TELMEKON CHALLENGE · TEAM ROBOG

Robot manuals that answer with receipts

Hybrid retrieval over 1,351 cited chunks: keywords, meaning and exact specs, in English, German and Italian.

1,351
cited chunks

22
robot and accessory products

3
languages, one answer policy

THE PROBLEM

Manuals are where support answers go to die

Exact values matter

A 20 kg payload or a page reference must be quoted, not paraphrased.

Users don’t speak manual

"The robot keeps bumping into tables" never matches "obstacle fault".

Chatbots invent specs

Without evidence, a model fills gaps with plausible torque values and page numbers.

Our rule: no evidence, no answer.

INGESTION · RUNS ONCE

From 63 PDFs to tagged, citable chunks

Clean markdown

Converted from PDF, with page markers kept.

→

1,351 chunks

Text, table and figure, each tagged with product, page, section and language.

→

2,039 spec rows

Parameter index for verbatim values with page numbers.

→

ChromaDB

Dense vectors in a local store, built once per machine.

Validation gate: 0 errors. 38 profile cards answer "which robot fits this site?"

THE CORPUS

Every chunk is cited, none is invented

706

text chunks

169

table chunks, values kept as cells

476

figure chunks, captions and legends

0 validation errors

Unique ids, non-empty text, complete metadata.

19 of 22 in all three languages

English, German and Italian editions, side by side.

RETRIEVAL ENGINE

Four retrievers vote, one ranking wins

BM25 keywords

Exact words: CC1, Notausschalter, 20 kg

EN/DE/IT lexicon

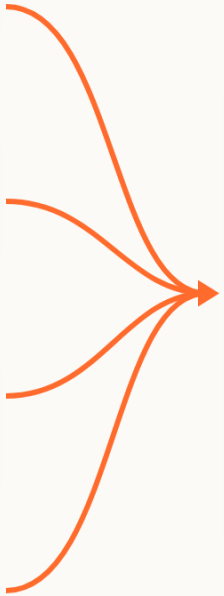
payload ≈ Zuladung ≈ carico

Parameter index

Verbatim spec values with page numbers

Dense vectors

Meaning and paraphrase via ChromaDB



Reciprocal Rank Fusion

Score = sum of 1 / (60 + rank) across lists.

Agreement between retrievers beats one lucky hit.

Output: top-k chunks, de-duplicated across EN/DE/IT editions, citations attached.

WHY WE NEED ALL OF THEM

Each retriever has a blind spot

Retriever	Wins at	Fails at
BM25 keywords	Codes, numbers, exact terms	Paraphrase and synonyms
Dense vectors	Meaning, cross-lingual phrasing	Rare codes and numbers
Parameter index	Verbatim spec values with page	Only what the tables state
EN/DE/IT lexicon	Translated terms like Zuladung	Terms missing from the list

Fusion keeps each strength and drops each weakness.

MULTILINGUAL BY POLICY

Ask in English, German or Italian

English

466

chunks indexed

German

469

chunks indexed

Italian

416

chunks indexed

Three products have no Italian edition: FlashBot, GW1 and Pager. We answer in Italian from EN or DE sources and say so on screen.

TRUST BY CONSTRUCTION

Citations are metadata, never generated text

1. Retrieve

Chunks arrive with product, section and page straight from the store.

2. Compose

Gemini writes the answer when a key is set; a deterministic formatter is the floor.

3. Validate

Uncited sentences go. Every number must appear in its cited chunk.

Offline check: 28/28 answers grounded, 100% of numbers verified

MEASURED TODAY

Numbers we can defend on stage



Harness checks on a small set, not a user study. The suite runs 274 tests.

HONEST STATUS

What we prove after the live run

Production index

gemini-embedding-2 vectors in ChromaDB, one command per machine.

Hybrid numbers

Dense plus keyword metrics are filled in only after the run.

Live Gemini answers

Exercised against the live API, then checked by the validator.

Not built yet: auth, rate limiting, streaming answers, a Telegram client.

Ask the manuals. Check the receipts.

```
python -m retrieve "Wieviel Zuladung hat der FlashBot?"
```

```
python -m api # chat at localhost:8000
```

```
curl -s localhost:8000/health
```

Say "hybrid" only when vector_active is true. Say "Gemini" only when provider is gemini.