	Resource Intensity	Scalability	Automation Potential	Real-World Applicability	Numerical Reducibility
Traditional Benchmarks	✓	✓	/	X	✓
Challenges and Competition	0	0	0	✓	0
Red Teaming and Capability Discovery	X	X	0	✓	X
Real-World Deployment Studies	X	X	×	/	X
Ablation Studies and Systematic (Stress) Testing	X	✓	✓	0	✓