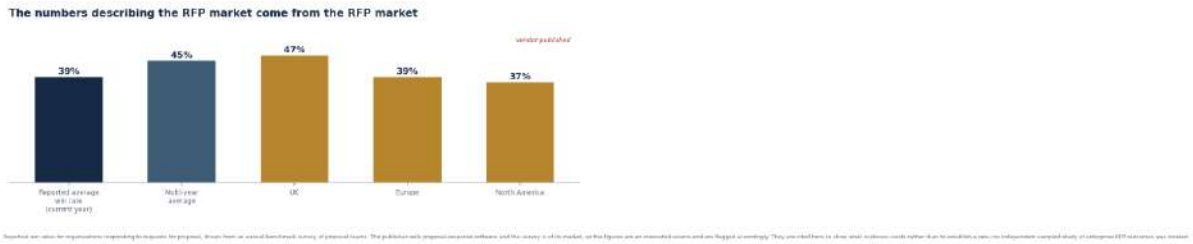

Figure 8. Reported win rates for organizations responding to requests for proposal, from an annual benchmark survey of proposal teams. The publisher sells proposal-response software and the survey is of its market, so the figures are an interested source.

The most substantial published data on request-for-proposal outcomes comes from annual benchmark surveys run by vendors of proposal-response software, conducted among proposal professionals and published as industry reports. The recent editions survey on the order of fifteen hundred organizations and report average win rates in the high thirties to mid forties percent, with regional variation, and they include figures on response volumes, response times, and the relationship between tooling and win rates. The work appears carefully done and the sample sizes are substantial. It is also, unavoidably, a survey of the customers and prospects of a company selling software to improve the outcomes it measures, conducted on a self-selecting population of organizations that respond to proposals frequently enough to have proposal professionals. It should be read as an interested source and it is flagged as one throughout this article.

What does not exist, so far as this research could establish, is an independent sampled study of enterprise software selection outcomes: how often incumbents win, how often the eventual winner was identified before the process began, how requirements are typically authored, or how frequently weights change after responses are opened. The absence is understandable, since private buyers have no obligation to disclose any of it and little incentive to invite scrutiny. It nonetheless means that the strongest quantitative claims circulating in this area, including striking figures about incumbent win rates relative to challengers, rest on practitioner assertion rather than measurement. One advisory firm states that challengers arriving only at the request-for-proposal stage win a low single-digit percentage of the time while incumbents win the large majority, and attributes the gap to requirement authorship rather than product merit. That is a plausible claim consistent with the mechanisms described in this article, made by an experienced practitioner, and it is not a study.

This publication's practice is to say so rather than to launder an assertion into a statistic by repetition, which is how most enterprise technology figures acquire their apparent authority. The honest position is that the mechanisms described in this article are documented, that public procurement law treats them as serious enough to prohibit, that the regulated domain shows evaluation defects occurring at a material rate under the strictest available conditions, and that nobody has measured how often they determine private enterprise outcomes because nobody can. A buyer does not actually need the

population statistic. It needs to know whether these mechanisms operated in its own last three selections, and that is a question it can answer from its own records in an afternoon.

The absence of independent measurement in this field has a further consequence worth naming, which is that it prevents buyers from learning collectively. In domains where outcomes are recorded and published, practice improves because participants can see which approaches produce better results and adjust. Enterprise software selection has no such feedback: each organization runs its processes, forms private impressions of how they went, and has no way to compare its approach against anyone else's on any objective measure. The result is that the practices described in this article persist not because anyone has established that they work but because nobody has established that they do not, and the same avoidable errors are rediscovered independently by organization after organization.

11 The fairness case: the process earns its keep

This article has been critical of a process that is close to universal, and fairness requires setting out the substantial case in its favour, because a reader who concluded that competitive selection should be abandoned would be adopting a position worse than the one criticised here.

The strongest argument for the formal process is that its alternative is not a better process but no process. Where structured competitive selection is absent, enterprise software is bought through relationship, familiarity, and executive preference, and those channels have failure modes considerably worse than the ones described here. A structured process forces the organization to articulate what it needs, exposes the decision to more than one perspective, creates a record that can be examined afterward, and gives internal stakeholders a legitimate route to influence a choice that will affect their work for a decade. Buyers who have watched a major system selected because a senior executive met a vendor at a conference will not be tempted to romanticise the informal alternative.

The second argument is that the process delivers auditability and defensibility that have independent value. In regulated industries, in public bodies, and in any organization whose decisions may be scrutinised by auditors, boards, or courts, the ability to demonstrate that a significant expenditure followed a documented and reasoned procedure is worth real money, quite apart from whether it improved the choice. A selection that produced the right answer by an undocumented route is a liability in a way that the same answer reached through a recorded process is not.

The third argument comes from the data this article has relied on. The federal effectiveness rate of roughly half, presented earlier as evidence that evaluation defects are common, is equally evidence that the correction mechanism works: agencies do take corrective action when challenged, and the independent forum does sustain protests where the record warrants. A system in which half of challenges yield relief is a system in which challenge is worthwhile, which is a functioning accountability loop rather than a broken one. The regulated domain looks flawed compared to an ideal and it looks excellent compared to a private process with no challenge mechanism whatsoever.

A fourth point deserves emphasis because this article's central mechanism can be read too broadly. Requirements that narrow the field are not inherently improper, and there are entirely legitimate reasons for a specification to exclude most of the market: genuine interoperability constraints, an existing technology standard the organization has committed to, regulatory obligations, safety requirements, or a deliberate strategic decision to consolidate on a platform already in place. A buyer with a considered reason to restrict the field should restrict it and record why. The pathology this article describes is not narrow requirements but unexamined ones, where the narrowing was inherited from a vendor document and nobody involved can articulate what operational need it serves. The remedy is not wider requirements; it is requirements whose authorship and rationale the buyer can explain.

A fifth point in fairness concerns the position of vendors, who are the parties bearing much of the cost of the practices described here and who are frequently cast as their beneficiaries. A vendor invited into a selection it cannot win because the requirements were written around a competitor incurs substantial cost for nothing: proposal teams, subject matter experts, demonstration preparation, travel, and legal

review, spent on an outcome that was determined before the invitation arrived. Vendors are generally aware when this is happening, since experienced sales organizations recognise a specification written around a rival, and they participate anyway because declining an invitation carries its own costs with the buyer and in the market. The practice therefore imposes a recurring tax on the losing side of every staged competition.

This has a consequence buyers should care about, which is adverse selection in the field. Vendors triage their opportunities, and the ones with the strongest positions are the most able to decline processes they judge to be predetermined. Over time a buyer known for staged competitions attracts responses from vendors with spare capacity and loses the attention of those in high demand, so the field that assembles is systematically weaker than the market. The buyer never observes this, because it sees the responses it received rather than the ones it did not, and it may conclude that the market offers less than it does. Running genuine competitions is therefore not only a matter of fairness to vendors but a matter of the buyer's own access to the best available options.

12 Move the rigour to the requirement

The constructive principle follows directly from the diagnosis and can be stated in one line: move the rigour from the evaluation to the requirement, because that is where the decision is made. Everything in the protocol that follows is an application of this single idea.

In practice the reallocation begins with how requirements are generated. The reliable method is to derive them from the organization's own processes rather than from any external document: walk the actual workflow, identify what the organization does, what it needs the system to support, where the current arrangement fails, and what would constitute an improvement, and write requirements that describe those needs in terms of outcomes rather than mechanisms. A requirement that states what the business must be able to accomplish permits any product that accomplishes it, whereas a requirement that states how the system must work permits only products built that way. The first is a specification of need and the second is a specification of a product, and the discipline is to keep every requirement in the first category unless there is a recorded reason for the second.

The second element is provenance. Every requirement should be traceable to a business reason that someone in the organization can articulate, and any requirement whose origin cannot be explained should be struck. This sounds bureaucratic and it takes less time than expected, because most requirement sets contain a large proportion of lines that nobody can defend once asked, having arrived from a template or a prior exercise. Running that filter routinely removes a substantial share of the constraints on the field, and each removed constraint restores a competitor to viability. A buyer who has struck every requirement that no one can justify has done more to open the competition than any amount of care in scoring will achieve.

The third element is the treatment of the weighting as a first-class decision. The weights should be derived from the same business reasoning as the requirements, set and documented before any response is opened, approved at the level appropriate to a decision that can reverse the outcome, published to respondents so they can address what matters, and left unchanged once scoring begins. If circumstances truly require a change, the change and its reasons should be recorded before the revised scores are calculated, so that the sequence is visible afterward. None of this is difficult. It is simply unusual, because the weighting has been treated as a mechanical detail of a process whose real work was assumed to happen elsewhere, and the argument of this article is that the assumption is backwards.

A practical objection to process-derived requirements deserves an answer, since it is the most common reason buyers reach for templates. The objection is that an organization purchasing in an unfamiliar category does not know what is possible, and that writing requirements purely from its own current processes risks specifying an improved version of what it already does rather than what the market could offer. This is a real risk and the answer is sequencing rather than abandonment. The organization should first document its own processes, needs, and failure points in its own language, and only then survey what the market offers, using that survey to identify possibilities it had not considered and to revise its statement of need.

The order matters because it determines which document anchors the other. An organization that surveys the market first and writes requirements second produces requirements shaped by the market's categories. An organization that documents its needs first and surveys second has a stable reference against which to evaluate what it learns, and can distinguish a capability it truly wants from one that merely sounds impressive in a demonstration. The market survey is valuable in both sequences; only in the second does the organization retain ownership of its own definition of the problem, which is the property that keeps the eventual competition open.

13 A selection protocol, and a scoring rubric

The principles above combine into a protocol a buyer can adopt, and a rubric a governance body can use to judge whether a selection is deciding anything. Figure 9 sets out the discipline.

Figure 9. Running a selection that is actually a selection: write requirements from process, fix the weights before opening responses, separate adviser from implementer, test on your own data, and record the reasoning rather than only the result.

The protocol runs as follows. Derive requirements from the organization's own processes and express them as outcomes rather than mechanisms, striking any requirement whose business rationale cannot be stated. Set and publish the weighting before any response is opened, approve it at a level commensurate with its influence, and do not revise it once scoring begins. Require any adviser to disclose its partner relationships and delivery interests before drafting begins, and prefer an adviser that will not bid for the implementation. Replace or supplement vendor-controlled demonstrations with tests the buyer controls, using its own data including the difficult parts, its own exception scenarios, and its own users. And document the reasoning behind the award rather than only the arithmetic, so the decision can be examined later by people who were not present.

A scoring rubric

The dimensions below distinguish a selection that decides from one that documents.

Dimension	A real selection	Procedural theatre
Requirement source	Derived from the buyer's own processes	Template or vendor capability list
Requirement form	Outcomes the business must achieve	Mechanisms one product happens to use
Provenance	Every line traceable to a stated need	Inherited lines nobody can justify
Weighting	Fixed and published before opening	Set late, adjusted after scores are visible
Adviser	Interests disclosed; does not bid to deliver	Writes requirements, then implements
Demonstration	Buyer's data, scenarios, users, exceptions	Vendor script on vendor data
Record	Reasoning documented, not only scores	A scorecard and an award letter

A selection scoring in the left column can be examined afterward by someone who was not involved, and the examination will show why the winner won. A selection scoring in the right column produces a defensible file and an outcome that was substantially determined before the competition opened. The rubric does not make selection easy, and it does not guarantee a better product. It ensures that the process is doing the work it appears to be doing, which is the minimum a buyer should require of an exercise that consumes months and determines a decade.

14 Buying without a competition, deliberately

A conclusion that follows from this analysis, and that most buyers resist, is that some purchases should not be competed at all, and that running a competition to legitimise a decision already made is worse than not running one. The alternative to procedural theatre is not always a better competition; sometimes it is an honest sole-source decision.

Consider the situations in which the outcome is truly predetermined by circumstances rather than by manipulation. An organization deeply committed to a platform, whose staff are trained on it, whose integrations are built for it, and whose adjacent modules come from the same supplier, may reasonably conclude that extending that platform is the right answer before examining alternatives, because the switching costs and integration advantages are decisive and known. An organization with a regulatory or interoperability constraint that only one or two products satisfy is in a similar position. In these cases a competition serves no decision-making function, because the decision follows from the constraints, and running one imposes real costs on the buyer and on the vendors invited to lose.

Those costs deserve to be named, since they are frequently treated as free. Each invited vendor commits substantial professional effort to responding, preparing demonstrations, and attending meetings, and where the outcome was never in question that effort is a transfer from the vendor to the appearance of the buyer's diligence. The buyer bears its own cost in months of staff time. And there is a reputational cost that accumulates: vendors compare notes, they form views about which buyers run genuine processes, and a buyer known for staged competitions will find that serious competitors decline to participate, which degrades the quality of the field in the selections where the buyer actually wants a contest.

The honest alternative is a documented sole-source decision, which public procurement provides for explicitly through justification and approval requirements and which private buyers can adopt without any regulatory prompting. The buyer records why this product is the answer, what alternatives were considered and why they were rejected, what the constraints are that make the choice near-inevitable, and what commercial terms were negotiated in the absence of competitive tension, which is the real risk of sole sourcing and should be addressed directly through benchmarking rather than through a fictitious competition. This produces a better record than a staged competition, because it states the actual reasoning rather than concealing it behind a scorecard, and it is more defensible on review for exactly that reason. A buyer that reserves competitive selection for decisions that are actually open, and documents sole-source reasoning where they are not, will run fewer competitions, run them better, and be taken more seriously by the market when it does.

There is a further category worth distinguishing from both the open competition and the honest sole source, which is the structured trial. Where a buyer is uncertain between two or three credible options and the differences that matter cannot be established from documents, the most informative approach is frequently to run a limited paid engagement with more than one candidate, scoped to a real piece of work, and to compare what happens. This costs more than a demonstration and less than a full implementation, and it produces evidence about the qualities that determine long-term satisfaction,

including how the vendor behaves when something does not work, how quickly its people understand the buyer's operation, and whether its product handles the buyer's awkward cases.

The structured trial addresses the central weakness of the conventional process, which is that it assesses vendors on their performance in a sales context rather than in a working relationship. Every input a buyer receives during a selection is produced by people whose function is to win the selection, and those people are frequently not the ones the buyer will work with afterward. A paid trial engages delivery staff on real work under realistic constraints, which is the closest approximation to the eventual relationship that can be obtained before committing to it. Buyers who can afford the approach report that it changes their ranking more often than any other single input, which is a strong indication that it is supplying information the rest of the process does not.

15 Conclusion: decide first, then document

The request for proposal has become the standard instrument of enterprise software selection because it appears to answer the fundamental problem of buying complex systems under uncertainty: it substitutes a documented, criteria-based, multi-party assessment for the fallible judgment of individuals. That substitution is worth something, and this article has argued that it delivers less than it appears to, because the assessment operates on a field that was determined before the assessment began, using weights that were set before any evidence arrived, informed by demonstrations that the assessed parties controlled, and frequently advised by a party positioned to profit from the answer.

The evidence is necessarily indirect, since private buying leaves no public record, and this article has been explicit about that limitation rather than filling it with confident statistics. What can be established is that public procurement law identifies requirement authorship, brand-name specification, adviser conflict, and the absence of external review as the mechanisms capable of subverting a competition, and prohibits or constrains each of them; that in the regulated domain, where those constraints apply and an independent forum reviews challenges, roughly half of protests yield some relief and the most common grounds concern the evaluation itself; that the scoring technique underlying the scorecard has been contested in the literature for decades; and that the only substantial published figures on proposal outcomes come from companies selling proposal software. Private enterprise buying operates without the constraints, without the forum, and without the record.

What follows is not that competitions should be abandoned. It is that the diligence should be moved to where the decision is made. Requirements should be derived from the organization's own processes and expressed as outcomes, with every line traceable to a need someone can articulate and every unjustifiable line struck. Weights should be treated as the consequential decision they are: derived from business reasoning, fixed and published before responses are opened, approved at an appropriate level, and left alone. Advisers should disclose their interests and, where possible, should not bid to deliver what they helped specify. Demonstrations should be replaced or supplemented by tests the buyer controls, on the buyer's own data including the parts that are difficult. And where the answer is truly determined by constraints rather than by comparison, the buyer should say so and document the reasoning rather than staging a contest whose result is known.

The organizations that select well are not the ones with the most elaborate scorecards. They are the ones that can explain, months afterward and to someone who was not in the room, why the requirements said what they said, why the weights were set where they were set, who wrote each and on what basis, and what the winning product demonstrated on the organization's own data. That explanation is the entire substance of a selection, and it is available to any buyer willing to produce it. The alternative is a process that produces a file rather than a decision, and a decade of living with a system that was chosen, in the part of the process nobody examined, by someone with an interest in the answer.

16 Methodology, caveats, and sources

Methodology

- This article draws on primary regulatory and legislative sources, published bid-protest reporting, the peer-reviewed multi-criteria decision literature, and vendor and advisory materials, current to mid-2026. Supply Chain Research is independent and accepts no payment from the research firms, advisers, or software vendors discussed.
- Because private enterprise selection generates no public record, the argument relies on documented mechanisms and on the regulated procurement domain, where the same practices are constrained and reviewed. The limits of that inference are stated in the text rather than concealed.

Caveats

- No independent sampled study of private enterprise request-for-proposal outcomes was located. The published win-rate and volume figures come from vendors of proposal-response software surveying their own market, and are flagged as interested throughout.
- Figures on incumbent and challenger win rates are practitioner assertions from an advisory firm, not sampled research. They are reported as assertions and should not be treated as measurements.
- Bid-protest data describe United States federal procurement and cannot be extrapolated to private buying. They are used to establish that evaluation defects occur at a material rate under the strictest available conditions, not to quantify private practice.
- The methodological critique of analytic hierarchy approaches is contested. The originators published a defence in the same journal issue, and a substantial subsequent literature argues both sides. This article takes no position on the resolution.
- Figures 1, 3, 4, 6, 7, and 9 are conceptual illustrations of structure or arithmetic rather than measured data, and are labelled as such. Figure 4 in particular uses invented scores to demonstrate a property of weighted scoring.
- Regulatory provisions are summarised rather than quoted, and are current as described at the time of writing. Buyers subject to procurement law should consult the current text and their own counsel rather than relying on this summary.

Sources

[1] US Government Accountability Office. Bid Protest Annual Report to Congress for Fiscal Year 2024.

[2] US Government Accountability Office. Bid Protests: Key Features and Trends.

[3] European Union. Directive 2014/24/EU on public procurement, including Article 42 on technical specifications.

[4] Court of Justice of the European Union. Case C-424/23, on specifications with the effect of excluding competing products.

[5] US Federal Acquisition Regulation. FAR 11.104, brand name or equal purchase descriptions.

[6] US Federal Acquisition Regulation. FAR 6.302-1, only one responsible source.

[7] Dyer, J.S. Remarks on the Analytic Hierarchy Process, Management Science 36(3), 1990.

[8] Loopio. RFP statistics and win rates, annual benchmark survey (interested source: vendor of proposal-response software).

Additional context drawn from published analyst toolkits and requirement templates marketed to buyers, from advisory commentary on requirement authorship and incumbent advantage, and from practitioner material on the distinction between vendor demonstrations and buyer-controlled proofs of concept. Sources with a commercial interest in the matters they describe are identified as such. This article is analysis, not legal or procurement advice, and its conclusions should be validated against your own circumstances and obligations before any decision.
