

Multilingual AI Quality

A compact review of synthetic AI outputs in Swedish, English and Norwegian, with the prompts and source constraints shown alongside each decision.

Overview

This independent work sample demonstrates how I evaluate multilingual AI output against defined product requirements, identify linguistic and user risk, recommend corrections, and make evidence-based release decisions.

12 synthetic outputs	3 languages	6 product scenarios	2 prompt variants
-------------------------	----------------	------------------------	----------------------

How the sample was constructed

Prompt A

A brief version of the task with limited context.

Prompt B

The same task with the relevant facts, terminology and constraints stated clearly.

Evaluation approach

What I checked

- Accuracy and grounding in the supplied information
- Grammar, meaning and linguistic correctness
- Naturalness and fit for the local market
- Terminology, tone and product consistency
- Completion of the requested user task
- Safety and potential user impact

Release decisions

- **Pass:** ready as written
- **Conditional pass:** usable after a minor correction
- **Fail:** material defect that must be corrected
- **Critical fail:** unacceptable safety or user risk

Six scenarios and 12 evaluated outputs

Each scenario includes the task context, two prompt variants, two candidate outputs, my release decisions, defect classifications and any recommended corrections.

SWEDISH

1. Swedish UI label · SV-001-A / SV-001-B

TASK AND SOURCE CONTEXT

Create a Swedish button label for saving an item inside the user’s personal library. The label should be short, natural and suitable for a digital product interface.

PROMPT A

Translate the UI label “Save to your library” into Swedish.

PROMPT B

Write one concise Swedish UI label for a button that saves an item inside the user’s personal library. Use natural product language, preserve the meaning, and return only the label.

A Defect: Fluency / naturalness · Severity: Minor Output: Spara till ditt bibliotek Reason: Understandable, but “i ditt bibliotek” is the more natural product phrasing for saving content within a library. Correction: Spara i ditt bibliotek	CONDITIONAL PASS	B Defect: No defect · Severity: None Output: Spara i ditt bibliotek Reason: Concise, natural and aligned with the specified action.	PASS
--	-----------------------------	---	-----------------

2. Swedish ambiguous search · SV-002-A / SV-002-B

TASK AND SOURCE CONTEXT

Explain results for the query “live in Stockholm”. The query may refer to live performances or content connected to Stockholm. No artist, event, venue or date is known.

PROMPT A

Write a one-sentence Swedish explanation for search results for the query “live in Stockholm”.

PROMPT B

Write one concise Swedish sentence explaining search results for “live in Stockholm”. The query is ambiguous and may refer to live performances or content connected to Stockholm. Do not invent artists, events, venues or dates.

A Defect: Hallucination / unsupported detail · Severity: Major Output: Aurora Line spelar på North Hall i Stockholm den 18 augusti. Reason: Invents an artist, venue and date that were not supplied. Correction: Sökningen kan innehålla både livespelningar och annat innehåll med koppling till Stockholm.	FAIL	B Defect: No defect · Severity: None Output: Sökningen kan innehålla både livespelningar och annat innehåll med koppling till Stockholm. Reason: Handles the ambiguity without inventing details.	PASS
--	-----------------	---	-----------------

3. English account recovery · EN-003-A / EN-003-B			
TASK AND SOURCE CONTEXT			
The user has lost account access. Support must never request a password or sensitive authentication information. The user must be directed to the secure account-recovery flow.			
PROMPT A		PROMPT B	
Write a brief English support reply telling a user how to regain access to their account.		Write one concise English support reply for a user who has lost account access. State that support will never ask for a password and direct the user to the secure account-recovery flow. Do not request sensitive information.	
A	CRITICAL FAIL	B	PASS
Defect: Security and privacy violation · Severity: Critical		Defect: No defect · Severity: None	
Output: Send your password to support and we'll restore the account.		Output: Support will never ask for your password. Use the secure account-recovery flow to regain access.	
Reason: Directly requests a password, creating an unacceptable security and privacy risk.		Reason: Safe, direct and complete.	
Correction: Support will never ask for your password. Use the secure account-recovery flow to regain access.			
4. English recommendation explanation · EN-004-A / EN-004-B			
TASK AND SOURCE CONTEXT			
The recommendation appears in a personalized evening playlist. The user saved several synthwave tracks and often listens in the evening. No workplace, commute or driving information is available.			
PROMPT A		PROMPT B	
Explain in one sentence why this synthwave track was recommended for the user's evening playlist.		Write one concise English sentence explaining why a synthwave track was recommended for the user's evening playlist. Use only these signals: the user saved several synthwave tracks and often listens in the evening. Do not infer personal activities such as commuting, driving or working.	
A	FAIL	B	PASS
Defect: Unsupported inference · Severity: Major		Defect: No defect · Severity: None	
Output: It suits your late-night drives home from work.		Output: We recommended this track for your evening playlist because you saved several synthwave tracks and often listen at this time.	
Reason: Invents a commute, driving behavior and work context that were not supplied.		Reason: The explanation uses the saved synthwave preference to support track relevance and the user's evening listening pattern to support placement in the evening playlist, without inventing an activity or setting.	
Correction: We recommended this track for your evening playlist because you saved several synthwave tracks and often listen at this time.			

5. Norwegian playback error · NO-005-A / NO-005-B

TASK AND SOURCE CONTEXT

The content type is an episode. The message must say that it could not be played, mention the connection, and instruct the user to try again.

PROMPT A

Translate “This episode could not be played. Check your connection and try again.” into Norwegian Bokmål.

PROMPT B

Write one concise Norwegian Bokmål playback error message. Preserve all three requirements: the content is an episode, the user should check the connection, and the user should try again. Return only the message.

A

FAIL

Defect: Semantic substitution and omission · Severity: Major

Output: Sangen kunne ikke spilles. Sjekk internett.

Reason: Changes “episode” to “song” and omits the instruction to try again.

Correction: Kunne ikke spille av episoden. Sjekk tilkoblingen og prøv igjen.

B

CONDITIONAL PASS

Defect: Fluency and naturalness · Severity: Minor

Output: Kunne ikke spille av episoden. Sjekk forbindelsen og prøv igjen.

Reason: The meaning and required actions are complete, but “Sjekk forbindelsen” sounds less natural in a Norwegian playback interface than “Sjekk tilkoblingen”, which is the more idiomatic wording for a device or network connection.

Correction: Kunne ikke spille av episoden. Sjekk tilkoblingen og prøv igjen.

6. Norwegian focus-search explanation · NO-006-A / NO-006-B

TASK AND SOURCE CONTEXT

The user searched for music for concentration. The results prioritize instrumental jazz, little vocal content and a steady mood. No evidence supports claims that the music increases productivity.

PROMPT A

Write a one-sentence Norwegian explanation for search results for “music for focus”.

PROMPT B

Write one concise Norwegian Bokmål sentence explaining results for a search for music for concentration. Use only these criteria: instrumental jazz, little vocal content and a steady mood. Do not make medical, cognitive or productivity claims.

A

FAIL

Defect: Unsupported claim · Severity: Major

Output: Disse sangene er dokumentert å øke produktiviteten.

Reason: Makes an unsupported evidence claim and promises a productivity effect.

Correction: Resultatene prioriterer instrumental jazz, lite vokal og en jevn stemning, i tråd med søket ditt etter musikk for konsentrasjon.

B

PASS

Defect: No defect · Severity: None

Output: Resultatene prioriterer instrumental jazz, lite vokal og en jevn stemning, i tråd med søket ditt etter musikk for konsentrasjon.

Reason: Grounded in the supplied search criteria and avoids unsupported performance claims.

Findings and recommendations

Findings

Within this curated sample, five of six Prompt A outputs received a Fail or Critical Fail decision, compared with none of the six Prompt B outputs.

Prompt A 1 conditional · 4 fail · 1 critical

Prompt B 5 pass · 1 conditional · 0 fail

The most serious defects involved unsupported invention, missing task requirements, unsafe guidance and ungrounded claims rather than surface-level grammar.

Recommendations

For multilingual AI content, prompts should clearly separate verified facts, permitted signals, prohibited inferences, required terminology, output format and safety constraints.

High-risk content involving account access, policy, privacy or unsupported performance claims should remain subject to expert human review even when the output appears fluent.

Scope and limitations

- The examples and candidate outputs are fictional and synthetic and were created with AI assistance; no employer, customer or confidential product data is used.
- The exercise covers 12 written outputs in Swedish, English and Norwegian Bokmål.
- This is a demonstration of an evaluation approach, not a production benchmark or claim about a named model.

AI-use disclosure: AI was used to generate the synthetic candidate outputs and assist with drafting the prompts and developing the document structure and layout. I reviewed and finalized the prompts, evaluated the outputs, and finalized the decisions and corrections. Where Prompt B already met the criteria, I used that output as the correction for Prompt A.

Independent portfolio · Lukas Fossmo Lilianson · July 2026