

Morpheus.AI: Evaluation results on benchmark papers (Prompt 2)

BENCHMARK SUMMARY					
Model:	claude-sonnet-4-5-20250929				
Total papers processed:	10				
Successfully completed:	5				
Failed/Incomplete:	5				
Total PNGs generated:	205				
Total CSVs generated:	7				
Min/Max score:	0				
Duration:	50m 48s				
#	Paper	Status	Score	PNGs	CSVs
1	Szabo2010_clean.pdf	✓ Done	5/7	7	0
2	Thapa2024_clean.pdf	✓ Done	7/7	32	1
3	Carpenter2024_clean.pdf	✓ Done	7/7	124	1
4	Atwell2025_clean.pdf	✓ Done	7/7	17	1
5	Berkemeier2023_clean.pdf	✗ max_itera	1/7	0	1
6	Herron2022_clean.pdf	✓ Done	7/7	25	3
7	Scheel2021_clean.pdf	✗ max_itera	0	0	0
8	Mu2024_clean.pdf	✗ max_itera	0	0	0
9	Luque2023_clean.pdf	✗ max_itera	0	0	0
10	Berkhout2025_clean.pdf	✗ max_itera	0	0	0
Success Rate:	50.0%				
Perfect Score	(7/7): 4/10				