

LLM Metadata Benchmark Report

Introduction

To assess whether metadata generation by a Large Language Model (LLM) is sufficiently reliable for documents published via the GPP-WOO publication platform, a benchmark test was conducted. For this benchmark, 20 documents were selected via OpenOverheid (<https://open.overheid.nl/home>). For each document in this test set, the official metadata was compared with the metadata generated by the LLM. The model used for this report is Mistral Large 2411. The comparison focuses exclusively on the metadata fields that are actually used within the publication process of the GPP-WOO platform. The applied comparison methodology differs per field type. Depending on the nature of the field, fuzzy string similarity, set overlap, or semantic overlap is used. Based on this, a score is calculated per field and per document, after which aggregated results provide insight into the overall performance of the LLM.

Comparison Methodology

- Title → Fuzzy string similarity
- Identifiers → Set overlap (Jaccard)
- Information categories → Fuzzy semantic overlap
- Document types → Fuzzy overlap
- Publisher / Responsible authority / Author → Fuzzy string similarity

Numerical Results

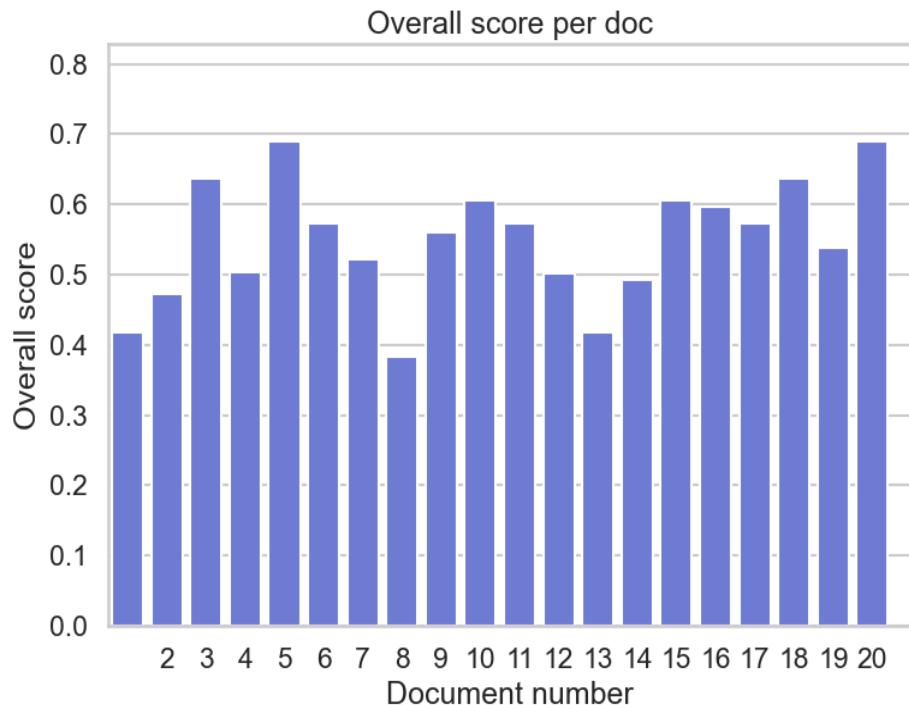
Overall average score: 0.549

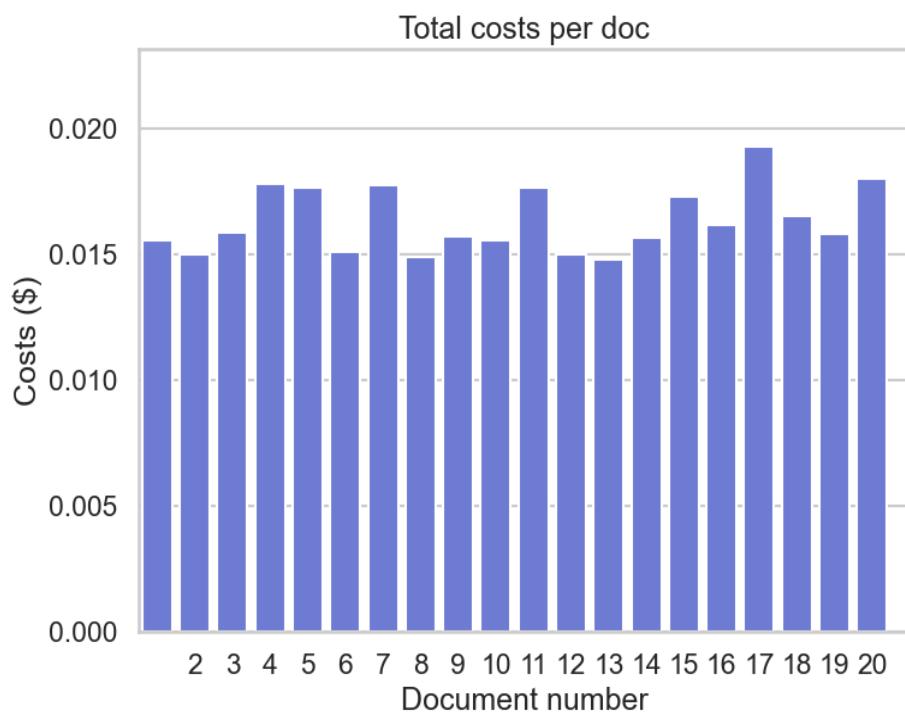
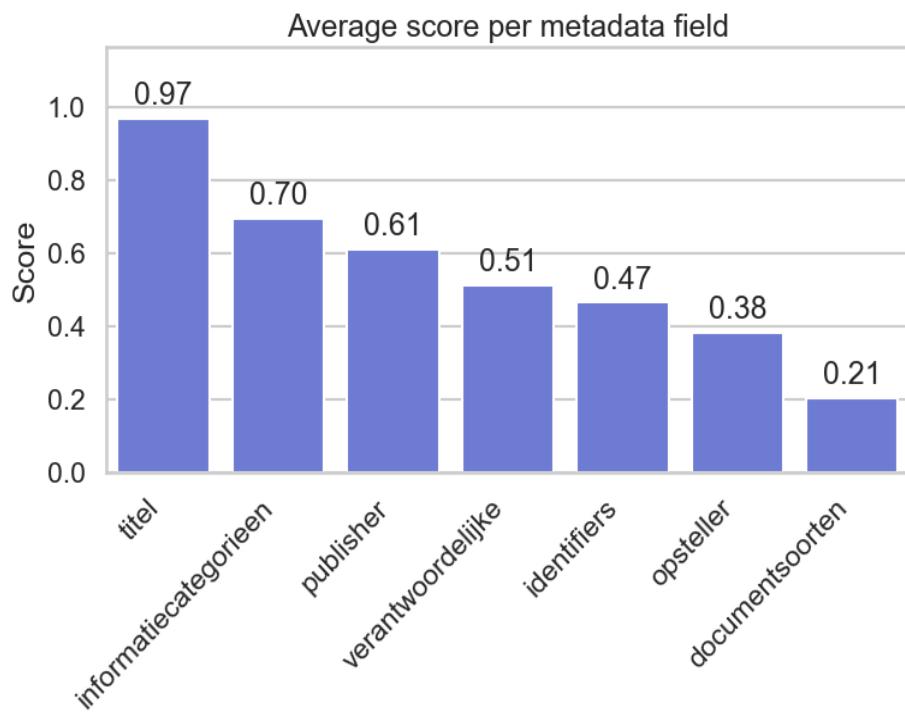
Metadata field	Average score
titel	0.97
informatiecategorieen	0.696
publisher	0.611
verantwoordelijke	0.514
identifiers	0.467
opsteller	0.383
documentsoorten	0.205

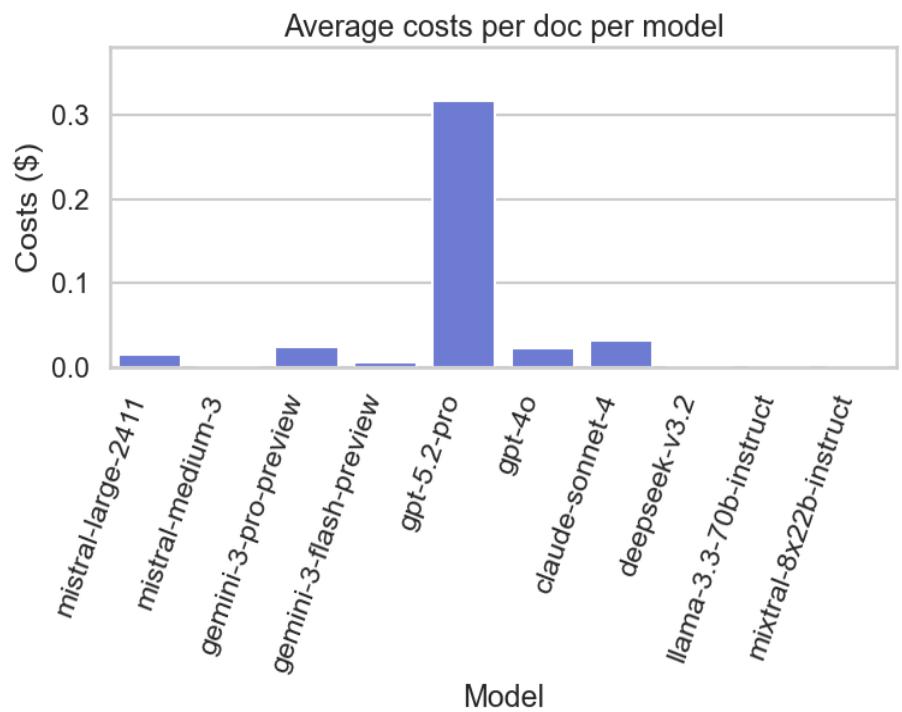
Average number of LLM tokens per document: 5432
Average cost per document for Mistral Large 2411: \$0.016

Model	Average cost per document (USD)
mistral-large-2411	\$0.0160
mistral-medium-3	\$0.0040
gemini-3-pro-preview	\$0.0250
gemini-3-flash-preview	\$0.0060
gpt-5.2-pro	\$0.3160
gpt-4o	\$0.0240
claude-sonnet-4	\$0.0330
deepseek-v3.2	\$0.0020
llama-3.3-70b-instruct	\$0.0010
mixtral-8x22b-instruct	\$0.0010

Visualizations







4. Example Comparisons

Doc	Field	Official	LLM	Score
oep-3497	titel	Wet van 28 januari 2026, houdende wijziging van de begrotingsstaten van het Ministerie van Klimaat en Groene Groei (XXIII) voor het jaar 2025 (wijziging samenhangende met de Najaarsnota)	Wet van 28 januari 2026, houdende wijziging van de begrotingsstaten van het Ministerie van Klimaat en Groene Groei (XXIII) voor het jaar 2025 (wijziging samenhangende met de Najaarsnota) - VERBETERING	1.0
oep-3497	identifiers	['stb-2026-35-n1']	['stb-2026-35-n1', 'Staatsblad 2026, 35-n1']	0.5
oep-3497	informatiecategorieën	['wetten en algemeen verbindende voorschriften']	['WETTEN_AVV']	0.75
oep-3497	documentsoorten	['Staatsblad']	['BESLUIT']	0.235
oep-3497	publisher	ministerie van Justitie en Veiligheid	Ministerie van Justitie en Veiligheid	1.0
oep-3497	verantwoordelijke	ministerie van Justitie en Veiligheid	Ministerie van Klimaat en Groene Groei	0.63
oep-3497	opsteller	ministerie van Justitie en Veiligheid	Onbekende organisatie	0.345
oep-608d	titel	Wet van 11 februari 2026 tot wijziging van de Omgevingswet (legalisering PAS-projecten)	Wet van 11 februari 2026 tot wijziging van de Omgevingswet (legalisering PAS-projecten) - Verbetering Staatsblad 2026, 36	1.0
oep-608d	identifiers	['stb-2026-36-n1']	['stb-2026-36-n1', 'ISSN 0920-2064']	0.5
oep-608d	informatiecategorieën	['wetten en algemeen verbindende voorschriften']	['WETTEN_AVV']	0.75
oep-608d	documentsoorten	['Staatsblad']	['BESLUIT']	0.235
oep-608d	publisher	ministerie van Justitie en Veiligheid	Ministerie van Justitie en Veiligheid	1.0
oep-608d	verantwoordelijke	ministerie van Justitie en Veiligheid	Ministerie van Justitie en Veiligheid	1.0
oep-608d	opsteller	ministerie van Justitie en Veiligheid	Onbekende organisatie	0.345
oep-8bfd	titel	Besluit van 12 februari 2026, houdende vaststelling van het tijdstip van inwerkingtreding van de Wet regels integriteit en vervolgsfuncties bewindspersonen (Stb. 2025, 269)	Besluit van 12 februari 2026, houdende vaststelling van het tijdstip van inwerkingtreding van de Wet regels integriteit en vervolgsfuncties bewindspersonen (Stb. 2025, 269)	1.0
oep-8bfd	identifiers	['stb-2026-34-n1']	['stb-2026-34-n1', 'Stb. 2025, 269', 'ISSN 0920-2064']	0.333
oep-8bfd	informatiecategorieën	['wetten en algemeen verbindende voorschriften']	['WETTEN_AVV']	0.75
oep-8bfd	documentsoorten	['Staatsblad']	['BESLUIT']	0.235
oep-8bfd	publisher	ministerie van Justitie en Veiligheid	Ministerie van Justitie en Veiligheid	1.0
oep-8bfd	verantwoordelijke	ministerie van Justitie en Veiligheid	Onbekende organisatie	0.345
oep-8bfd	opsteller	ministerie van Justitie en Veiligheid	Onbekende organisatie	0.345