

Sample audit snapshot

Project type: Hausa prompts, responses, transcripts, annotations, and speech/text dataset rows.

Sample size: 20 synthetic rows. **Use case:** pre-training, evaluation, pilot review, localization QA, or customer-facing readiness.

Deliverable: severity-coded issue log, score summary, Hausa-specific notes, and practical fixes.

Overall verdict

Usable with targeted fixes

The sample contains enough usable Hausa to continue, but high-risk rows should be corrected and rechecked before use in training, evaluation, pilots, or customer-facing workflows.

- Ready or minor edits: **14 / 20**
- Needs human review: **4 / 20**
- High-risk or not ready: **2 / 20**
- Main risks: meaning drift, unnatural Hausa, entity/number handling, and inconsistent sensitive guidance.

What the buyer receives

- Issue-coded report with severity, category, and item IDs.
- Plain-English explanation of each Hausa quality risk.
- Hausa-specific notes on fluency, tone, cultural fit, terminology, and transcript issues.
- Practical fix recommendations or reviewer guideline notes.
- Production-readiness decision: ready, minor edits, human review needed, or reject.

Scorecard

Review dimension	Score	What it means
Meaning accuracy	3.8 / 5	Mostly preserved, but 2 items changed user intent.
Natural Hausa fluency	3.6 / 5	Understandable, with translationese and stiff wording.
Terminology consistency	3.7 / 5	Core terms need a small glossary before scale-up.
Cultural and local fit	4.0 / 5	Generally suitable, but some advice feels imported.
Safety and user trust	4.2 / 5	Mostly safe, with one high-risk guidance issue.
Dataset readiness	3.5 / 5	Good pilot material, not clean enough for unchecked training.

Starter pilot

A practical first step is a focused audit of **100-300 Hausa prompts, responses, transcripts, annotations, app strings, or dataset rows**.

The goal is to identify the highest-risk patterns early, improve reviewer guidance, and decide whether the data is ready for training, evaluation, pilots, or customer-facing use.

Sample issue log

ID	Data type	Finding	Severity	Practical fix
HAI-001	Prompt-response	The Hausa answer changes the task from resetting a password to creating a new account.	Major	Preserve the exact user action and rewrite the answer in clear Hausa.
HAI-002	Speech transcript	A payment amount is transcribed incorrectly: <i>dubu biyar</i> becomes <i>dubu biyu</i> .	Critical	Add number/entity verification and flag mixed Hausa-English audio for reviewer check.
HAI-003	UI/localization	<i>Danna nan don ci gaba gaba</i> is understandable but duplicated and unnatural.	Minor	Use concise UI Hausa such as <i>Taba nan don ci gaba</i> for mobile contexts.
HAI-004	Support/safety	The answer gives imported emergency guidance without local context or product-approved wording.	Critical	Replace with approved safety guidance and localize the escalation path.