



GALORATH

AI-Enhanced Benchmarking for IT Project Performance

Industry

Energy Sector

Products used

AI Solutions Services

Background

A global enterprise in the energy sector needed to evaluate the cost and schedule performance of approximately 500 IT projects completed since 2021. These projects supported critical business functions, including exploration, supply chain, manufacturing, and commercial operations. Leadership sought an objective benchmarking process that could reveal performance patterns, identify outliers, and inform future investment decisions.

Benchmarking Challenges

Traditional benchmarking approaches, categorical matching and basic statistical analysis, proved inefficient to deal with large volumes of data. They failed to account for nuanced project characteristics such as delivery methodology, system complexity, or functional scope, which often drive actual costs. **Additional obstacles included:**

Disparate data sources with inconsistent taxonomies and naming conventions.

Manual classification processes that would have taken weeks to complete.

Limited comparability across diverse IT project types, reducing stakeholder confidence in benchmarking results.

AI-Driven Benchmarking Approach

Galorath implemented an AI-enhanced benchmarking solution that leveraged large language models (LLMs) to transform static datasets into dynamic, contextually aware knowledge repositories. **The solution included:**

Semantic labeling and contextual categorization: LLMs interpreted unstructured project descriptions, technical specifications, and requirements, applying consistent taxonomies across more than 5,000 data points spanning 66 attributes.

Similarity scoring: AI generated nuanced similarity measures that captured subtle but cost-critical commonalities beyond categorical matching. The AI generated scoring was used as the underpinning variable to produce boxplots and scatterplots of the data.

Dynamic benchmarking: Projects could be compared and refined by delivery method (Agile, hybrid, waterfall), operational platform (Sever, Cloud, Mobile) and solution type (SaaS, COTS, in-house), as well as across industries.

Human-in-the-loop validation: Expert oversight ensured traceability, defensibility, and accountability in AI-generated classifications.

Results: Accuracy, Speed, and Stakeholder Confidence

The AI-enabled process delivered substantial improvements over traditional methods:

- **Contextual accuracy:** Benchmarks reflected real cost drivers rather than surface-level categories.
- **Speed and efficiency:** Weeks of manual effort were reduced to a single day while improving consistency.
- **Stronger confidence:** Stakeholders trusted evidence-based, semantically aligned comparisons that explained outliers with clarity.
- **Future-proofing:** The organization established a repeatable, AI-enabled process to continuously refine benchmarking as new projects and technologies emerge.

By embedding AI into the benchmarking process, the enterprise achieved **faster, more accurate insights that supported stronger investment decisions across its global IT portfolio.**

[Request a consultation](#)