

CFR_R

Experts	1.00	0.75	0.76	0.73	0.46	0.57	0.64	0.69	-0.25	0.73	0.58
GPT 5.1	-0.75	1.00	0.89	0.89	0.90	0.87	0.96	0.92	0.05	0.84	0.89
GPT 5.1 *	-0.76	0.89	1.00	0.97	0.67	0.71	0.88	0.66	0.17	0.64	0.76
Flash 2.5	-0.73	0.89	0.97	1.00	0.68	0.71	0.85	0.67	0.16	0.70	0.82
Flash 2.5 *	-0.46	0.90	0.67	0.68	1.00	0.85	0.93	0.90	0.01	0.82	0.81
Grok 4.1	-0.57	0.87	0.71	0.71	0.85	1.00	0.85	0.87	-0.25	0.84	0.89
Grok 4.1 *	-0.64	0.96	0.88	0.85	0.93	0.85	1.00	0.84	0.12	0.78	0.84
Sonar	-0.69	0.92	0.66	0.67	0.90	0.87	0.84	1.00	-0.25	0.89	0.81
Sonar *	-0.25	0.05	0.17	0.16	0.01	-0.25	0.12	-0.25	1.00	-0.42	0.07
Sonnet 4.5	-0.73	0.84	0.64	0.70	0.82	0.84	0.78	0.89	-0.42	1.00	0.79
Sonnet 4.5 *	-0.58	0.89	0.76	0.82	0.81	0.89	0.84	0.81	0.07	0.79	1.00
	Experts	GPT 5.1	GPT 5.1 *	Flash 2.5	Flash 2.5 *	Grok 4.1	Grok 4.1 *	Sonar	Sonar *	Sonnet 4.5	Sonnet 4.5 *

 CFR_{R+P}

Experts	1.00	0.75	0.72	0.73	0.58	0.53	0.59	0.75	-0.20	0.76	0.61
GPT 5.1	-0.75	1.00	0.91	0.89	0.88	0.91	0.91	0.86	-0.02	0.81	0.87
GPT 5.1 *	-0.72	0.91	1.00	0.99	0.67	0.72	0.89	0.65	0.15	0.66	0.77
Flash 2.5	-0.73	0.89	0.99	1.00	0.68	0.71	0.88	0.67	0.16	0.70	0.81
Flash 2.5 *	-0.58	0.88	0.67	0.68	1.00	0.94	0.78	0.96	-0.15	0.84	0.89
Grok 4.1	-0.53	0.91	0.72	0.71	0.94	1.00	0.75	0.87	-0.22	0.81	0.87
Grok 4.1 *	-0.59	0.91	0.89	0.88	0.78	0.75	1.00	0.72	0.27	0.69	0.78
Sonar	-0.75	0.86	0.65	0.67	0.96	0.87	0.72	1.00	-0.29	0.92	0.84
Sonar *	-0.20	-0.02	0.15	0.16	-0.15	-0.22	0.27	-0.29	1.00	-0.33	0.14
Sonnet 4.5	-0.76	0.81	0.66	0.70	0.84	0.81	0.69	0.92	-0.33	1.00	0.79
Sonnet 4.5 *	-0.61	0.87	0.77	0.81	0.89	0.87	0.78	0.84	0.14	0.79	1.00
	Experts	GPT 5.1	GPT 5.1 *	Flash 2.5	Flash 2.5 *	Grok 4.1	Grok 4.1 *	Sonar	Sonar *	Sonnet 4.5	Sonnet 4.5 *