

Mistral-Nemo-Instruct-2407								LLaMA 3.1 Instruct-8B								Gemini 2.5 Flash								Qwen3-8B								Qwen3-14B							
EN	ORG	100%	97.0%	82.4%	68.1%	51.9%	45.5%	22.6%	100%	87.3%	55.6%	59.4%	54.2%	27.6%	13.0%	100%	95.1%	80.0%	70.4%	58.3%	46.4%	9.4%	100%	94.1%	87.3%	68.9%	50.9%	38.2%	14.6%	100%	100%	97.5%	83.0%	70.6%	78.9%	20.5%			
	SS	100%	94.3%	81.0%	65.0%	61.8%	51.5%	6.7%	100%	100%	98.9%	93.4%	71.9%	58.9%	30.8%	100%	94.5%	73.2%	74.1%	55.6%	23.8%	9.3%	100%	85.8%	62.5%	48.8%	47.2%	29.6%	2.0%	100%	79.1%	66.7%	42.9%	40.0%	36.8%	0.0%			
	SO	100%	90.5%	77.7%	52.3%	55.6%	36.7%	12.8%	100%	97.3%	90.6%	82.3%	69.4%	56.9%	21.1%	100%	85.3%	69.6%	64.6%	47.8%	30.8%	2.9%	100%	84.1%	68.7%	48.8%	40.0%	26.9%	5.1%	100%	83.0%	59.6%	53.1%	53.3%	18.2%	5.0%			
	PSS	100%	93.4%	91.7%	71.4%	60.7%	43.5%	9.6%	100%	97.5%	96.4%	88.4%	84.8%	74.3%	27.1%	100%	92.5%	89.8%	47.6%	30.0%	20.0%	10.0%	100%	87.0%	81.0%	58.1%	36.0%	20.0%	8.8%	100%	82.2%	75.7%	58.3%	30.8%	14.3%	0.0%			
	POS	100%	93.1%	90.2%	70.2%	50.0%	72.7%	8.6%	100%	94.4%	96.7%	93.0%	89.6%	61.1%	23.9%	100%	89.2%	79.3%	48.1%	66.7%	36.4%	11.1%	100%	90.9%	77.3%	54.5%	33.3%	21.4%	13.3%	100%	83.6%	69.7%	75.0%	30.8%	8.3%	0.0%			
DE	ORG	100%	93.0%	73.0%	63.8%	37.0%	27.3%	16.1%	100%	95.5%	82.2%	68.8%	45.8%	48.3%	4.3%	100%	92.7%	73.3%	63.0%	45.8%	21.4%	10.6%	100%	92.9%	71.4%	55.6%	39.6%	20.6%	9.2%	100%	93.3%	90.0%	59.6%	41.2%	31.6%	11.3%			
	SS	100%	91.1%	86.0%	66.7%	50.0%	27.3%	10.0%	100%	91.4%	80.7%	42.6%	45.6%	26.8%	8.3%	100%	87.2%	76.8%	55.6%	48.1%	38.1%	7.0%	100%	88.7%	81.9%	61.0%	55.6%	33.3%	8.0%	100%	79.1%	64.4%	57.1%	50.0%	31.6%	5.9%			
	SO	100%	90.5%	79.8%	61.5%	41.7%	26.7%	20.5%	100%	91.8%	77.6%	59.7%	40.3%	33.3%	6.8%	100%	90.8%	71.4%	62.5%	43.5%	30.8%	14.3%	100%	85.6%	76.1%	48.8%	36.7%	42.3%	12.8%	100%	87.5%	68.1%	56.2%	60.0%	18.2%	5.0%			
	PSS	100%	91.0%	84.4%	60.3%	35.7%	39.1%	8.2%	100%	89.9%	81.1%	63.8%	34.8%	25.7%	5.1%	100%	90.6%	85.7%	66.7%	50.0%	40.0%	10.0%	100%	94.0%	79.4%	67.4%	76.0%	35.0%	4.4%	100%	84.4%	73.0%	58.3%	61.5%	35.7%	3.4%			
	POS	100%	95.7%	84.8%	57.4%	47.2%	31.8%	6.9%	100%	93.0%	86.2%	59.2%	45.8%	22.2%	8.8%	100%	90.4%	69.0%	66.7%	33.3%	18.2%	7.4%	100%	92.6%	81.8%	60.6%	66.7%	39.3%	13.3%	100%	87.7%	75.8%	50.0%	38.5%	50.0%	13.0%			
FR	ORG	100%	93.0%	91.9%	70.2%	59.3%	40.9%	14.5%	100%	90.0%	82.2%	46.9%	54.2%	17.2%	4.3%	100%	93.9%	88.9%	66.7%	62.5%	46.4%	10.6%	100%	95.3%	85.7%	73.3%	62.3%	38.2%	6.2%	100%	96.6%	91.2%	78.7%	50.0%	39.5%	12.6%			
	SS	100%	89.4%	74.0%	60.0%	44.1%	30.3%	15.0%	100%	91.4%	92.0%	73.8%	52.6%	35.7%	8.3%	100%	82.6%	69.6%	48.1%	37.0%	38.1%	4.7%	100%	84.0%	72.2%	63.4%	50.0%	18.5%	14.0%	100%	81.4%	73.3%	57.1%	40.0%	15.8%	5.9%			
	SO	100%	83.8%	70.2%	75.4%	36.1%	43.3%	15.4%	100%	94.5%	90.6%	71.0%	54.8%	35.3%	12.0%	100%	81.7%	75.0%	64.6%	34.8%	23.1%	11.4%	100%	86.4%	70.1%	68.3%	33.3%	26.9%	7.7%	100%	81.2%	78.7%	59.4%	53.3%	18.2%	25.0%			
	PSS	100%	91.0%	76.0%	69.8%	28.6%	26.1%	9.6%	100%	89.9%	95.5%	69.6%	50.0%	28.6%	9.3%	100%	88.7%	61.2%	66.7%	50.0%	70.0%	3.3%	100%	81.0%	74.6%	51.2%	64.0%	25.0%	7.4%	100%	80.0%	73.0%	58.3%	61.5%	28.6%	6.9%			
	POS	100%	94.0%	83.7%	66.0%	50.0%	27.3%	12.1%	100%	95.8%	91.1%	74.6%	52.1%	44.4%	12.4%	100%	86.7%	79.3%	55.6%	50.0%	45.5%	7.4%	100%	86.0%	77.3%	78.8%	23.8%	42.9%	6.7%	100%	82.2%	69.7%	56.2%	38.5%	16.7%	8.7%			
ES	ORG	100%	93.0%	75.7%	51.1%	18.5%	9.1%	3.2%	100%	94.5%	64.4%	43.8%	33.3%	6.9%	0.0%	100%	87.8%	82.2%	51.9%	41.7%	28.6%	5.9%	100%	96.5%	87.3%	66.7%	41.5%	41.2%	8.5%	100%	100%	92.5%	83.0%	55.9%	31.6%	7.3%			
	SS	100%	96.7%	89.0%	75.0%	58.8%	51.5%	18.3%	100%	96.6%	98.9%	91.8%	71.9%	55.4%	17.9%	100%	94.5%	83.9%	59.3%	55.6%	33.3%	14.0%	100%	96.2%	80.6%	61.0%	36.1%	40.7%	14.0%	100%	90.7%	71.1%	61.9%	50.0%	26.3%	5.9%			
	SO	100%	94.3%	87.2%	58.5%	61.1%	40.0%	15.4%	100%	100%	91.8%	82.3%	66.1%	49.0%	13.5%	100%	93.6%	82.1%	58.3%	30.4%	53.8%	5.7%	100%	93.2%	73.1%	65.9%	53.3%	34.6%	7.7%	100%	89.3%	76.6%	37.5%	46.7%	54.5%	10.0%			
	PSS	100%	96.7%	88.5%	76.2%	71.4%	34.8%	13.7%	100%	97.5%	98.2%	85.5%	71.7%	62.9%	7.6%	100%	84.9%	83.7%	47.6%	40.0%	40.0%	0.0%	100%	92.0%	84.1%	67.4%	64.0%	55.0%	11.8%	100%	90.0%	67.6%	50.0%	30.8%	28.6%	13.8%			
	POS	100%	99.1%	93.5%	68.1%	47.2%	36.4%	12.1%	100%	100%	97.6%	87.3%	56.2%	63.9%	15.0%	100%	91.6%	89.7%	55.6%	50.0%	18.2%	22.2%	100%	95.0%	79.5%	57.6%	47.6%	21.4%	11.1%	100%	95.9%	75.8%	50.0%	53.8%	33.3%	4.3%			
IT	ORG	100%	94.0%	82.4%	51.1%	51.9%	18.2%	11.3%	100%	91.8%	84.4%	62.5%	37.5%	13.8%	13.0%	100%	95.1%</																						