

BM25 tool usage statistics						
CLI script to extract those values from the logs: <code>cat *.jsonl jq -r 'select(.type == "tool_result" and .bm25_top_k != null) .bm25_top_k' sort uniq -c</code>						
For TOP1 retrieval	SWE			LCA		
Qwen2.5	10	20	30	10	20	30
Run1 - all	216	0	76	100	0	46
Run1 - best	140	160	0	78	0	74
Average	178	80	38	89	0	60
Qwen3						
Run1 - all	5	234	12	4	112	3
Run1 - best	7	184	12	13	172	9
Average	6	209	12	8.5	142	6
For TOP5 retrieval						
Qwen2.5						
Run1 - all	1	0	299	2	0	171
Run1 - best	0	0	300	1	0	159
Average	0.5	0	299.5	1.5	0	165
Qwen3						
Run1 - all	1	252	99	0	149	33
Run1 - best	1	117	101	0	128	62
Average	1	184.5	100	0	138.5	47.5
For TOP10 retrieval						
Qwen2.5						
Run1 - all	1	0	302	1	0	151
Run1 - best	0	0	300	1	0	151
Average	0.5	0	301	1	0	151
Qwen3						
Run1 - all	4	348	24	6	180	5
Run1 - best	4	171	47	5	174	21
Average	4	259.5	35.5	5.5	177	13
Usage statistics per tool						
CLI script: <code>cat qwen2.5/*.jsonl jq -r 'select(.type == "tool_result") if .bm25_top_k then "bm25_top_k" elif .extract_relevant then "extract_relevant" elif .view_readme then "view_readme" else "unknown" end' sort uniq -c</code>						
For file views: <code>cat qwen2.5/*.jsonl jq -r 'select(.type == "tool_decision" and .tool_name == "view_file") .tool_args.file_path' wc -l</code>						
For unique file views: <code>cat qwen2.5/*.jsonl jq -r 'select(.type == "tool_decision" and .tool_name == "view_file") .tool_args.file_path' sort -u wc -l</code>						
SWE-bench Lite						
Run 1						
<i>Qwen2.5-coder 32B</i>	<i>Total</i>	<i>Average per repo</i>	<i>Notes</i>			
bm25_top_k	1795	6	Several calls may happen as the agent refers to tools available in a previous turn of the conversation.			
extract_relevant	1798	6				
view_file	12056	40.2	Some files are viewed many times. This number also includes request to view invalid files.			
view_readme	0	0	This one is never called!			
Unique file views	599	2				

[illegible]

[illegible]