

How Much Should Your AI Development Efforts Cost?

October 14, 2026

© 2026 Unison. All Rights Reserved.

Agenda

- Motivation for this research
- Where traditional estimation breaks down
- AI Lifecycle Phases
- What should change with ISBSG Submissions to address AI development projects



© 2026 Unison. All Rights Reserved.

Motivation

- AI enabled functionality
  - No longer experimental or peripheral
  - Often central to mission critical capabilities in many domains
- AI cost estimation requires estimators to reframe how they think
- Need a structured estimation framework focused on AI specific activities that is about cost realism, not hype



© 2026 Unison. All Rights Reserved.

Further Motivation

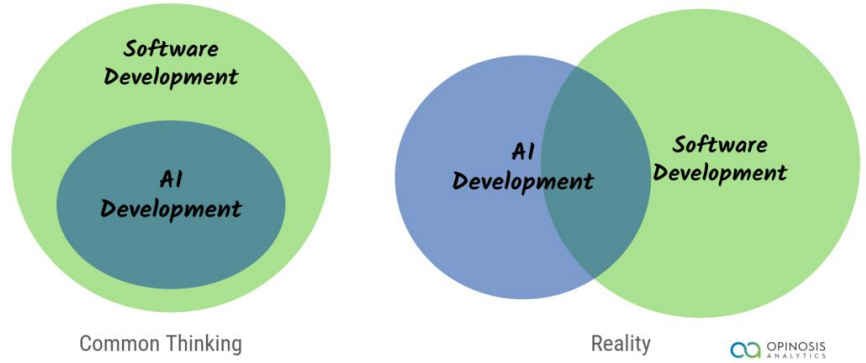
- AI is no longer optional, but estimation hasn't caught up
  - Code still matters – but no longer the dominant driver of cost
  - Data quantity and quality dominate
  - Experimentation and iteration as algorithms are selected and models are trained impact effort and schedule
  - Learning continues post deployment
  - Traditional estimation struggles with AI focused activities:
    - Data Acquisition
    - Data Preparation
    - Data Annotation
    - Algorithm Selection
    - Model Tuning and Training



© 2026 Unison. All Rights Reserved.

Traditional Software Estimation Breaks Down for AI

- Code is no longer the dominant artifact
  - Data volume and quality
  - Experimentation cycles
- Development is inherently experimental
  - Learning occurs through iterative activities such as training, evaluation and refinement
- Effort scales nonlinearly
  - Compliance constraints
  - Retraining
- Sustainment costs dominate
  - Monitoring
  - Infrastructure operations
  - Data drift and retraining

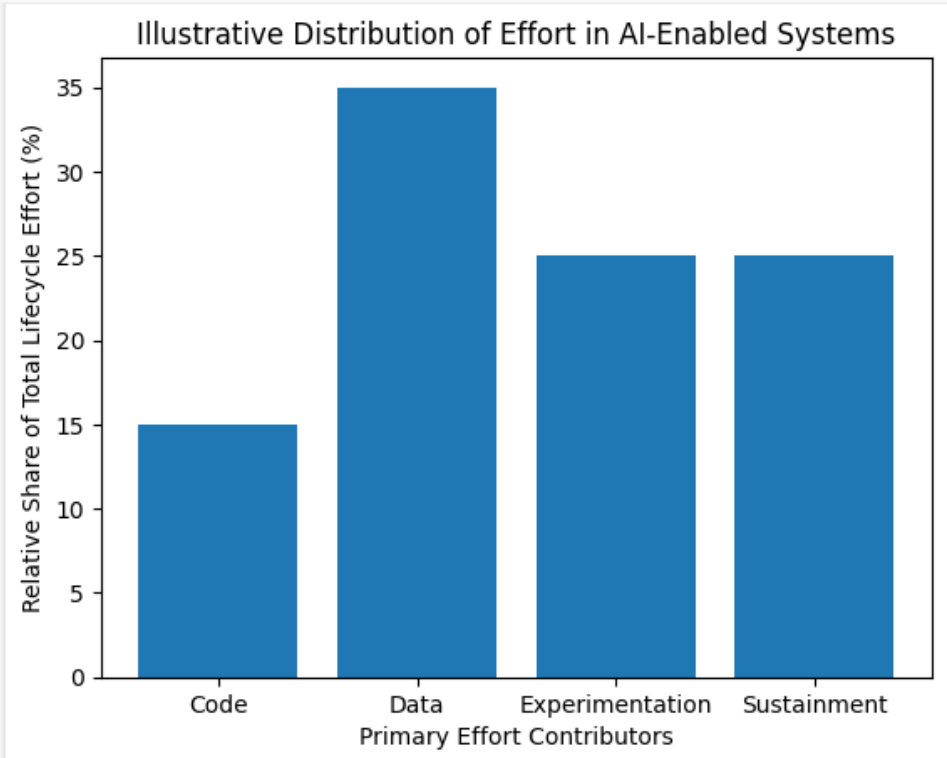


If you size AI like code, you systematically underestimate it.

© 2026 Unison. All Rights Reserved.

Traditional Software Estimation Breaks Down for AI

- Code is no longer the dominant artifact
- Development is inherently experimental
- Effort scales nonlinearly
- Sustainment costs can be a significant share of lifecycle cost.



If you size AI like code, you systematically underestimate it.

Amershi et al., *Software Engineering for Machine Learning,ICSE 2019*
Sculley, et al.;., Hidden Technical Debt in Machine Learning Systems, NeurIPS, 2015

© 2026 Unison. All Rights Reserved.

Why Traditional Size Measures Fail



© 2026 Unison. All Rights Reserved.

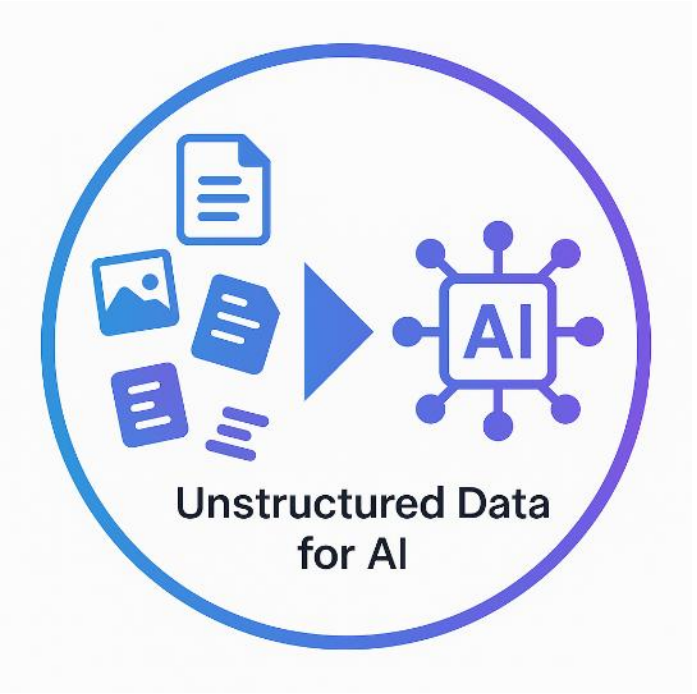
AI Lifecycle Phases that Drive Cost



© 2026 Unison. All Rights Reserved.

Data is A Major Cost Driver

- Data Acquisition and Ingestion – Data needs to be identified, collected, ingested and stored
  - Data volume, freshness, and diversity
  - Licensing and access fees
  - Legal and compliance reviews
  - Engineering effort for ingestion pipeline
  - Compute and data storage
- Data Preparation – Raw data needs to be cleansed, transformed, integrated and validated
  - Number of data sources and heterogeneity of data schema(s)
  - Data quality
  - Regulatory and security requirements
  - Degree of automation vs. manual intervention



Data is not an input – it is a cost driver.

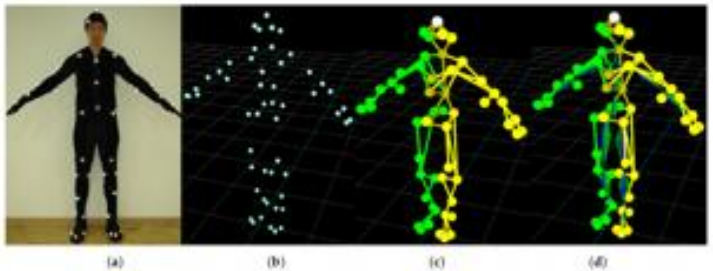
© 2026 Unison. All Rights Reserved.

Data Annotation: Not All Labels are Equal

- Annotation type
  - Classification
  - Segmentation
  - Sequence labeling
- Consensus Factor/Quality Expectations
- Rework Rate
- Complexity and Ambiguity of dataset(s)





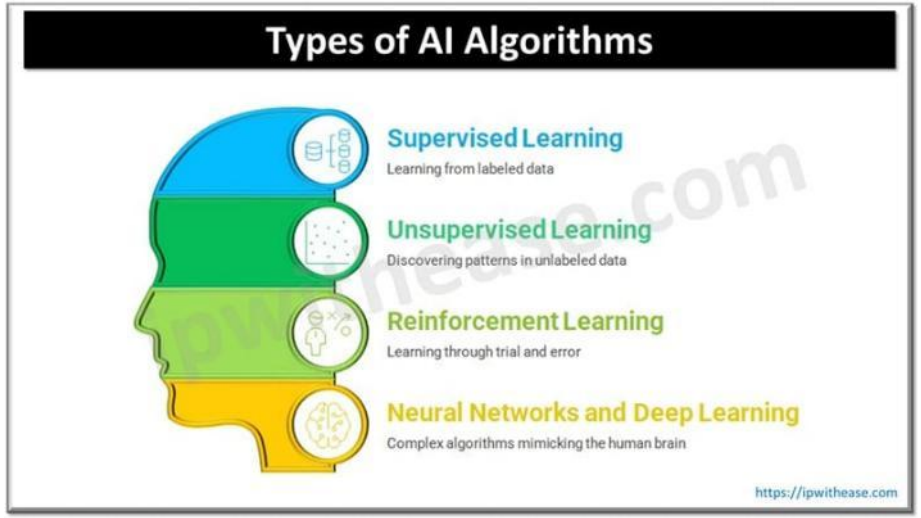


Annotation type can vary cost by orders of magnitude

© 2026 Unison. All Rights Reserved.

Algorithm Choice is a Cost Decision

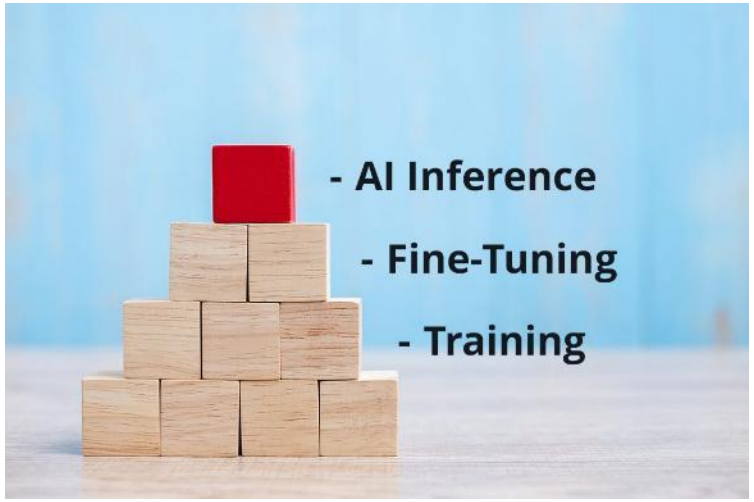
- Every algorithm choice trades performance for cost, risk
  - Compute
  - Iterations
  - Certification
  - Deployment feasibility



Accuracy is not the only objective – affordability and defensibility matter

© 2026 Unison. All Rights Reserved.

Model Training and Tuning: The Hidden Multiplier



- Models need to be trained to learn from the training data
  - Dataset size and feature dimensionality
  - Hardware type (CPU, GPU, TPU) and utilization
  - Training iterations
  - Hyperparameter searches multiply cost
  - Parallelization and optimization strategies
- Compute + Labor + Schedule Risk

This is where budgets quietly disappear

© 2026 Unison. All Rights Reserved.

Sustainment Is Continuous Re-Development

- AI sustainment is not post-development maintenance but rather continuous re-development under operational conditions
  - Model drift
  - Retraining cycles
  - Inference computing
  - Monitoring and performance validation
  - Compliance, auditability and security updates



© 2026 Unison. All Rights Reserved.

Recommended ISBSG Data Collection for AI Projects

- Hours for AI Specific Activities
  - Data Acquisition
  - Data Preparation
  - Data Annotation/Labeling
  - Algorithm Selection
  - Model Training and Tuning
- Type of AI Project (Computer vision, NLP, Gen AI, Predictive Machine Learning, etc.)
- Data type (text, image, audio, video, tabular or multimodal)
- Data Volume
- Data Source (internal, public, commercial/licensed)

© 2026 Unison. All Rights Reserved.

Recommended ISBSG Data Collection for AI Projects

- Data Preparation requirements (cleansing, formatting, deduplication, integration, privacy review, data quality remediation)
- Data Annotation/Labeling Required? Percentage requiring annotation/labeling
- Level of model experimentation required (low, medium, high)
- Frequency of model updates or retraining
- Domain expertise required (General Knowledge, Business Knowledge, Professional Expertise, Specialized Technical Expertise, Mission Critical Expert Knowledge)
- Level of governance, privacy, security, validation required (low, medium, high)

© 2026 Unison. All Rights Reserved.

AI cost does not equal code cost!



© 2026 Unison. All Rights Reserved.

Q&A

© 2026 Unison. All Rights Reserved.

Backup

© 2026 Unison. All Rights Reserved.

Recommended ISBSG Data Collection for AI Assisted SW Development

- How much AI assistance in each of the phases of software development
- AI Tools used in which activities
- Team experience with AI tools used on project
- Project information available to the AI tool
- How were AI generated code/changes checked

© 2026 Unison. All Rights Reserved.