Score ID	Criteria	Weight	Anthropic	OpenAI	Google DeepMind	Meta	x.AI
	1. Risk Identification	25%	28%	27%	17%	33%	5%
	1.1 Classification of Applicable Known Risks	**40%**	**30%**	**30%**	**18%**	**13%**	**13%**
C1	1.1.1 Risks from literature and taxonomies are well covered	50%	50%	50%	25%	25%	25%
C2	1.1.2 Exclusions are justified and documented	50%	10%	10%	10%	0%	0%
	1.2 Identification of Unknown Risks (Open-ended red teaming)	**20%**	**0%**	**3%**	**0%**	**0%**	**0%**
	1.2.1 Internal	70%	0%	3%	0%	0%	0%
C3	1.2.1.1 Adequate methodology (includes resources, time, and access to the model)	70%	0%	0%	0%	0%	0%
C4	1.2.1.2 Appropriate expertise to properly identify hazards	30%	0%	10%	0%	0%	0%
	1.2.2 Third parties	30%	0%	3%	0%	0%	0%
C5	1.2.2.1 Appropriate expertise to identify hazards	33%	0%	0%	0%	0%	0%
C6	1.2.2.2 Adequate resources/time/access to the model	33%	0%	10%	0%	0%	0%
C7	1.2.2.3 Commitment to non-interference with findings	34%	0%	0%	0%	0%	0%
	1.3 Risk Modeling	**40%**	**41%**	**36%**	**25%**	**69%**	**1%**
C8	1.3.2 The company uses risk models for all the risk domains identified, and the risk models are published (with potentially dangerous information redacted)	40%	50%	75%	50%	90%	0%
	1.3.1 Risk modeling methodology	40%	40%	10%	12%	58%	2%
C9	1.3.1.1 Methodology precisely defined	70%	50%	10%	10%	75%	0%
C10	1.3.1.2 Mechanism to incorporate red teaming findings	15%	10%	10%	10%	10%	0%
C11	1.3.1.3 Prioritization of severe and probable risks	15%	25%	10%	25%	25%	10%
C12	1.3.4 Third-party validation of risk models	20%	25%	10%	0%	50%	0%