How can you tell if an AI prototype is ready for everyday use?

For: Paušalni obrt · Obrtnici · j.d.o.o. · d.o.o.

Practical business guidance · Reviewed: 2026-11-19

Applicability: Practical business guide

Provide the prototype with typical and problematic examples. Measure how much time remains after human verification, what errors occur, and whether the tool can identify cases that require manual review.

Agree on the criteria before the test.

First, write down what constitutes a good result and what would constitute a serious error. On sample or appropriately anonymised data, test the typical case, incomplete entry and exception. Also process the same examples manually for comparison.

Measure the total time, including corrections. Before wider implementation, determine who reviews the output, when the tool stops and how to revert to the manual procedure. After changing the model or instructions, repeat the key tests.

Example of query sorting

AI correctly suggests a category for most sample queries, but when the message is incomplete it invents a priority. Introduce a "requires manual review" tag and retain human approval before proceeding.

Specific next steps

1. Write the acceptance criterion and list the serious errors.
2. Prepare typical and edge cases with the expected outcome.
3. Measure the savings after checks and corrections.
4. Try a limited application under supervision and note when to stop the test.

Limits of application

The method is used to verify a specific process. When using real data and in sensitive procedures, please separately check the terms of use and obligations.

Sources

NIST · AI Risk Management Framework, testing and measurement

https://airc.nist.gov/airmf-resources/airmf/5-sec-core/

Original page

https://adriatrends.hr/poslovni-vodic/en/pitanja/kako-testirati-ai-prototip-prije-uvodenja