

Benchmarking Application Development Contracts



Introduction

The ISBSG repository contains a vast array of project data and includes data for projects undertaken in an agile way of working. This enables analysis of the differences between traditional projects and agile projects.

ISBSG collects industry data, where output is measured using ISO/IEC standardized and therefore objective, repeatable, auditable methods, such as Nesma, IFPUG and COSMIC function points.

Typical key metrics based on function points are:

- Project Delivery Rate (PDR)¹: Hours spent per function point
- Cost efficiency: Cost (or Price) per function point
- Quality: Defects per function point (in test and/or 1st month of production)
- Delivery Speed: Function points delivered per calendar month.

The ISBSG 'New Development & Enhancement' repository contains thousands of completed projects for which these metrics are calculated. This allows organizations to make better decisions based on facts instead of opinions.

In this short report we'll dive into the benefits of using functional size and ISBSG data to benchmark IT Contracts. There will be a particular focus on the Application Development phase, which usually costs a large portion of the allocated IT budget!

Why Benchmarking IT Contracts?

In the current global IT landscape, large enterprises increasingly rely on long-term outsourcing arrangements with major system integrators. These contracts, often spanning several years and valued at hundreds of millions of dollars, are intended to provide both scalability and access to specialized expertise. However, a systemic challenge arises in the governance of these relationships: as contract durations extend, the initial competitive alignment of the agreement inevitably begins to drift from prevailing market realities.

¹ The PDR is the inverse of the universal concept of Productivity (output/input) as it is easier to process for human minds, which usually struggles with metrics with many decimals

For infrastructure-oriented services—such as wide-area networking, server management, and end-user computing—this drift is relatively straightforward to mitigate. Benchmarking firms can leverage standardized, quantifiable KPIs, such as "cost per server" or "cost per unit of workplace," to conduct comparative analyses against established market data. These metrics provide a transparent mechanism to verify whether the cost structure remains commercially viable or if an adjustment is required.

The situation becomes significantly more complex with application development and maintenance services. While auditing annual rate card adjustments is routine, this provides only a partial view of value. Vendors typically index their hourly rates upward annually, leading to increased expenditure. Yet the actual output—the tangible functionality delivered for that investment—remains a "black box." Without a standardized methodology to measure the volume of delivered functionality, cost increases are often conflated with productivity or value improvements.

Adopting Functional Size Measurement (FSM)—such as Nesma, IFPUG, or COSMIC FPA—is the essential prerequisite to breaking this black box. By transitioning the conversation from "cost per hour" to "cost per unit of functionality," organizations can finally apply the same rigorous benchmarking standards to application services that they currently apply to infrastructure, ensuring that the value received remains commensurate with the market-aligned investment.

The Limitations of Output-Obscurity in Application Services

The core difficulty in benchmarking application services lies in the fundamental disconnect between financial inputs and operational outputs. In infrastructure management, the service is defined by tangible assets (servers, endpoints, network bandwidth), making the unit of measurement clear and directly correlated to cost. Conversely, application services are often procured under time-and-material or capacity-based models that prioritize effort over output. While indexation clauses allow vendors to raise hourly rates annually—effectively increasing the contract's total cost of ownership—the lack of an objective, size-based metric for the software being developed masks the true productivity of the delivery team.

In this environment, the customer is often left assessing the "billable hours" rather than the "value created." When the delivered software volume is not measured, it becomes impossible to determine if a rising bill is the result of increasing scope, technical debt, or a degradation in delivery efficiency. The vendor's development process remains a black box where time is consumed, but the resulting functional volume is never quantified. Without a transparent measurement of output, the customer loses the ability to distinguish between cost-effective innovation and inefficient delivery, effectively rendering the contract's benchmarking clause impotent for the largest portion of their IT budget.

Only by quantifying the work in standardized units can organizations transition from subjective cost auditing to objective value-based performance management.

The Governance Vacuum: Why the Black Box Persists

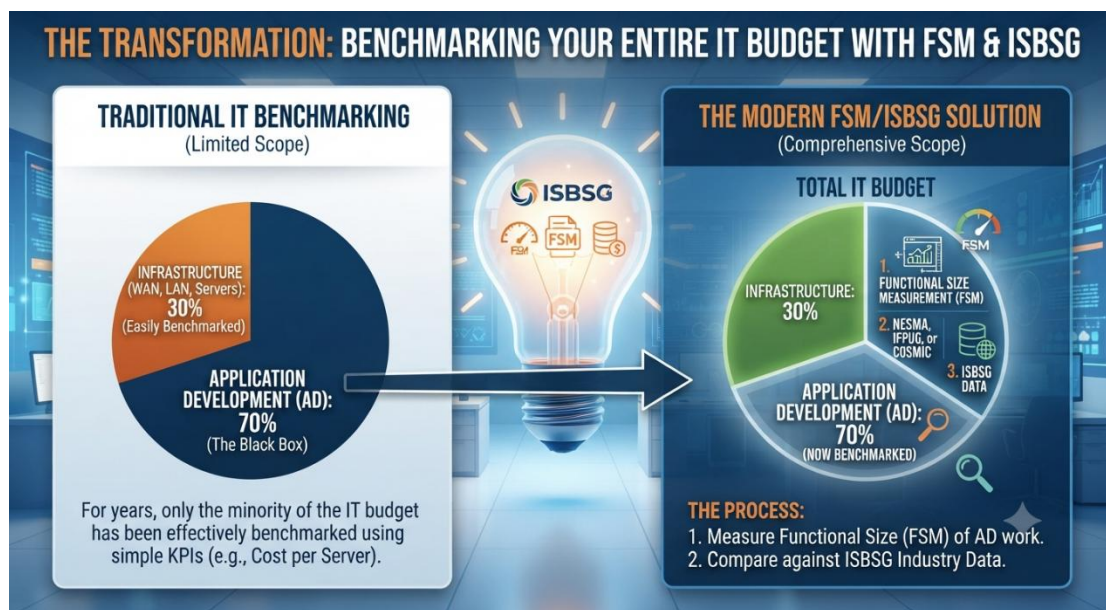
The paradox of massive IT outsourcing expenditure is that while these contracts represent hundreds of millions of dollars, they frequently operate in a state of accountability paralysis. Both parties have historically been incentivized to maintain the status quo, effectively leaving the delivery of software functionality unmanaged and unmeasured.

From the customer's perspective, the inability to measure functional output creates a profound governance vacuum. Enterprise procurement and contract management teams are often equipped to audit financial compliance—ensuring that the billed rates match the contract's indexed schedule—but they lack the technical visibility to challenge the efficiency of the underlying development. Because they cannot quantify the "what" (the functionality delivered), they are forced to rely entirely on the vendor's reported "how" (the hours spent). This asymmetry of information means the customer remains a passive observer, paying for the accumulation of effort rather than the realization of business value.

The vendor, conversely, operates under a set of incentives that are often diametrically opposed to transparency. In time-and-material or capacity-based contracts, every hour spent is an hour invoiced. When the vendor is not contractually obligated to report functional output, they have little reason to prioritize the "efficiency" of delivery. In fact, under this model, increasing productivity—delivering more functionality in fewer hours—can actually be counterproductive to their revenue goals.

This alignment of perverse incentives means that neither party seeks to open the black box. The vendor is satisfied with the predictable revenue stream of billable hours, and the customer, lacking the tools to prove otherwise, treats the rising cost of software as an unavoidable cost of doing business. This systemic lack of ownership turns the contract into a financial commitment that drifts further from market conformity with every passing year, unchecked by any objective performance standard.

Benchmarking the Vendor: A High-Level Process



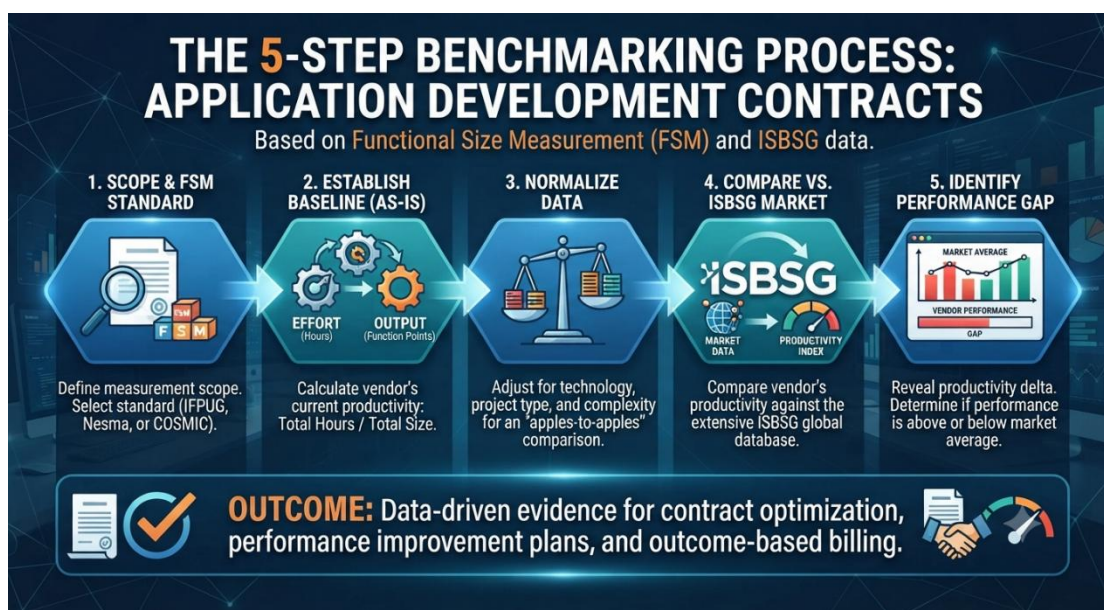
To effectively benchmark an Application Development (AD) contract and move beyond simple rate-card analysis, organizations must shift to an objective, evidence-based approach. The benchmarking process requires a structured methodology that transforms "black box" expenditures into clear productivity metrics that can be compared against industry standards.

The process follows a logical flow:

- **Define the Scope and Measurement Standard:** Establish a clear foundation by selecting a recognized Functional Size Measurement (FSM) standard—Nesma, IFPUG, or COSMIC FPA—to define what constitutes a "unit of work." Consistency is paramount; the same rules must be applied to the current contract as those used to populate the benchmarking database.
- **Establish the Baseline (The "As-Is" State):** Aggregate the actual functional output of the vendor's recent delivery cycles. By measuring the realized functionality (e.g., in Function Points) against the total hours consumed, the organization calculates the current delivery productivity. This is the crucial step of turning "effort" into "output."
- **Normalize the Data:** Not all projects are comparable. Data must be normalized to account for factors that influence productivity, such as the programming language, the type of application (new development vs. enhancement), and the complexity of the environment. This ensures an "apples-to-apples" comparison against the market.
- **Leverage ISBSG Data for Market Comparison:** Once the vendor's performance is quantified and normalized, it is compared against the

extensive, industry-standard data repository provided by the ISBSG. By benchmarking the organization's current productivity (hours per unit of function) against the performance of similar projects within the ISBSG dataset, the organization can objectively determine if the vendor's delivery efficiency is in line with the current market reality.

- **Identify the Performance Gap:** The final analysis reveals the "productivity delta". If the vendor requires significantly more hours to deliver the same functional size compared to the market average, the benchmarking result provides the objective evidence needed to initiate performance improvement plans, renegotiate contract terms, or adjust the vendor's delivery model.



This structured process effectively removes the ambiguity from the relationship. It changes the conversation from a subjective disagreement over "invoiced hours" to a data-driven dialogue about "delivery performance," empowering the customer to hold the vendor accountable for market-aligned productivity.

The Role of ISBSG Data as a Neutral Benchmark

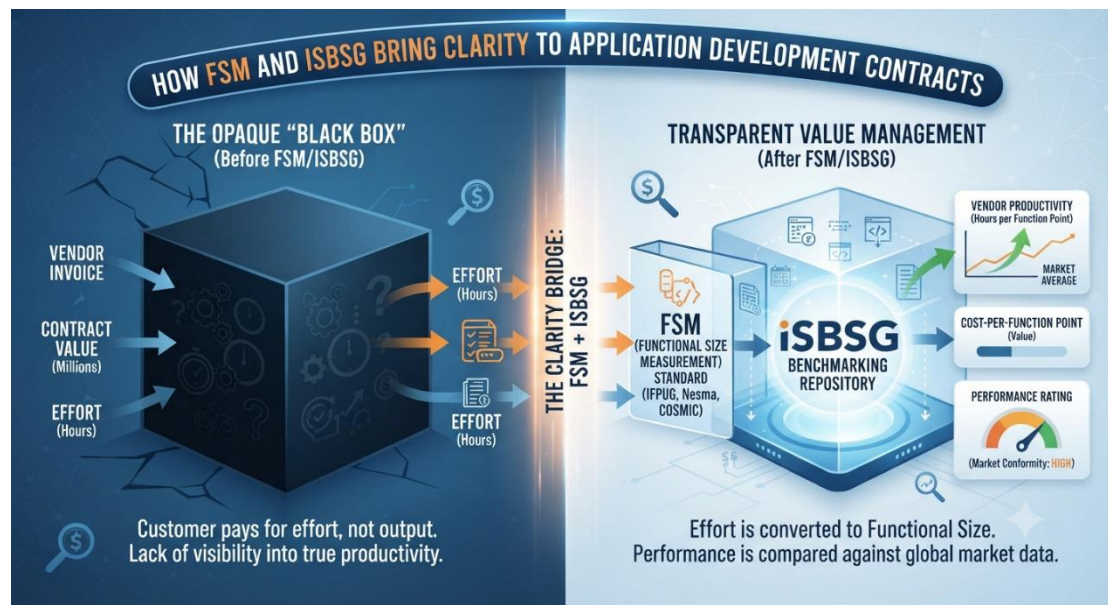
The effectiveness of this benchmarking process hinges entirely on the quality and objectivity of the comparison data. In a high-stakes contract dispute or periodic review, neither the vendor's internal logs nor the client's own historical data—which may be biased by legacy processes—are sufficient. The ISBSG data repository provides the necessary third-party, industry-recognized "ground truth."

Because ISBSG data is derived from thousands of projects across diverse industries, technologies, and geographic regions, it offers a robust statistical baseline. When a client benchmarks their vendor using ISBSG, they are not relying on a proprietary, "black box" algorithm from a specific consulting firm. Instead, they are measuring

performance against a transparent, peer-reviewed global standard. This neutrality is the ultimate defence against vendor pushback; it is difficult to refute a performance assessment that is backed by an objective, aggregate industry average.

By utilizing this external benchmark, the client can precisely calculate the **Productivity Index**, revealing the exact percentage deviation of the vendor's performance compared to the market. This creates a powerful, indisputable basis for contract optimization. When the data shows that a vendor is consistently delivering below the market-average productivity, it provides a clear mandate for corrective action—whether that involves revising delivery processes, demanding efficiency gains, or restructuring the contract to shift from effort-based billing to outcome-based or fixed-price arrangements.

Ultimately, the integration of FSM with ISBSG data transforms the benchmarking clause from a passive legal requirement into a dynamic, strategic tool for ensuring that investment in application development translates directly into realized business value.



Conclusion: From Cost Management to Value Realization

The reliance on time-and-material or capacity-based models has allowed the true productivity of application development to remain obscured within a "black box" for far too long. As long as IT contracts are governed solely by rate cards and indexed hourly billing, organizations will continue to pay for effort rather than value, effectively ceding control over their largest IT investments. The path to rectifying this imbalance lies in the adoption of Functional Size Measurement and the systematic use of objective, industry-recognized benchmarks.

By quantifying the output of development teams—regardless of the vendor or the project—and comparing these metrics against the verified, global repository of

ISBSG data, clients can finally break the cycle of uncontrolled cost escalation. This transition is not merely a technical exercise; it is a critical shift toward mature IT governance. When organizations replace subjective assumptions with the empirical evidence provided by ISBSG benchmarks, they transform their outsourcing contracts from stagnant financial obligations into high-performing partnerships, ensuring that every dollar spent is aligned with competitive, market-standard productivity.

If you wish to do your own analysis, or if you are interested to use the ISBSG data for cost estimation, benchmarking, performance measurement, procurement, etc., please subscribe to the data here: <https://www.isbsg.org/project-data/>

The International Software Benchmarking Standards Group (ISBSG)

The ISBSG is a not-for-profit organization founded in 1997 by a group of National Software Metrics Associations. Their aim was to promote the use of IT industry data to improve software processes and products.

ISBSG is an independent international organization that collects and provides industry data for software development projects and maintenance & support activities. It aims to help all organizations (commercial and government, suppliers and customers) in the software industry to understand and to improve their performance and decision making.

ISBSG sets the standards of software data collection, software data analysis and software project benchmarking processes and is considered to be the international thought leader in these practices.

The ISBSG mission is to support commercial and public organizations to improve the estimation, planning, control and management of IT software projects and/or maintenance and support contracts.

To achieve this, ISBSG maintains and grows 2 repositories of IT software development/maintenance & support data. This data originates from trusted, international IT organizations and can be obtained for a modest fee from the website www.isbsg.org/project-data/

Help us to collect data

ISBSG is always looking for new data. In return for your data submission, we issue a free benchmark report that shows the performance in your project or contract against relevant industry peers.

Please submit your data through one of the forms listed on <http://isbsg.org/submit-data/>

A specific Agile/Scrum data collections questionnaire can be downloaded here:

<https://cutt.ly/4vnuXVT>

Partners

This page will help you to find an ISBSG partner in your country:

<https://www.isbsg.org/board/>