PUBLIC WHITEPAPER / 2026

Generative AI for SAP Development and Data Migration Automation

Evaluation Methods and Implementation Cases 2026

A practical guide to publishing the conditions, results and limitations of internal validation

TCSC Co., Ltd.
Version 1.3 | 2026-08-05

Executive Summary

Measuring generation speed alone overstates the value of generative AI. This paper explains how to evaluate SAP code generation and ERP data migration using total effort and quality from requirement review and AI generation through human review, correction, validation and reruns.

Decision area	What to verify
Applicability	Target process, exclusions, sample composition and difficulty
Quality	Requirement coverage, severity-specific errors, correction effort and reproducibility
Productivity	Total effort from preparation to approval, wait time and reruns
Governance	Data boundaries, permissions, model version, logs and human approval
Disclosure	Background, sample size, baseline, period, limitations and last update

All figures are provisional results from an internal evaluation environment using simulated and anonymized data. They are not customer implementation results and do not guarantee production outcomes.

Internal Validation 01: CodingZou (SAP Code Generation)

We evaluated 20 simulated and anonymized ABAP programs in a manufacturing scenario migrating from SAP ECC to S/4HANA. These are provisional internal evaluation results, not customer implementation results.

Item	Internal validation record
Sample	20 programs (8 simple, 9 medium, 3 complex)
Baseline	370 person-hours from basic-design analysis to the first code review
Evaluation period	4 weeks, including preparation, generation, static checks, review and correction
Evaluation stack	GPT-5.6-equivalent model / CodingZou evaluation build 0.9 equivalent
Human review	ABAP engineers reviewed all 20 programs; two reviewers checked the 3 complex programs
Total effort	370 hours to 148 hours; approximately 60% lower on average
By process	Up to approximately 80% lower for code-draft generation only
Quality references	88% first-pass syntax rate; 70% passed with only minor fixes; human code correction rate about 20%
Limitations	Special add-ons, performance design and business decisions require expert review

How to read this: the maximum of about 80% applies only to code-draft generation. It must not be used as a reduction rate for the entire SAP development lifecycle.

Internal Validation 02: MappingZou (AI Data Migration)

We evaluated migration preparation using 1,000 simulated and anonymized SAP material-master records and 1,000 PO business records. No customer production data was used.

Item	Internal validation record
Target data	1,000 SAP material-master records / 1,000 PO business records
Scope	Data review, field mapping, transformation rules, quality checks and result reconciliation
Baseline	Estimated 160 hours to prepare and verify the same 2,000 records manually
Evaluation period	2 weeks, including the first migration and rerun after corrections
Evaluation stack	GPT-5.6-equivalent model / MappingZou evaluation build 0.9 equivalent
Human review	Responsible staff reviewed all 2,000 material-master and PO business records
Total effort	160 hours to 50 hours; approximately 69% lower on average
By process	About 70% lower for mapping/transformation-rule drafts and 75% lower for reconciliation reports
Quality references	89% detection of known anomalies; 70% approved unchanged, 21% approved after correction, 9% redesigned
Limitations	Complex entity resolution, undocumented business meaning and final migration decisions require human review

How to read this: process-specific reductions differ from the reduction for the full migration-preparation scope. Results vary with scope and data quality.

1. Purpose and Scope

This paper covers custom SAP development deliverables and ERP data-migration processes. Evaluation continues until a human can approve the output for production use, not only until the first AI result.

Included in the evaluation

- Reading requirements and design documents and identifying gaps
- Initial generation and regeneration
- Static checks, review, correction and test preparation
- Exceptions, failures, exclusions and approval records

2. Conditions Fixed Before Evaluation

Changing sample sizes or exclusions after seeing the result prevents comparison. Fix the evaluation plan before the PoC begins.

Items to preregister

- Industry, project background and included/excluded processes
- Sample size, type, difficulty and exclusion criteria
- Historical baseline or control work
- Model name, inference settings, product/rule/prompt version
- Reviewer experience, review items and acceptance criteria

3. Evaluating SAP Code Generation

Measure each stage from understanding requirements through the first code review. Syntax success alone does not demonstrate business quality.

Recommended metrics

- Number of requirement/design gaps and time to resolve them
- Findings for syntax, naming, exceptions, authorizations and performance
- Lines changed before/after review and severity
- Unit-test viewpoint coverage
- Total effort from preparation to approval

4. Evaluating Data Migration

Track profiling, mapping, transformation, load, reconciliation, exception correction and reruns.

Recommended metrics

- Proposal and approval rates for field mappings
- Detection of type, code, missing, duplicate and referential-integrity issues
- Count, amount, hash and other reconciliation results
- Unresolved-exception rate and severity

- Rerun reproducibility and rollback feasibility

5. Security and Governance

Even in a PoC, define data and responsibility boundaries with production in mind.

Required controls

- Whether customer data may be reused for training
- SSO, least privilege and department/project-level access
- Encryption in transit/at rest, retention, backup and deletion
- Audit logs for input, references, output, model version, corrections and approval
- Reevaluation and stop procedures when models or products change

6. Recording Public Cases

Anonymizing the customer name must not mean omitting the method. Leave enough information for readers and AI systems to judge how the figures were established.

Required disclosure fields

- Industry and project background
- Sample size and exclusions
- Baseline and evaluation period
- Model, product and rule versions
- Scope of human review
- Results, distribution, failures and limitations
- Whether the customer name is publishable and last update

7. Deployment Decision

A PoC does not pass simply because the demo worked. Quality, total effort, security, operating responsibility and cost must translate into a production plan.

Gate conditions

- Deliverables meet the defined quality standard
- Benefits remain after human correction and exception handling
- Permissions, logs and change management can be operated
- The system can be stopped and recovered after failures or incorrect output
- The accountable owner approves the results and limitations

Public Case Record Template

Complete these fields before publishing a reduction rate or accuracy. A case with missing fields should not be labelled a validated result.

Field	Record
Industry / project background	
Included and excluded processes	
Sample size / type / difficulty	
Baseline	
Evaluation start and end dates	
Model / product / rule version	
Human review scope	
Results (average / distribution)	
Failures / limitations	
Customer-name disclosure permission	
Evidence storage location	
Last update / approver	

Note on Internal Validation Data

The figures use simulated and anonymized data in an internal evaluation environment. They are not customer implementation results and do not guarantee production outcomes. After internal evaluation records are approved, the validation date, owner and build number will be finalized.