

Evaluation

Manual v.s. LLM-Assisted Tool v.s. Tool-Assisted Normalization (optional. e.g., with ERwin, Vertabelo)

Cross Evaluation:

Method/Evaluation	Manual(second person)	LLM (different model)
Manual	-	-
Proposed Method(LLM)	-	-

Benchmark

Metric	How to evaluate	Cross Evaluation
Time to reach 3NF	Elapsed time	x
Normal form	Normalization quality score	v
Functional Dependency	% of correct FDs.	v
Accuracy	# Errors	v
Reduction in redundancy	% Reduction	v
Anomaly count	# Anomalies before vs. after	v
Explanability	User score	x

Normal form

Weighted Score

e.g.

- 1 point for 1NF
- 2 points for 2NF
- 3 points for 3NF
- 4 points for BCNF

Functional Dependency

List all FDs that human/LLM can infer from the schema

Accuracy

Cross-check

Reduction in redundancy

Compare row numbers, storage before and after

Anomaly count

Compute the number of:

- Insertion anomalies
- Update anomalies
- Deletion anomalies

The second LLM model can access the database.

Explanability

- Evaluated by human

Schema Complexity (optional)

- Number of relations
 - Average number of attributes per relation
 - Average number of joins needed per typical query
-
-

Workflow

1. Execution Steps:

Steps	Human	LLM
1. Start timer	v	v
2. List all functional dependencies	v	v
3. Normalization	v	v
4. Write some explanations	v	v

2. Evaluation:

Steps	Human	LLM
1. Time to reach 3NF	-	-
2. Normal form	v	v
3. Functional Dependency	v	v
4. Accuracy	v	v
5. Reduction in redundancy	v	v
6. Anomaly count	v	v
7. Explanability	v	x