RAG EVALUATION · RAGAS RETRIEVAL METRICS

# Context precision vs context recall vs context entity recall

All three grade the **retriever**, not the final answer. Here they all look at the same search result, and each one gives a different score.

## One example, three scores

USER QUESTION

Where is the Eiffel Tower, and when and why was it built?

REFERENCE ANSWER (GROUND TRUTH)

The Eiffel Tower is in Paris, France.[C1] It was completed in 1889[C2] as the entrance to the World's Fair.[C3]

RETRIEVED CHUNKS, IN RANKED ORDER

| | | |
|---|---|---|
| #1 | The Louvre, in Paris, is the most visited art museum in the world. | ✗ |
| #2 | The Eiffel Tower stands in Paris, France. It was completed in 1889. | ✓ |
| #3 | The tower is repainted about every seven years. | ✗ |
| #4 | It was built as the entrance arch for the 1889 Exposition Universelle. | ✓ |

underlined = claim  highlighted = entity

✓ / ✗ = useful for this question

## Context precision

Ranking and noise

**Are the useful chunks at the top of the list?**

OUR RANKING

| #1 | #2 | #3 | #4 |
|---|---|---|---|
| ✗ | ✓ | ✗ | ✓ |
| | P@2 = ½ | | P@4 = 2/4 |

(0.50 + 0.50) ÷ 2 useful = **0.50**

SAME CHUNKS, RERANKED

| #1 | #2 | #3 | #4 |
|---|---|---|---|
| ✓ | ✓ | ✗ | ✗ |
| P@1 = 1 | P@2 = 1 | | |

(1 + 1) ÷ 2 useful = **1.00**

**0.50** Good chunks found, but noise sits above them

## Context recall

Completeness, by meaning

**Did we retrieve everything needed to give the reference answer?**

EACH REFERENCE CLAIM → FOUND IN A CHUNK?

| | | |
|---|---|---|
| C1 | In Paris, France | ✓ #2 |
| C2 | Completed in 1889 | ✓ #2 |
| C3 | Entrance to the World's Fair | ✓ #4 |

An LLM judges meaning, so "Exposition Universelle" counts as the World's Fair. Rank order does not matter here.

3 supported ÷ 3 claims = **1.00**

**1.00** Nothing important is missing

## Context entity recall

Completeness, by named things

**Do the chunks mention the names, places, dates and things from the reference?**

ENTITIES IN THE REFERENCE → IN THE CHUNKS?

✓ Eiffel Tower  ✓ Paris  ✓ France  ✓ 1889  ✗ World's Fair

Entities are matched as sets. The chunk says "Exposition Universelle", a different name, so "World's Fair" does not match.

4 shared ÷ 5 entities = **0.80**

**0.80** One key entity is not named in the chunks

## Side by side

| | Context precision | Context recall | Context entity recall |
|---|---|---|---|
| SIMPLE QUESTION | Is the good stuff on top? | Did we get all the facts we need? | Did we get all the key names and numbers? |
| UNIT COUNTED | Retrieved chunks, and their rank | Claims in the reference answer | Entities in the reference answer |
| FORMULA | `Σ(P@k × relevant_k) ÷ relevant chunks in top K` | `reference claims supported by context ÷ all reference claims` | `\|entities in context ∩ entities in reference\| ÷ \|entities in reference\|` |
| INPUTS | Question, retrieved contexts, reference (or the generated answer) | Retrieved contexts, reference answer | Retrieved contexts, reference answer |
| CARES ABOUT ORDER? | Yes | No | No |
| LOW SCORE MEANS | Too much noise, or useful chunks ranked low | The retriever missed needed information | Key names, dates or terms are missing from the chunks |
| TYPICAL FIX | Add a reranker, lower top-K, use metadata filters | Raise top-K, improve chunking or embeddings, try hybrid search | Add keyword or hybrid search, keep entities inside chunks, add metadata |
| BEST FOR | Keeping the prompt short and focused | General question answering | Fact-heavy domains: history, products, legal, tourism |

**Remember:** Precision asks *"is it clean and well ordered?"* Recall asks *"is anything missing?"* Entity recall asks *"are the key names there?"*