

# How ISBSG Data Supports Output-Based Contracting



## Introduction

The ISBSG repository includes data for projects carried out in an agile way of working. This is useful since the analysis of differences between traditional projects and agile projects is valuable.

ISBSG collects industry data where output is measured using ISO/IEC standardized and therefore objective, repeatable, auditable methods. These include Nesma, IFPUG and COSMIC function point counting methods.

Typical key metrics based on function points are:

- Project Delivery Rate (PDR): 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 enables organizations to use this industry data for fact-based understanding and decision making.

**In this short report we'll dive into the benefits of using functional size and standardized metrics with ISBSG data. This is useful in 'fixing' 'Time and Materials' contracts – a bad practice used by agile software development teams.**

## The industry 'bad practice': Hourly Rates Contracting (T&M)

In the current landscape of agile application development, the reliance on input-based Time & Materials (T&M) contracts has become an industry "bad practice". It forces customers to bear project risks while failing to provide an objective oversight of value delivery. This creates a significant transparency gap where budgets are known but the actual value delivered remains obscured. This can lead to misaligned incentives for both customer and supplier organizations (see Figure 1).

To bridge this gap, organizations are increasingly looking towards output-based contracting models that utilize objective, standardized metrics. By leveraging industry-recognized data from the International Software Benchmarking Standards Group (ISBSG) alongside frameworks such as the Dutch practice guideline NPR 5333 (output-based measurement, performance management and contracting), and Easy Functional Sizing (EFS), stakeholders can establish a "win-win" model based on transparency, fair risk-sharing, and clear, data-driven performance management.

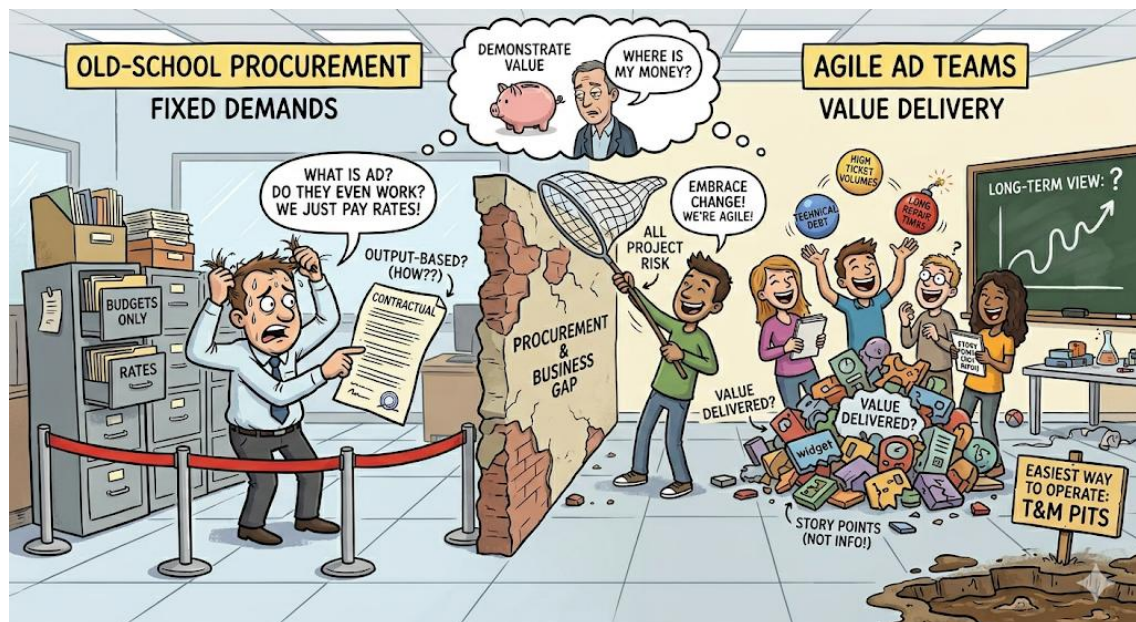


Figure 1: Challenges in determining value delivered from Agile Teams

### Why T&M Contracts fail

Time & Materials (T&M) contracts frequently fail because they misalign the fundamental incentives between customers and suppliers. They tether compensation to effort rather than outcomes. The supplier is financially rewarded for the time spent on development. This inadvertently penalizes efficiency and creates a disincentive for productivity improvements, as faster delivery directly reduces the supplier's revenue.

This dynamic shifts the focus from delivering "value for money" to mere "resource consumption," often resulting in bloated budgets and scope creep. An adversarial relationship may also occur where the client is forced to micromanage progress to ensure the developer's time is being used effectively, rather than focusing on the strategic goals of the software product.

Many organizations benchmark the hourly rates they pay for certain functions, but this only shows what the market pays per hour for certain roles, not what is delivered in that role. See hourly rates based on years of experience in Figure 2 below.

Country	Customer Role	Skill level	Years of Experience	Hourly Rate (EUR/Hour)		
				Peer Low	Peer Avg	Peer High
France	Program Sponsor	N.a.	N.a.	€ 277,93	€ 322,64	€ 362,11
France	Program Director	N.a.	N.a.	€ 299,30	€ 362,45	€ 413,69
France	Senior Managing Director	Senior	>10 years	€ 261,72	€ 326,96	€ 373,74
France	Managing Director	Confirmed	4-10 years	€ 239,11	€ 277,01	€ 312,52
France	Subject Matter Expert	Senior	>10 years	€ 129,26	€ 147,66	€ 166,70
France	Program Architect	Senior	>10 years	€ 156,80	€ 181,80	€ 204,30
France	Solution Architect	Confirmed	4-10 years	€ 115,60	€ 133,50	€ 153,20
France	Solution Architect	Senior	>10 years	€ 139,70	€ 160,10	€ 180,40
France	Project Manager	Junior	0-2 years	€ 76,00	€ 87,80	€ 100,10
France	Project Manager	Advanced	2-4 years	€ 104,20	€ 119,50	€ 134,60
France	Project Manager	Confirmed	4-10 years	€ 129,59	€ 150,63	€ 177,23
France	Project Manager	Senior	>10 years	€ 142,34	€ 167,00	€ 196,75
France	Program Manager	Confirmed	4-10 years	€ 155,90	€ 180,50	€ 203,80
France	Program Manager	Senior	>10 years	€ 188,70	€ 218,30	€ 246,30
France	Service Delivery Manager	Confirmed	4-10 years	€ 120,30	€ 140,10	€ 157,80
France	Service Delivery Manager	Senior	>10 years	€ 139,40	€ 160,20	€ 183,30
France	Scrum master	Advanced	2-4 years	€ 86,30	€ 98,20	€ 127,70
France	Scrum master	Confirmed	4-10 years	€ 94,60	€ 120,50	€ 148,90
France	Scrum master	Senior	>10 years	€ 112,78	€ 145,82	€ 180,12
France	Proxy Product Owner	Junior	0-2 years	€ 71,50	€ 81,90	€ 92,50
France	Proxy Product Owner	Advanced	2-4 years	€ 106,90	€ 124,40	€ 140,80
France	Proxy Product Owner	Confirmed	4-10 years	€ 135,00	€ 155,70	€ 179,20
France	Proxy Product Owner	Senior	>10 years	€ 167,50	€ 193,30	€ 221,10
France	Project Manager	Junior	0-2 years	€ 76,00	€ 87,80	€ 100,10
France	PMO	Advanced	2-4 years	€ 78,50	€ 90,10	€ 102,90

**Figure 2: An example of a ratecard benchmark.**

Trying to procure the lowest rate may backfire. You may use staff the supplier can't deploy for higher rates, because they lack experience, knowledge and/or skills (see Figure 3).



Figure 3: Low rates can equate to inexperience and skills

## Output-based contracting

The persistent shift toward agile development has often paradoxically reinforced an reliance on Time & Materials (T&M) contracting. This fundamentally treats software delivery as a commodity of effort rather than an achievement of business outcomes. The pressure to optimize delivery speed or simplify architecture is neutralized, often leading to "agile theater" where teams log hours under the guise of productivity while missing critical performance targets.

The customer becomes responsible for the financial consequences of scope creep, low velocity, or poor quality, with no contractual mechanism to tie payment to the actual functionality delivered. This misalignment creates a fragile relationship defined by uncertainty and constant oversight, as the client must audit the process itself—checking velocity, story points, and team composition—rather than simply verifying the value of the completed product.

In a Time & Materials (T&M) model, vendors are prevented from optimizing their performance because their primary revenue stream is directly tethered to the consumption of hours rather than the delivery of functional outcomes. Under this contractual structure, increases in productivity—such as implementing automation, simplifying architectural requirements, or refining team workflows—acts as a direct financial penalty to the vendor by reducing the total billable hours (see Figure 4).

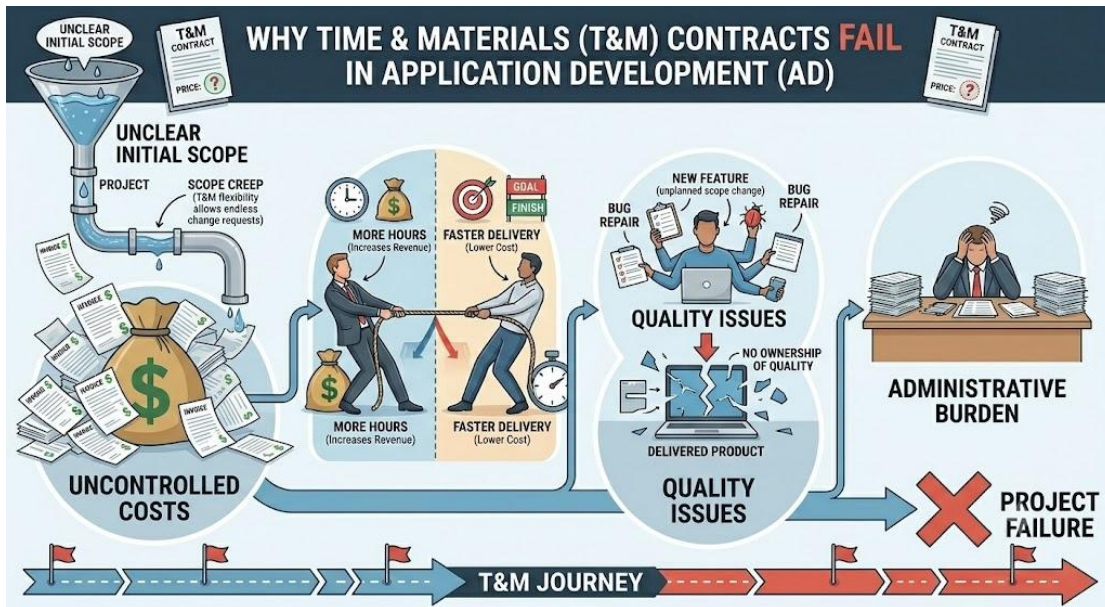


Figure 4: Why T&M contracts fail in application development

Consequently, the T&M model fosters a hidden incentive to maintain or even prolong current levels of effort, effectively decoupling the supplier’s business success from the client’s need for faster or more cost-effective software delivery. By prioritizing "resource utilization" over "value creation," it traps both parties in a cycle where performance optimization is economically irrational for the vendor, leading to stagnant delivery rates and perpetual reliance on effort-based billing (see Figure 5).

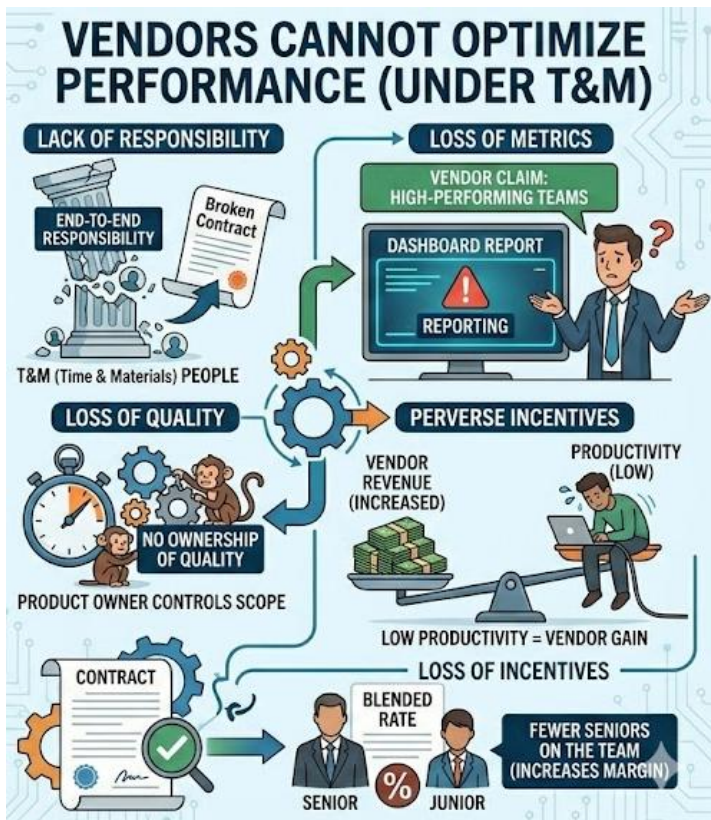


Figure 5: Why vendors cannot optimise performance under the T&M model

## Moving to better practices

To move beyond this industry 'bad practice', NPR 5333 provides a structured, objective framework designed to transition these engagements into outcome-oriented partnerships. Rather than managing inputs, this guideline forces a focus on four dimensions:

**Faster** (time-to-market)

**Better** (quality/defect density)

**Cheaper** (cost-per-unit-of-value)

**Happier** (team and stakeholder satisfaction).

However, the challenge has been how to quantify the "output" in a way that is both simple enough for agile teams to maintain and rigorous enough for enterprise-level benchmarking. This is where ISO standards for Functional Size Measurement like IFPUG, Nesma or COSMIC step in.

**Nesma Easy Functional Sizing (EFS)** was designed as the bridge. Recognized under ISO/IEC 25986, EFS strips away the complex, heavy-duty measurement overhead of traditional methods. It focuses purely on what the user needs: the logical files that hold data and the transactions (EIs and EOs) that move that data. By restricting the measurement to these two transaction types and logical files, EFS provides a clean, standardized "currency" for software size that is immune to differences in programming language, methodology, or team seniority.

**IFPUG Simple Function Points (SiFP)**, which measures only Logical Files and Transactions and multiply these with fixed factors.

When ISBSG performance data is layered onto this standardized output, organizations finally have a verifiable yardstick. Customers can stipulate contract terms based on clear functional targets, and suppliers can demonstrate their performance through objective, empirical data rather than speculative estimates.

This integration of functional size measurement and the NPR 5333 framework essentially fixes the "bad practice" of T&M. It moves the conversation from "how many hours did we spend?" to "how much functional value did we create for the budget invested?" By normalizing performance across the industry, this approach ensures that output-based contracting becomes a repeatable, scalable, and fair mechanism for modern software procurement.

## How ISBSG data supports output-based contracting

To effectively transition from input-based management to a high-performance, output-oriented culture, the Nesma-aligned framework—as detailed in the "Beyond

Velocity" masterclass curriculum—advocates for the implementation of five core metrics:

**Delivery Speed** (time-to-market)

**Productivity** (delivery rate per unit of size)

**Quality** (defect density)

**Cost-Efficiency** (budget adherence relative to output)

**Value** (functional units delivered per currency unit spent).

These five pillars are designed to create a balanced dashboard that prevents the dangerous, localized optimization often seen in traditional delivery, such as teams prioritizing raw speed at the severe expense of code maintainability or user requirements (see Figure 6).



**Figure 6: Core metrics for contracts**

While these metrics provide the structure for measurement, their utility is entirely dependent on the quality of the baselines against which they are compared. This is where the integration of ISBSG data becomes a critical, strategic differentiator. By leveraging the vast, global repository of industry-standard project data compiled by the ISBSG, organizations can move away from relying on internal historical data that may be skewed by inefficient practices or localized technical debt.

ISBSG data acts as the empirical foundation that enables the creation of realistic, market-calibrated targets. Specifically, this data allows leaders to:

- **Establish Comparative Baselines:** Organizations can compare their internal productivity and quality metrics against global peers operating with similar technologies, project sizes, and complexity levels.
- **Identify Performance Gaps:** ISBSG data illuminates whether a team is performing at, above, or below the industry standard, moving the dialogue from subjective opinion to evidence-based discussion.

- **Calibrate Realistic Targets:** Rather than setting arbitrary goals based on past performance, teams can set "stretch goals" that are grounded in what is demonstrably achievable across the industry.
- **Validate Investment Decisions:** Benchmarking against ISBSG data provides the objective justification needed to demonstrate that a project's budget is appropriately aligned with expected functional output.

Ultimately, by pairing this metrics framework with the empirical authority of ISBSG data, organizations create a "closed-loop" performance system. This ensures that the objectives set at the beginning of an engagement are not only measurable but also inherently fair and grounded in the reality of modern software development benchmarks.

### Why not Story Points?

Often we hear: "...we have metrics, our teams use story points, velocity and burndown charts to manage the delivery". However, story points are a subjective, relative effort estimation method, using non-standardized units that can't be compared between teams. It is impossible to use Story Points in contracts or for any other management information purpose.

Incorporating Function Point metrics into software contracts is essential for establishing an objective, transparent, and auditable foundation, whereas using Story Points for such purposes is fundamentally inappropriate. As illustrated in the provided comparison (Figure 7), Story Points function as a local, subjective "currency" that varies significantly between teams, making them impossible to benchmark against external industry standards. This subjectivity leads to "point inflation," where the lack of a standardized definition undermines the consistency and integrity of the measurement over time.

Furthermore, because Story Points focus heavily on effort and technical complexity—rather than the functional value delivered to the user—they foster a "black box" environment that diminishes stakeholder trust. In contrast, Functional Sizing Methods (FSM) rely on ISO-standardized units that are repeatable, auditable, and inherently grounded in user-delivered value. By utilizing FSM alongside empirical ISBSG data, organizations can compare their performance against thousands of real-world projects. This ensures that contracts are built upon verifiable industry reality rather than team-specific estimations.

Feature	Team-Specific Metrics (e.g., Story Points)	Objective Metrics (FSM + ISBSG Data)
<b>Unit of Measure</b>	Subjective "Points" (Local Currency)	ISO Standardized Functional Units
<b>Benchmarking</b>	Impossible; no external reference	Global; compares against thousands of projects
<b>Consistency</b>	Subjective; prone to "point inflation"	Repeatable and auditable across teams
<b>Value Focus</b>	Effort and technical complexity	User-delivered functional value
<b>Stakeholder Trust</b>	Low; perceived as a "black box"	High; grounded in industry reality

Figure 7: Team-specific vs Objective metrics

## Conclusions

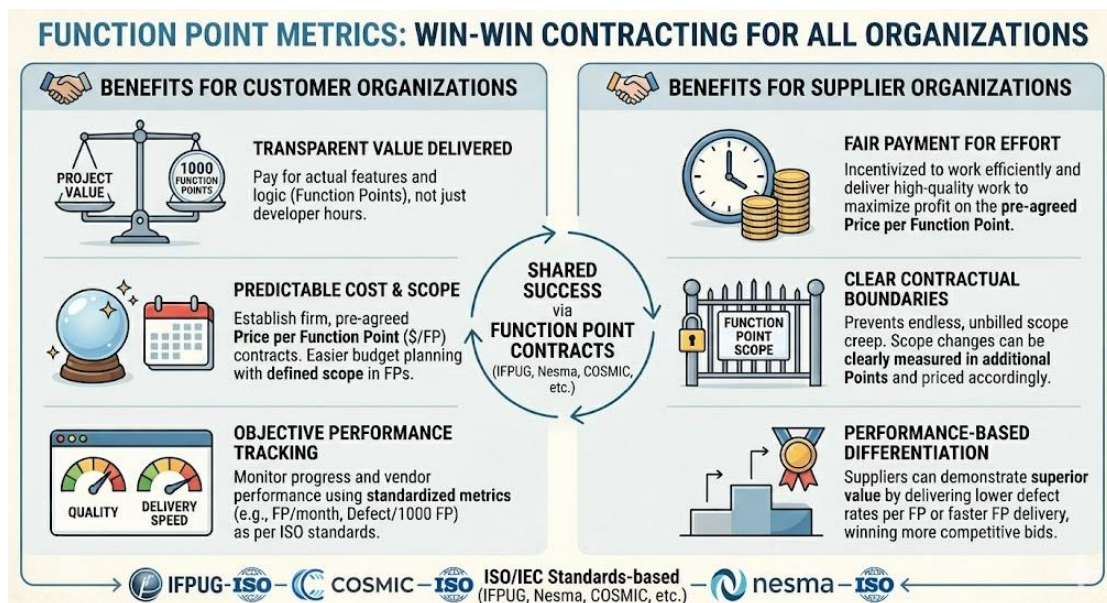


Figure 8: Benefits of function-point metrics

The transition from traditional, input-based Time & Materials contracts to a modern, output-oriented model is a strategic imperative for organizations seeking to eliminate the structural inefficiencies and transparency gaps that plague current software delivery. By replacing the focus on hours consumed with the objective measurement of functional value—using standardized functional size measurement metrics—vendors are incentivized to pursue genuine productivity improvements rather than labor-intensive workarounds (see Figure 8).

When these output metrics are paired with the empirical benchmarks provided by the ISBSG and managed through the balanced scorecard approach defined in NPR 5333, both customers and suppliers can operate within a framework of mutual trust and data-driven accountability. Ultimately, this shift does more than just optimize costs; it fosters a performance-based culture where the success of the vendor is perfectly aligned with the timely, high-quality, and cost-effective delivery of real business value to the customer.

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 of software development projects and maintenance & support activities. It helps 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/](http://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

Find an ISBSG partner in your country: <https://www.isbsg.org/board/>