

# MoDS: Moderating a Mixture of Document Speakers to Summarize Debatable Queries in Document Collections



**Nishant Balepur**

Tong Sun

Alexa Siu

Nedim Lipka

Franck Dernoncourt

Jordan Boyd-Graber

Puneet Mathur



Paper

ChatGPT struggles with our new task of **debatable query-focused summarization...**

So we build a **multi-LLM summarization** framework to beat existing methods!

Document Set: [1] [2] [3] [4] [5] [6]

Debatable Query: Is EU expansion and membership itself a good idea?

Debatable Query-Focused Summary (2 topics)

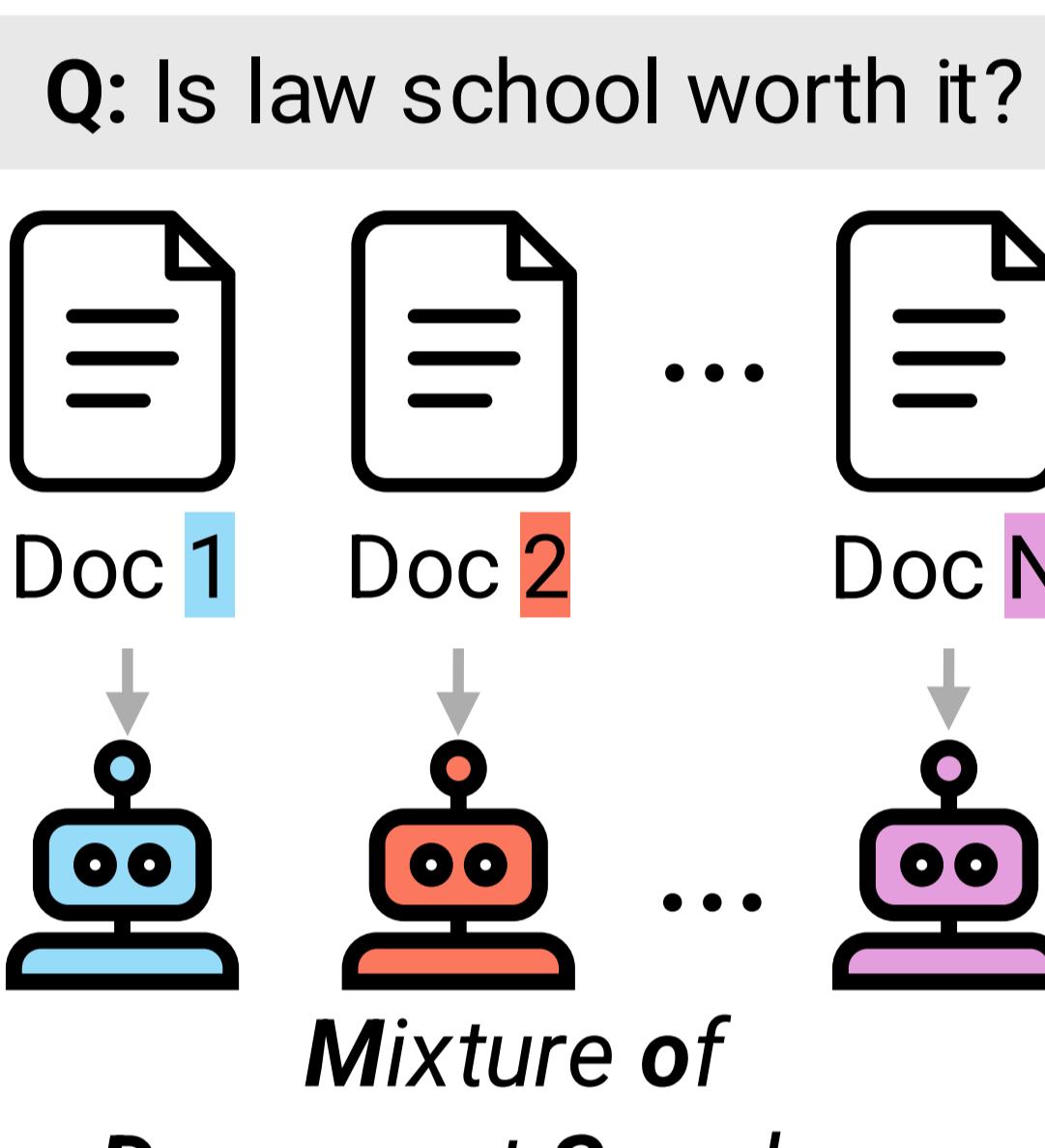
**Topic 1: Economic Growth**

EU expansion is expected to boost the GDP growth of acceding countries by 1.3-2.1% annually, while existing members could see a cumulative GDP increase of 0.7 percentage points [1]. Studies suggest that the economic gains from the accession of Central and Eastern European countries could reach €10 billion for the EU-15 and €23 billion for new members [1]. The accession is also seen as an opportunity for creating new markets and stabilizing the continent economically [2].

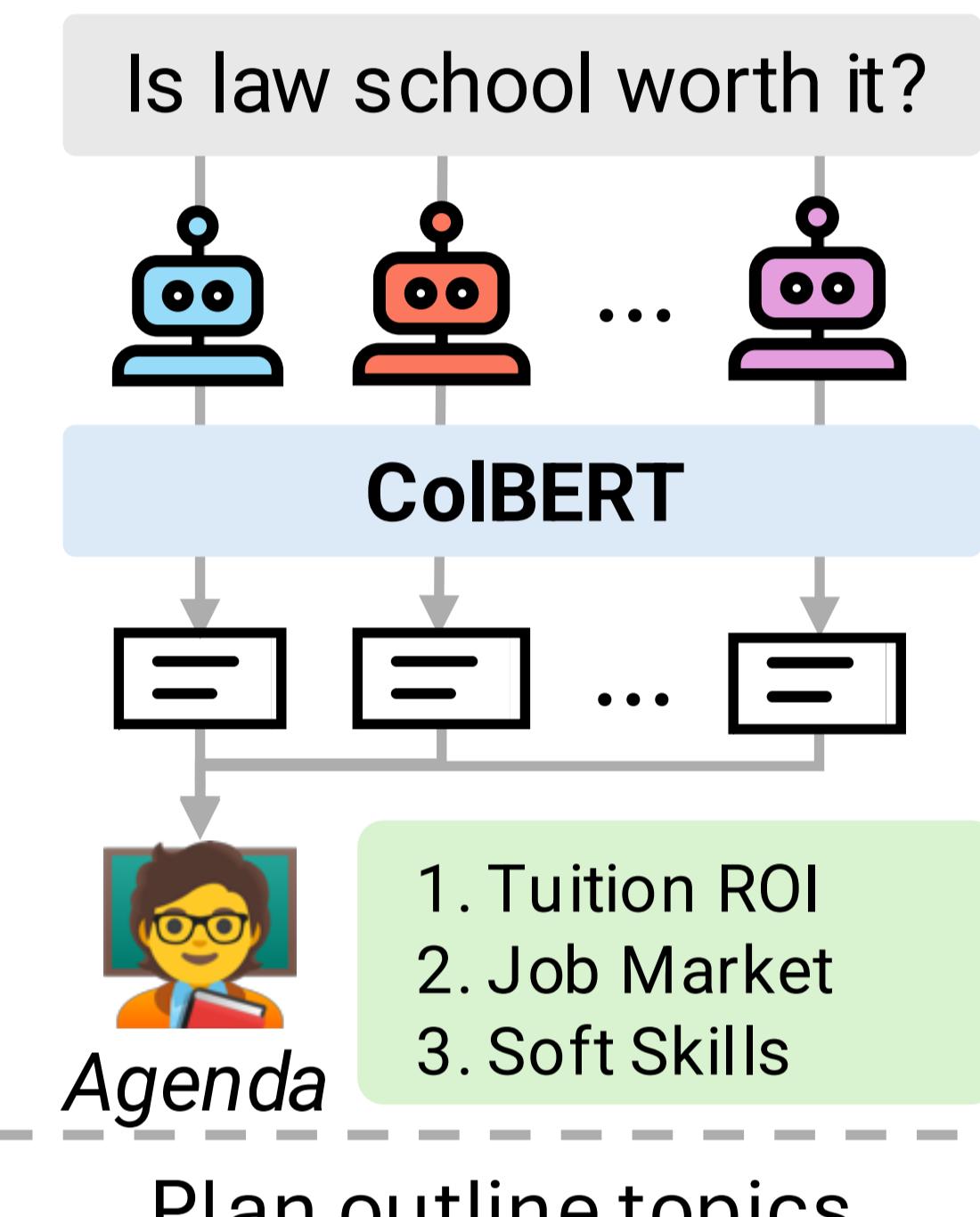
Imbalanced and Low Coverage

## MoDS: Moderating a Mixture of Document Speakers

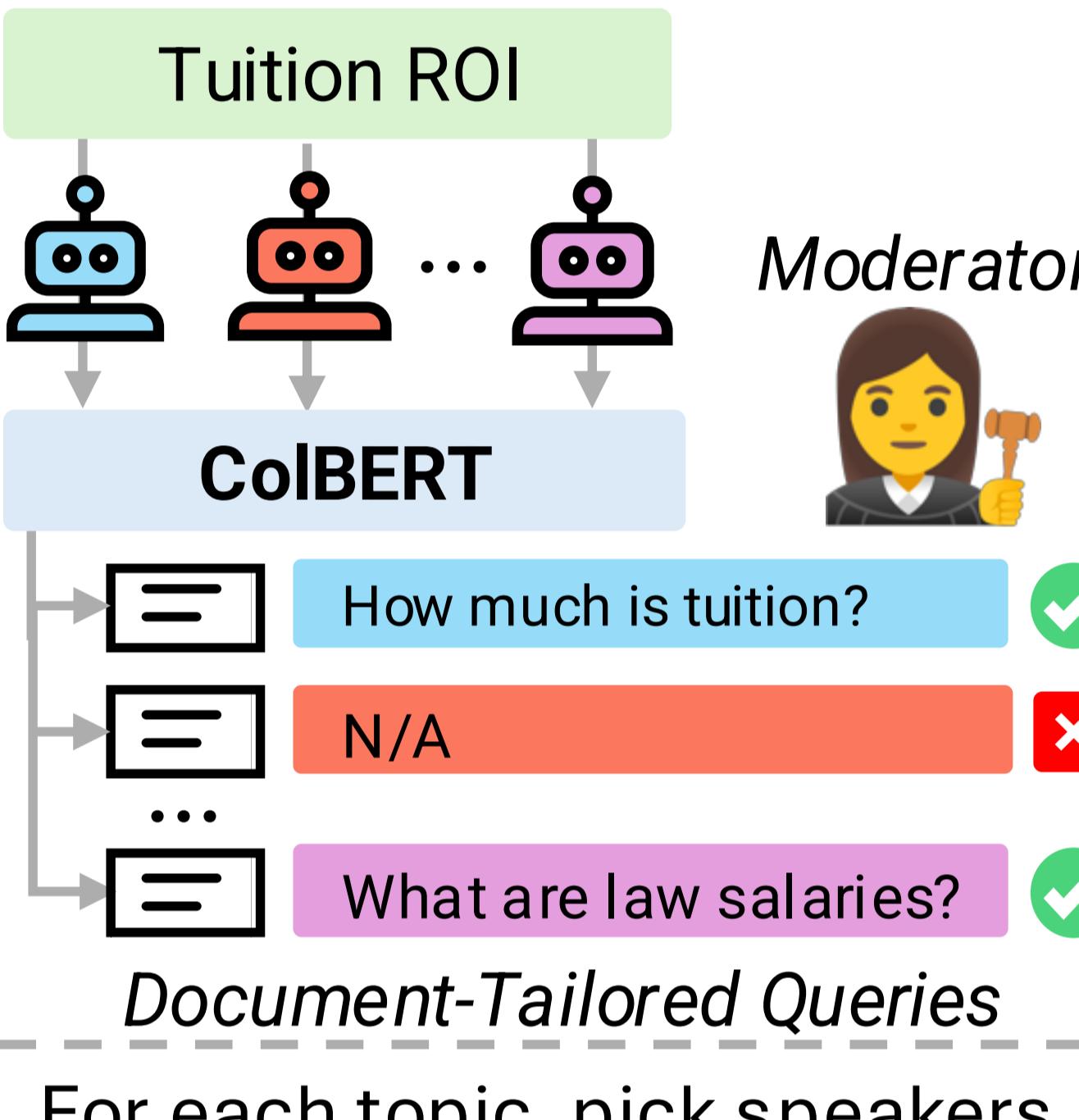
### Task Inputs (§3)



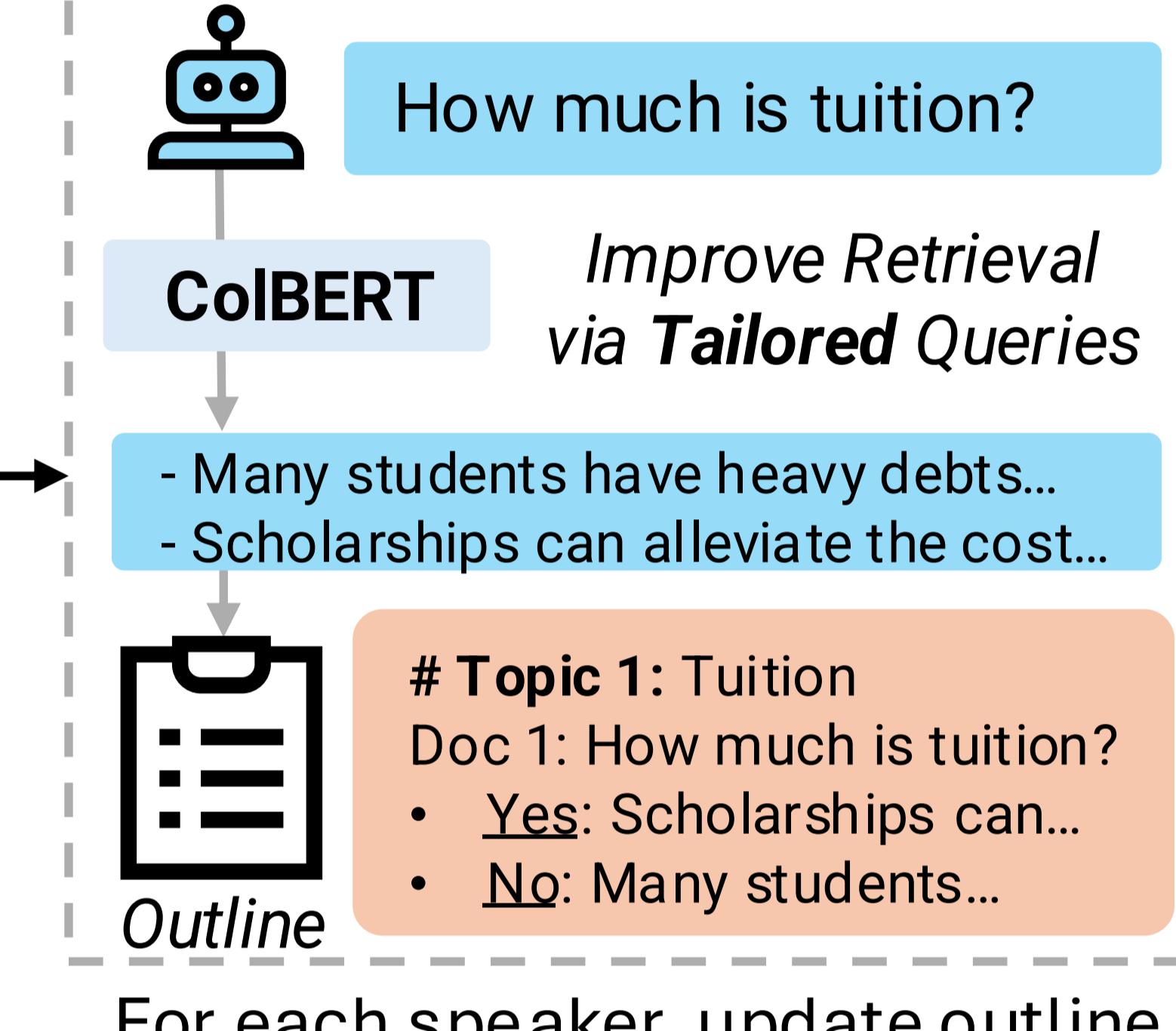
### Agenda Planning (§4.1)



### Speaker Selection (§4.2)



### Speaker Discussion (§4.3)



Make one speaker per doc

Plan outline topics

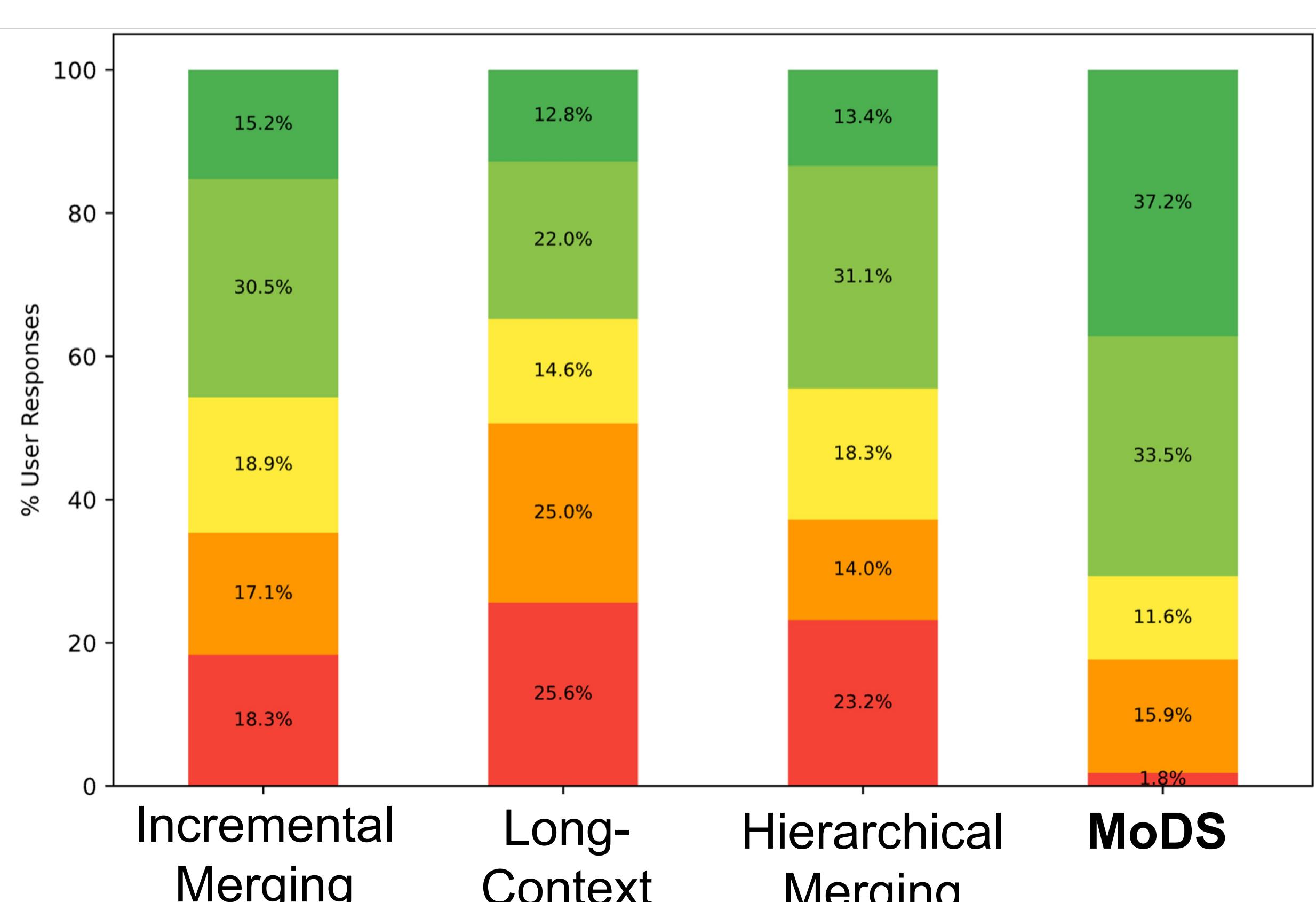
For each topic, pick speakers

For each speaker, update outline

**Finding 1:** MoDS beats all baselines in citation coverage and balance for summaries + topic paragraphs

# Top.	Model	Overall Summary			Topic Paragraphs			Cite Acc. (↑)	All / Avg Sents
		DC (↑)	Fair (↓)	Faithful (↓)	DC (↑)	Fair (↓)	Faithful (↓)		
3	MoDS-Topic (Ours)	0.8724*	0.0701*	0.0235*	0.6066*	0.1255*	0.0789*	0.982	8.99 / 3.00
	MoDS-All (Ours)	0.8457*	0.0786*	0.0273*	0.5508	0.1463*	0.0938*	0.987	8.87 / 2.96
	Long-Context	0.5877	0.2094	0.1790	0.2798	0.4336	0.4028	0.953	9.02 / 3.01
	RAG-All	0.6125	0.1544	0.1040	0.3229	0.3176	0.2701	0.997	9.01 / 3.00
	RAG-Doc	0.7171	0.1180	0.0664	0.3504	0.3233	0.2748	0.961	9.01 / 3.00
	Hierarchical	0.7868	0.0907	0.0374	0.3639	0.2980	0.2452	0.983	9.02 / 3.01
	Incremental-All	0.5566	0.2579	0.2089	0.3919	0.3243	0.2765	0.950	8.91 / 2.97
	Incremental-Topic	0.6152	0.2415	0.1970	0.4707	0.3128	0.2674	0.954	9.03 / 3.01
	Cluster	0.7102	0.1106	0.0725	0.3632	0.3106	0.2737	0.931	9.04 / 3.01
	RAG+Cluster	0.6811	0.1405	0.0894	0.3428	0.3200	0.2689	0.977	9.01 / 3.00

**Finding 2:** Humans find our summaries more balanced



**Finding 3:** Intermediate outlines are useful!

