

# Augmenting Large Language Models with Long-term Memory and Multimodal Capabilities

Rogerio Schmidt Feris  
Principal Scientist and Manager  
MIT-IBM Watson AI Lab, IBM Research

March 2024

# Humans have a remarkable ability to remember information over long time horizons (across modalities)

We can remember the name of a childhood friend ...

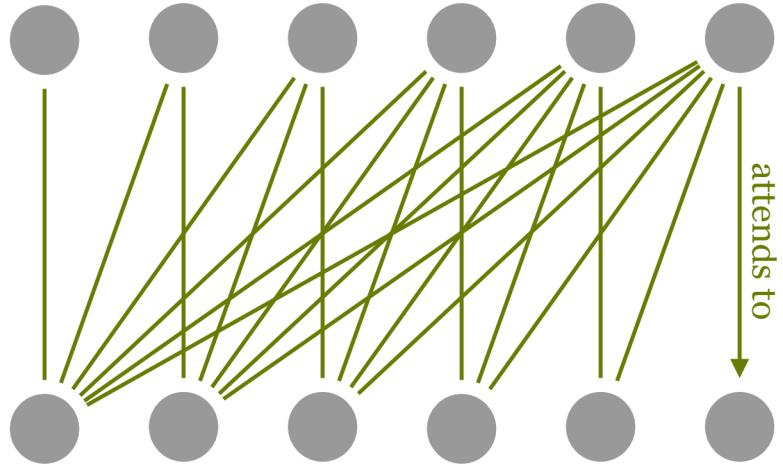


We recall historical facts from decades ago



We retrieve memories of songs across various time periods

# Large Language Models (LLMs) struggle to handle long input sequences

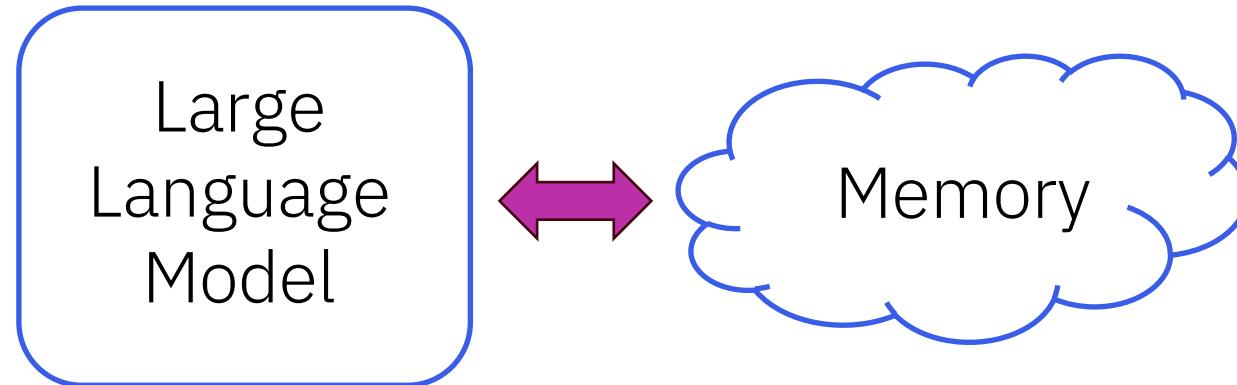


Memory/Computation complexity grows quadratically with sequence length due to transformer self-attention

Many long-range attention methods exist but cannot handle unlimited context window sizes and fail to retrieve information in the middle of the input [Liu et al., 2023]

# Promising Solution: Memory-Augmented LLMs

[Wu et al., 2022], [Tworkowski et al., 2023], [Wang et al., 2023], ...



## Limitations of existing approaches:

- Context window size is limited by the number of memory slots
- Require costly LLM fine-tuning / re-training
- Memory staleness

# Desiderata

We want to design a memory module that is:

- Simple to implement
- Training-free (plug-and-play integration with a frozen LLM)
- Capable of handing unbounded (“infinite”) long sequences
- Complementary to other efficient self-attention methods

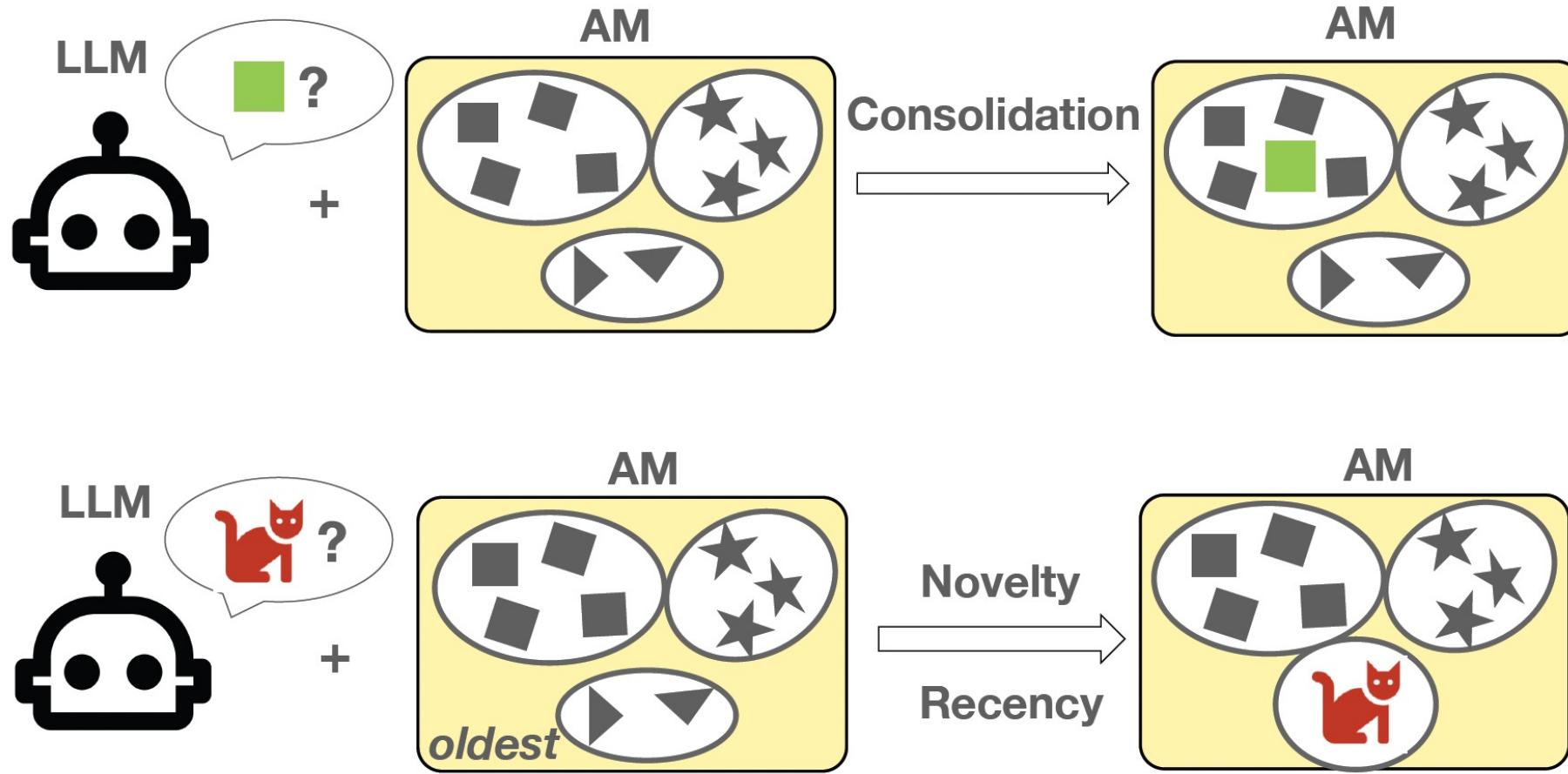
# Proposed Approach: CAMELoT

(Consolidated **A**ssociative **M**emory **E**nhan<sup>c</sup>ed **L**ong **T**ransformer)

Zexue He, Leonid Karlinsky, Donghyun Kim, Julian McAuley, Dmitry Krotov, and Rogerio Feris (Arxiv 2024)

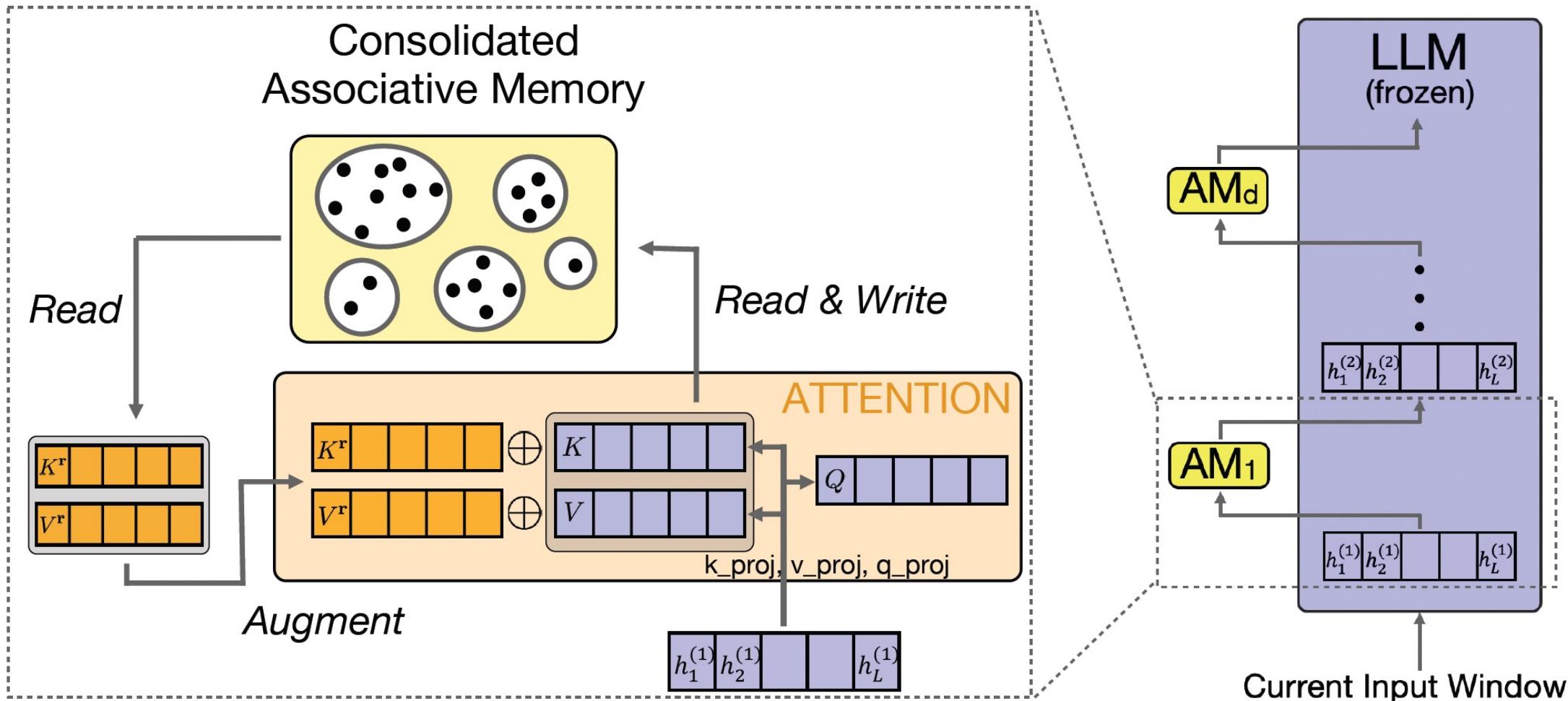


# Proposed Approach: CAMELoT

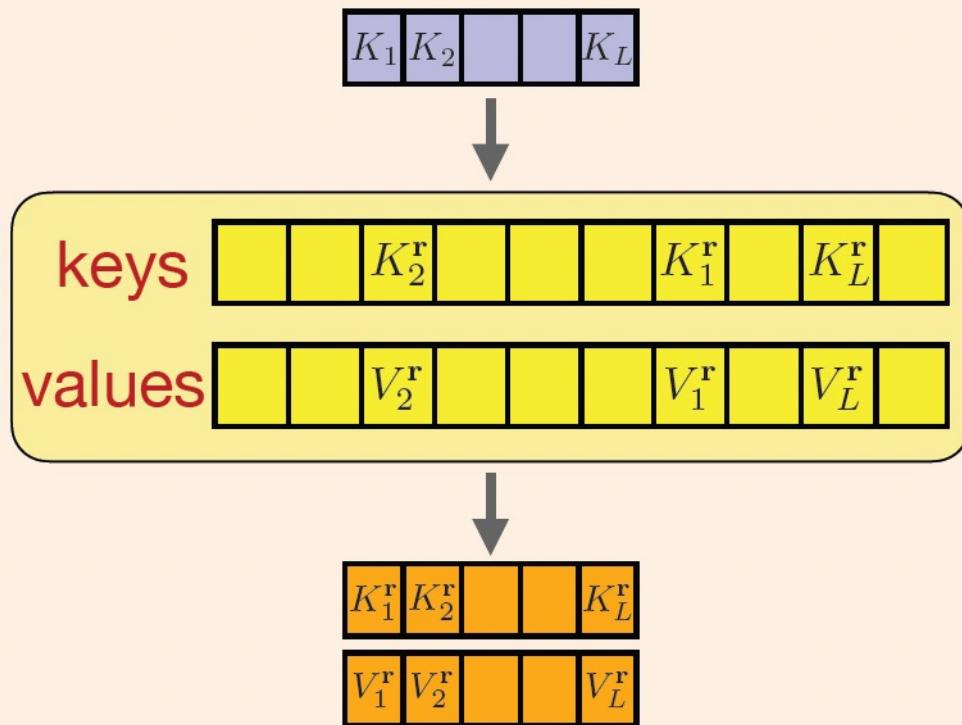


AM = Associative Memory

# CAMELoT Architecture



# Associative Memory – Read Operation



**Read**( $[K_1, \dots, K_L]$ , Memory)

$$K^r \leftarrow \emptyset, V^r \leftarrow \emptyset$$

**for**  $i = 1 \dots L$

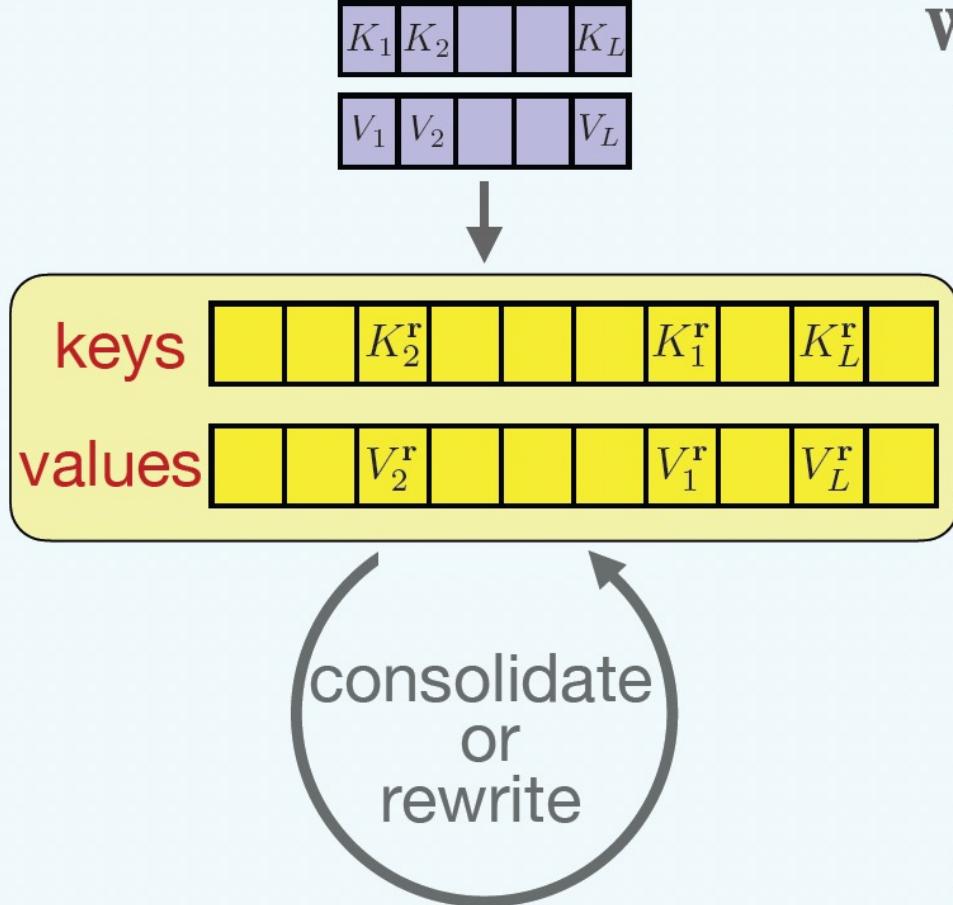
$$\hat{\mu}(i) = \underset{\mu}{\operatorname{argmax}} [\operatorname{sim}(K_{\mu}^{\text{mem}}, K_i)]$$

$$K^r \leftarrow K^r \cup K_{\hat{\mu}(i)}^{\text{mem}}$$

$$V^r \leftarrow V^r \cup V_{\hat{\mu}(i)}^{\text{mem}}$$

**return**  $K^r, V^r$

# Associative Memory – Write Operation



**Write**( $[K_1, \dots, K_L]$ ,  $[V_1, \dots, V_L]$ , **Memory**)

**for**  $i = 1 \dots L$

$$\hat{\mu}(i) = \underset{\mu}{\operatorname{argmax}} [\operatorname{sim}(K_{\mu}^{\text{mem}}, K_i)]$$

**if**  $\operatorname{sim}(K_{\hat{\mu}(i)}^{\text{mem}}, K_i) > R$

$$K_{\hat{\mu}(i)}^{\text{mem}} \leftarrow \varepsilon_{\hat{\mu}(i)} K_i + (1 - \varepsilon_{\hat{\mu}(i)}) K_{\hat{\mu}(i)}^{\text{mem}}$$

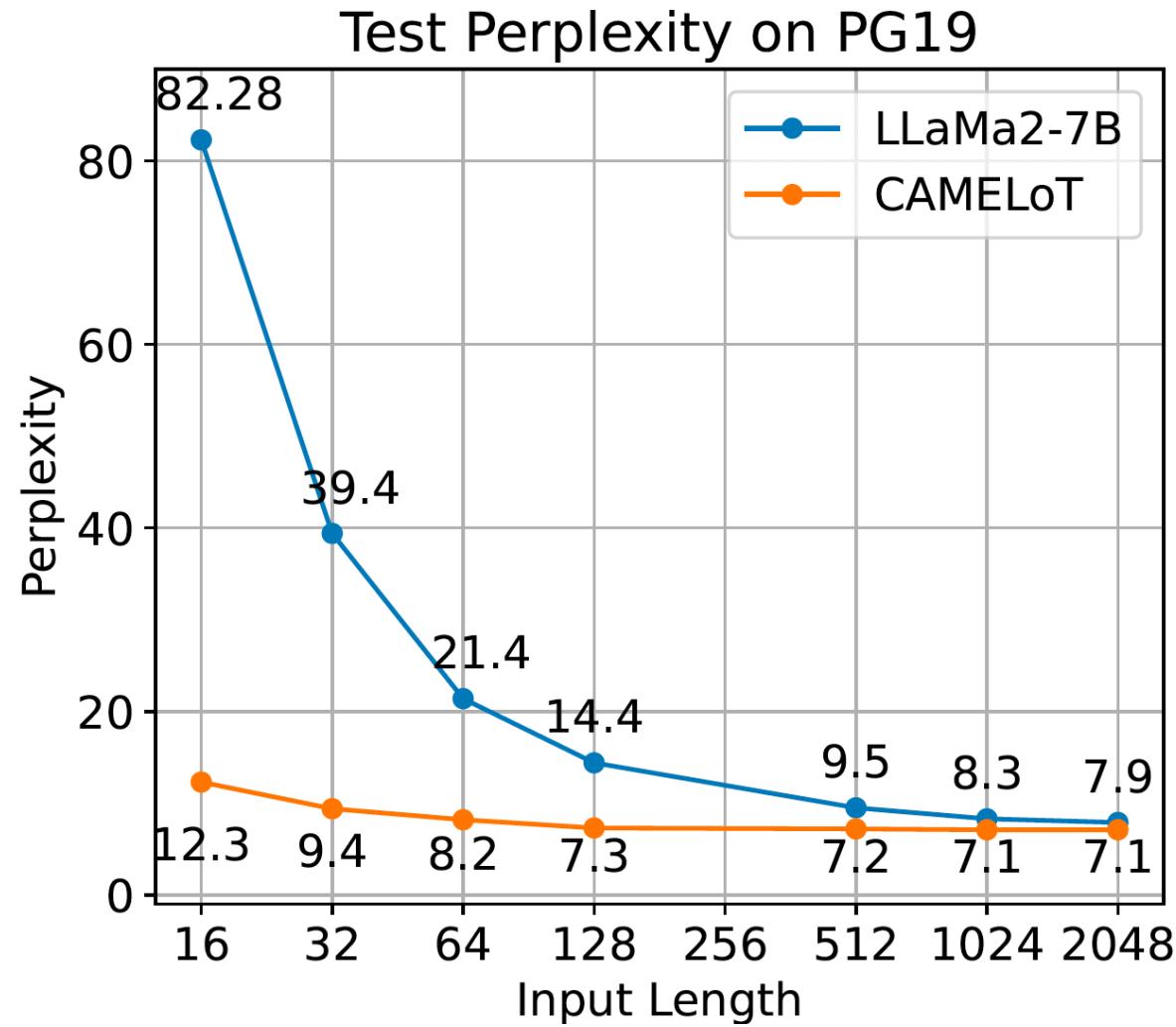
$$V_{\hat{\mu}(i)}^{\text{mem}} \leftarrow \varepsilon_{\hat{\mu}(i)} V_i + (1 - \varepsilon_{\hat{\mu}(i)}) V_{\hat{\mu}(i)}^{\text{mem}}$$

**else**

rewrite the oldest used slot  
with  $K_i, V_i$

**return** **Memory**

# Experimental Results



Strong performance even  
with a tiny context window

# Experimental Results

## Language modeling perplexity on PG-19, Arxiv, and wikitext-103

		PG-19 and Arxiv				wikitext-103		
		Input Length	Retrieved Mem.	PG-19	Arxiv	Input Length	Retrieved Mem.	Wikitext-103
LLaMa2-7B	512	None	9.54	5.99		512	None	16.0
	1024	None	8.33	4.98		1024	None	14.80
	2048	None	7.88	4.35		2048	None	14.46
	Avg	-	8.58	5.12		Avg	-	15.09
Transformer-XL	512	512	8.44	4.15		256	256	15.02
	1024	1024	8.27	3.81		512	512	14.21
	2048	2048	7.86	3.65		1024	1024	14.2
	Avg	-	8.19	3.87		Avg	-	14.48
Memorizing Transformers	512	512	8.12	3.82		256	256	14.18
	1024	1024	7.4	3.63		512	512	14.07
	2048	2048	7.34	3.62		1024	1024	14.39
	Avg	-	7.62	3.69		Avg	-	14.21
CAMELoT	512	512	7.24	3.61		256	256	14.06
	1024	1024	7.14	<b>3.60</b>		512	512	<b>14.00</b>
	2048	2048	<b>7.10</b>	<b>3.60</b>		1024	1024	14.34
	Avg	-	7.16	3.60		Avg	-	14.13

Up to 30%  
perplexity  
reduction  
on Arxiv

# Experimental Results

In-Context Learning (TREC dataset)  
with more demonstrations

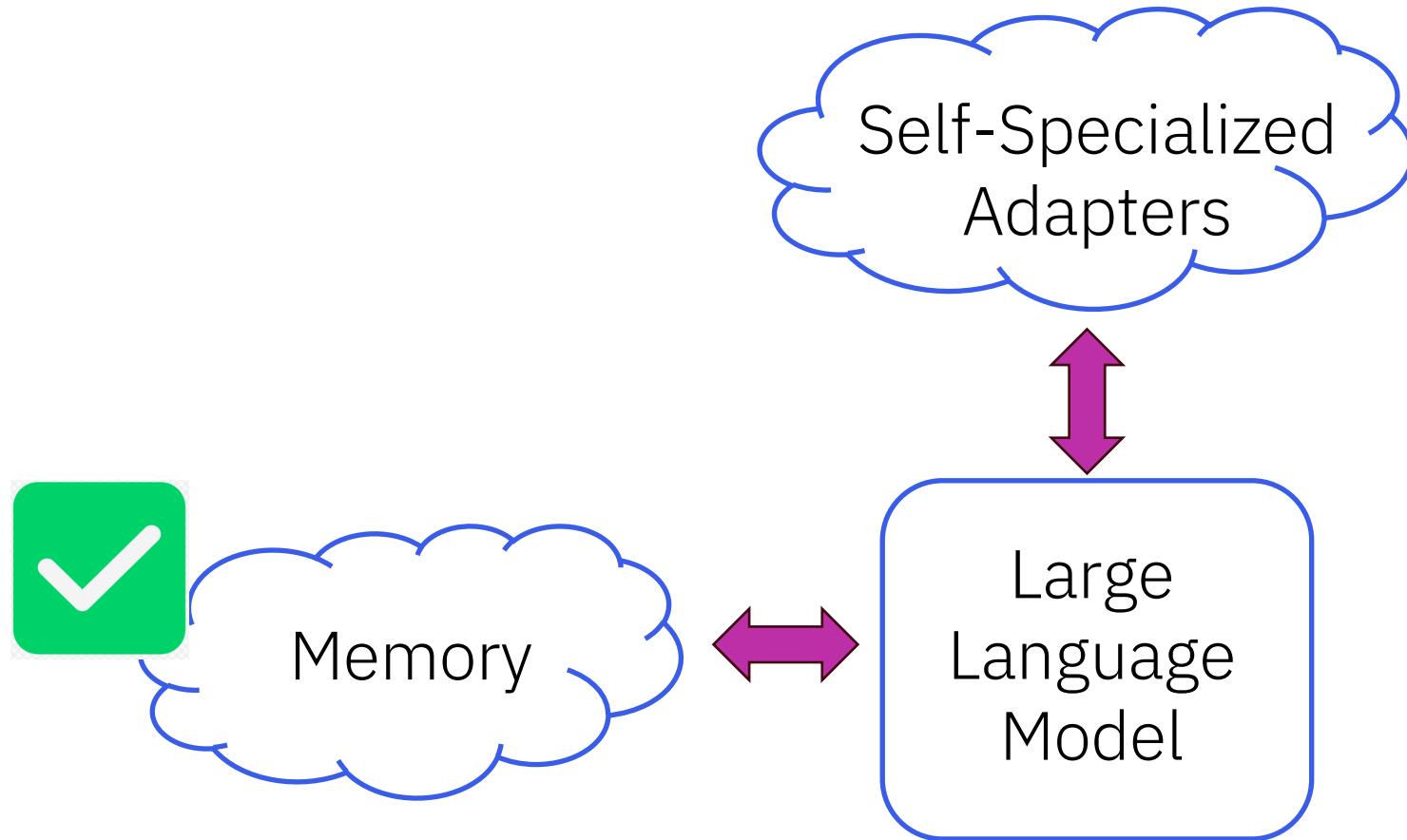
	Prompt Length	Accuracy
LLaMa2-7b	13 (0-shot)	0.15
CAMELoT	2k	0.21
	4k	0.25
	10k	0.42
	20k	0.44

Better performance for infrequent words

	Frequency			
	>10K	1K - 10K	100 - 1K	<100
LLaMa2-7B	2.75	5.08	9.77	25.96
CAMELoT	2.13	4.11	7.54	19.4
Relative Gain Over LLaMa2	22.1%	19.3%	22.8%	25.3%

Ablation study

Models	PPL
LLaMa2-7B	7.30
CAMELoT	<b>6.85</b>
CAMELoT w/o Read	> 20
CAMELoT w/o Recency	9.25
CAMELoT w/o Novelty	7.23
CAMELoT w/o Consolidation	7.00



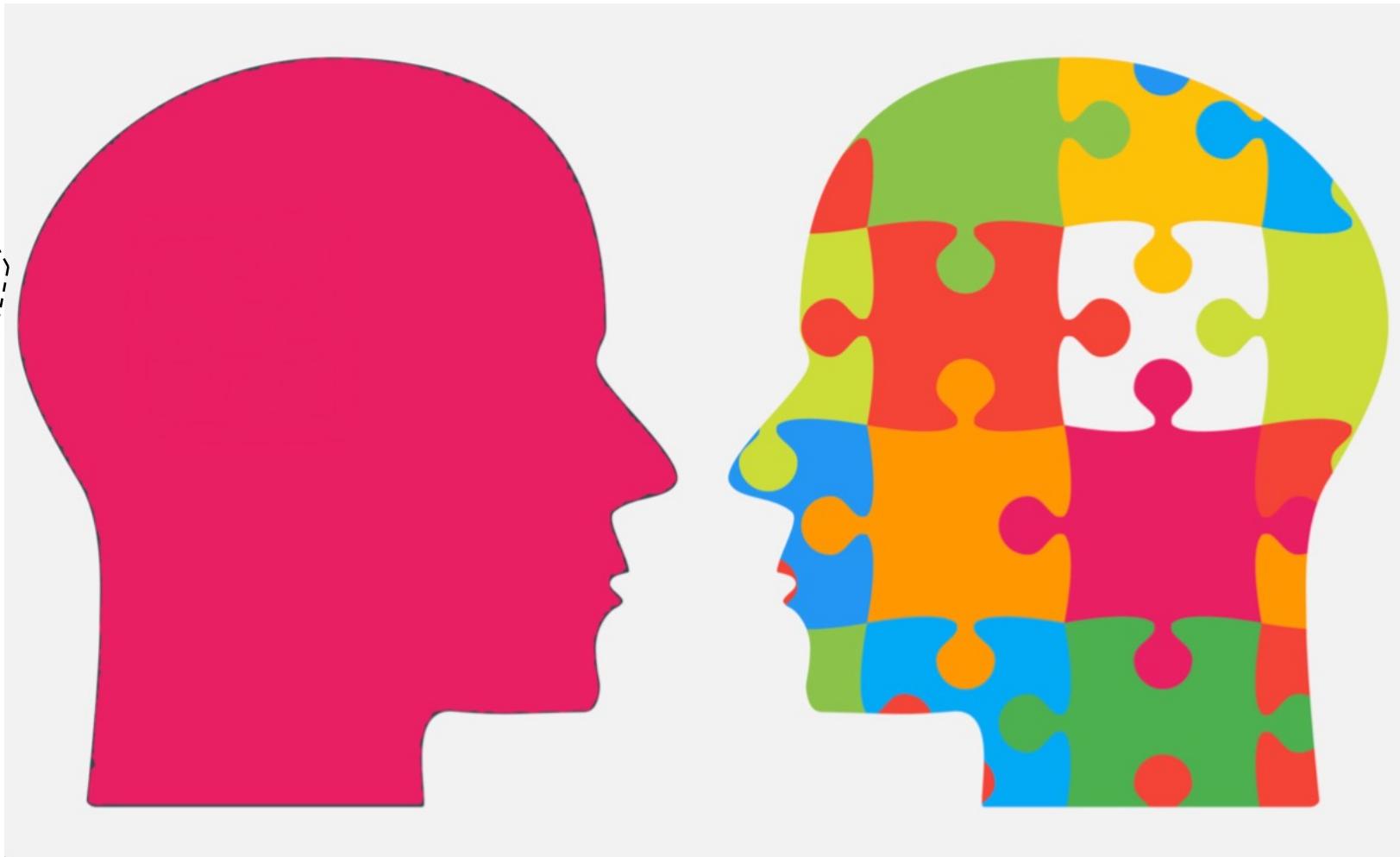
# Generalist



- Does well in many domains
- Requires more capacity (bigger models)

# Specialist

- Best in a specific domain
- Requires less capacity (smaller models)



# Generalist

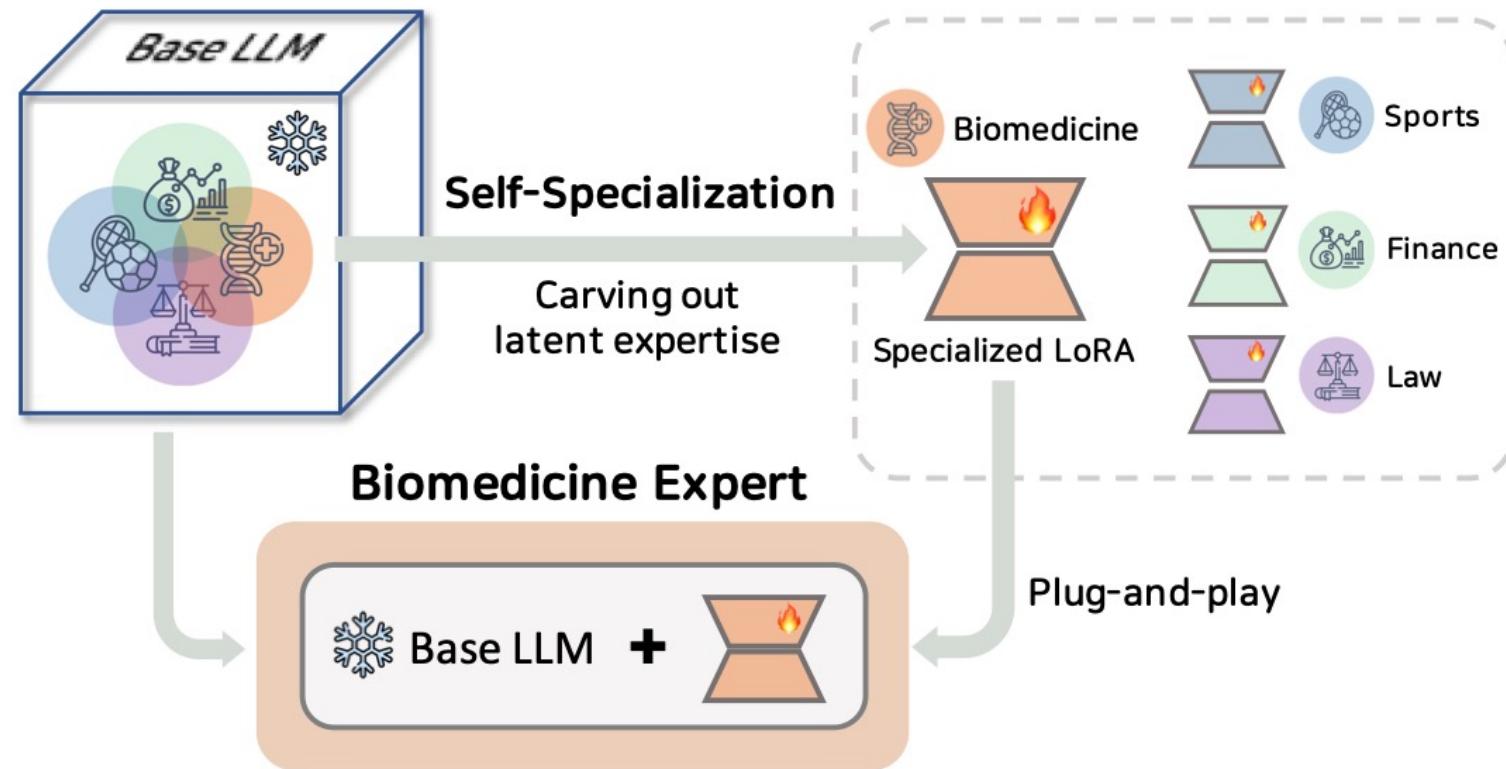
- Does well in many domains
- Requires more capacity (bigger models)

Can we adapt a general large language model to perform better in a specific domain?

- Naive Approach: Instruction fine-tuning, but it requires time-consuming and difficult to scale manual annotation

# SELF-SPECIALIZATION: UNCOVERING LATENT EXPERTISE WITHIN LARGE LANGUAGE MODELS

**Junmo Kang<sup>\*1</sup>**   **Hongyin Luo<sup>2</sup>**   **Yada Zhu<sup>3</sup>**   **James Glass<sup>2</sup>**   **David Cox<sup>3</sup>**  
**Alan Ritter<sup>1</sup>**   **Rogerio Feris<sup>3</sup>**   **Leonid Karlinsky<sup>3</sup>**



## Real-Data:

Domain-specific  
seed instructions  
(small set – e.g., 80)



### Instruction:

In this task, you are given a short article and question.  
Read the short article and answer the question.

### Input:

Short article: The extract from the opium poppy was tested on breast cancer cells and was found to inhibit the migration and invasion of breast cancer cells.

Question: Was the extract more potent in its inhibitory effect on the migration of breast cancer cells than its effect on the invasion of breast cancer cells?

## Real-Data:

Domain-specific  
seed instructions  
(small set – e.g., 80)



## Base LLM



## Synthetic Data:

Self-generation of  
domain-specific  
instructions

**Instruction:** Generate a list of drugs which can be used for the treatment of the given symptom.

**Instruction:** Given medication records, predict possible drug-drug interactions.

**Instruction:** You are given data of genetic variations and mutations, generate a comprehensive report.

⋮

**Instruction:** Provide an answer to the following question about the patient's medical history.

## Real-Data:

Domain-specific  
seed instructions  
(small set – e.g., 80)



## Base LLM



## Synthetic Data:

Self-generation of  
domain-specific  
instructions

**Instruction:** Generate a list of drugs which can be used for the treatment of the given symptom.

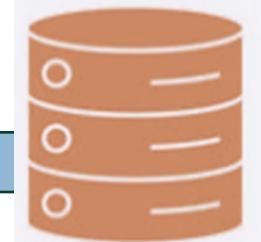
**Instruction:** Given medication records, predict possible drug-drug interactions.

**Instruction:** You are given data of genetic variations and mutations, generate a comprehensive report.

⋮

**Instruction:** Provide an answer to the following question about the patient's medical history.

Domain-specific  
knowledge  
(unlabeled)



(Instruction, Input, Output)

## Base LLM



## Real-Data:

Domain-specific  
seed instructions  
(small set – e.g., 80)



## Synthetic Data:

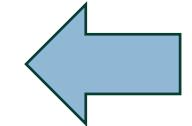
Self-generation of  
domain-specific  
instructions

Base LLM



## Specialized LLM

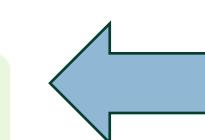
(e.g., Biomedicine)



Parameter-Efficient  
Fine Tuning (LoRA)



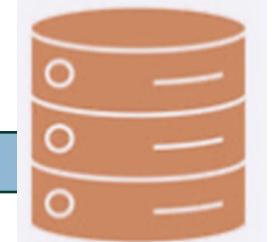
(Instruction, Input, Output)



Base LLM



Domain-specific  
knowledge  
(unlabeled)



**Instruction:** Generate a list of drugs which can be used for the treatment of the given symptom.

**Instruction:** Given medication records, predict possible drug-drug interactions.

**Instruction:** You are given data of genetic variations and mutations, generate a comprehensive report.

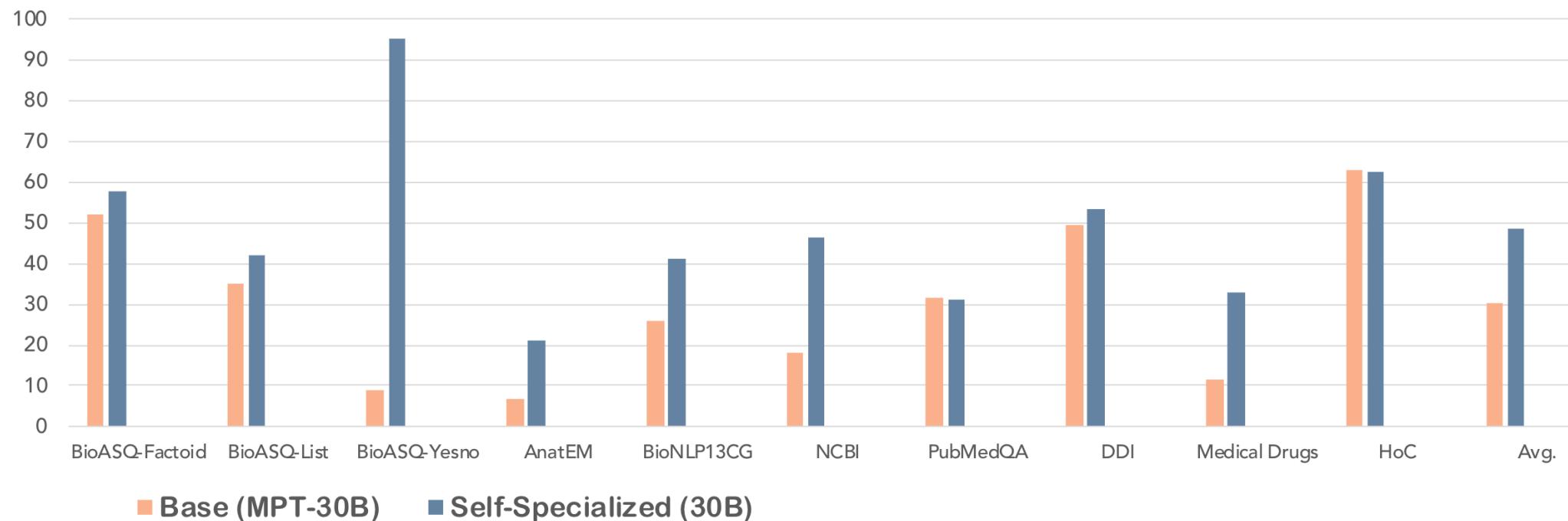
:

**Instruction:** Provide an answer to the following question about the patient's medical history.

# Main Results

- Our self-specialized model (30B) outperforms its base model, MPT-30B by a large margin (up to 18 points)

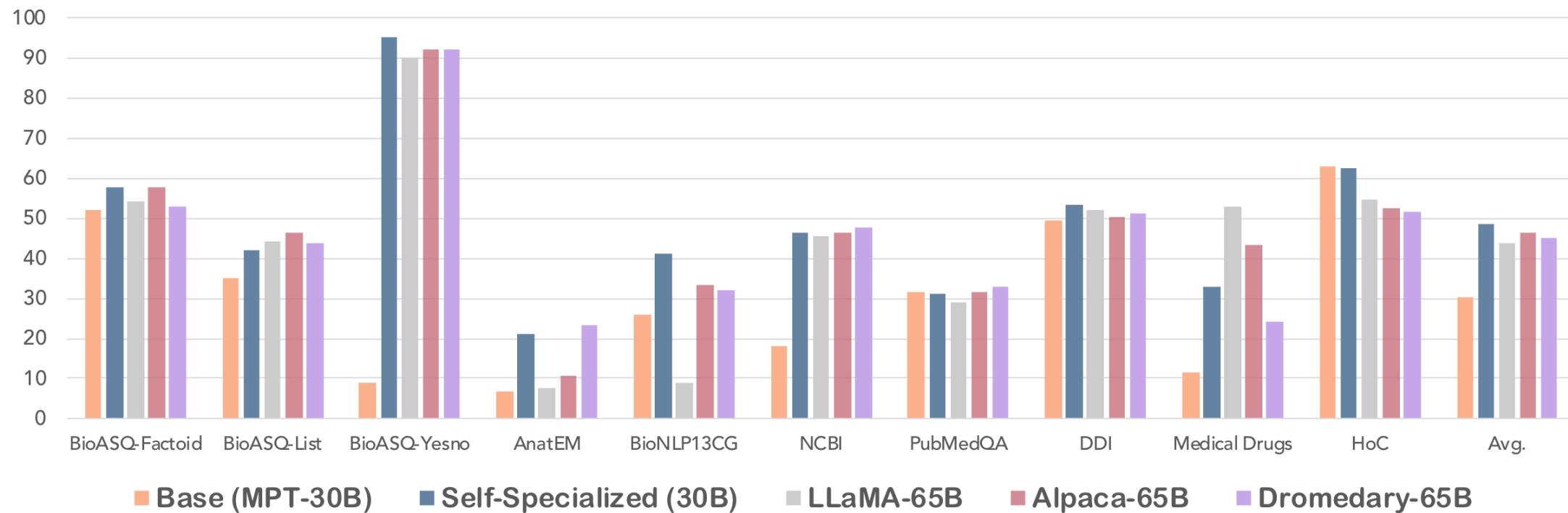
5-shot results in 10 datasets (biomedicine)

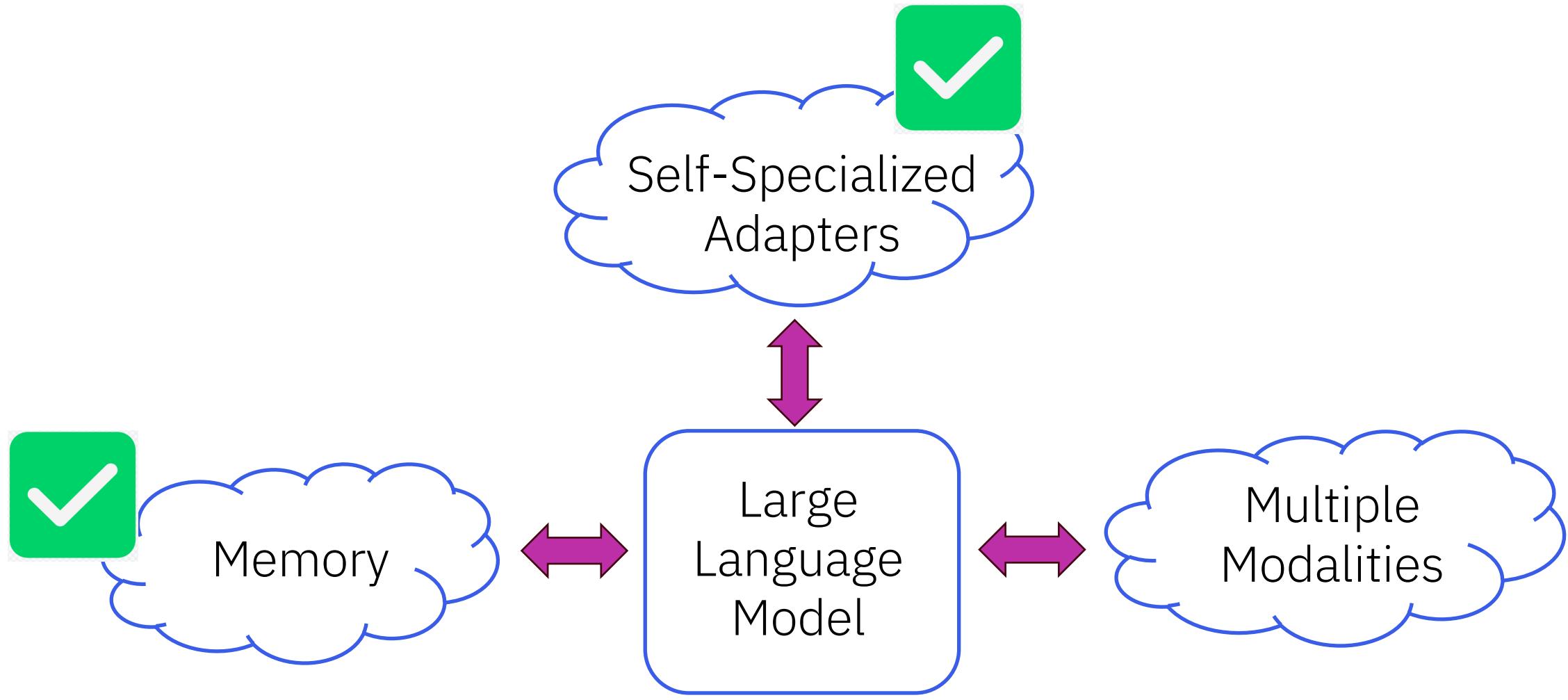


# Main Results

- Our self-specialized model (30B) even outperforms 65B models (2.2x larger), including Llama-65B

5-shot results in 10 datasets (biomedicine)



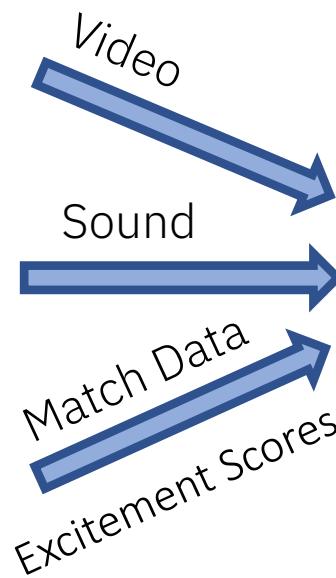
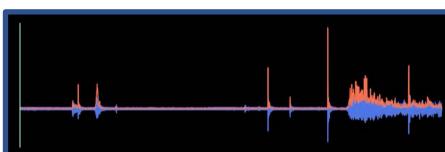


# AI Commentary

Wimbledon and US Open 2023

# Turning multi-sensory data into fluent commentary using Large Language Models

After a wide out serve by Peniston, Laaksonen returns the service with a short backhand. Peniston then approaches the net, hits a forehand, but Laaksonen strikes back with a magnificent backhand. What a beautiful shot! Peniston makes all efforts to save the point but hits the net with a volley unforced error. The game is now tied: 30-30



## Billion-scale IBM Language Model

- IBM Sandstone.3B model
- Pretrained on petabytes of data
- Fine-tuned on real and synthetic commentary data

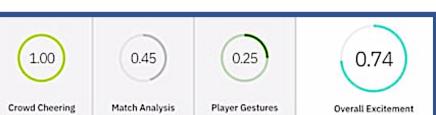
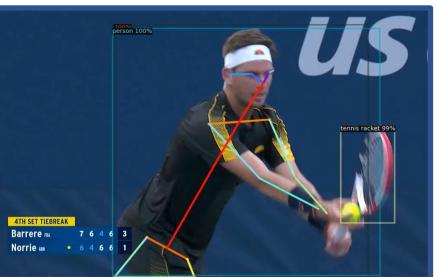


# Turning multi-sensory data into fluent commentary using Large Language Models

After a wide out serve by Peniston, Laaksonen returns the service with a short backhand. Peniston then approaches the net, hits a forehand, but Laaksonen strikes back with a magnificent backhand. What a beautiful shot! Peniston makes all efforts to save the point but hits the net with a volley unforced error. The game is now tied: 30-30

## Billion-scale IBM Language Model

- IBM Sandstone.3B model
- Pretrained on petabytes of data
- Fine-tuned on real and synthetic commentary data



Video  
Sound  
Match Data  
Excitement Scores

Projection into a unified embedding space  
(v1: Data to Text)

# Play-by-play metadata extraction using computer vision

Understanding every detail of the game in real-time



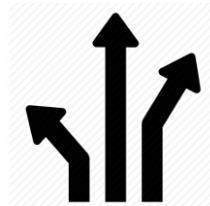
Court and Net  
Detection



Player Tracking  
and pose/keypoint estimation



Ball Tracking

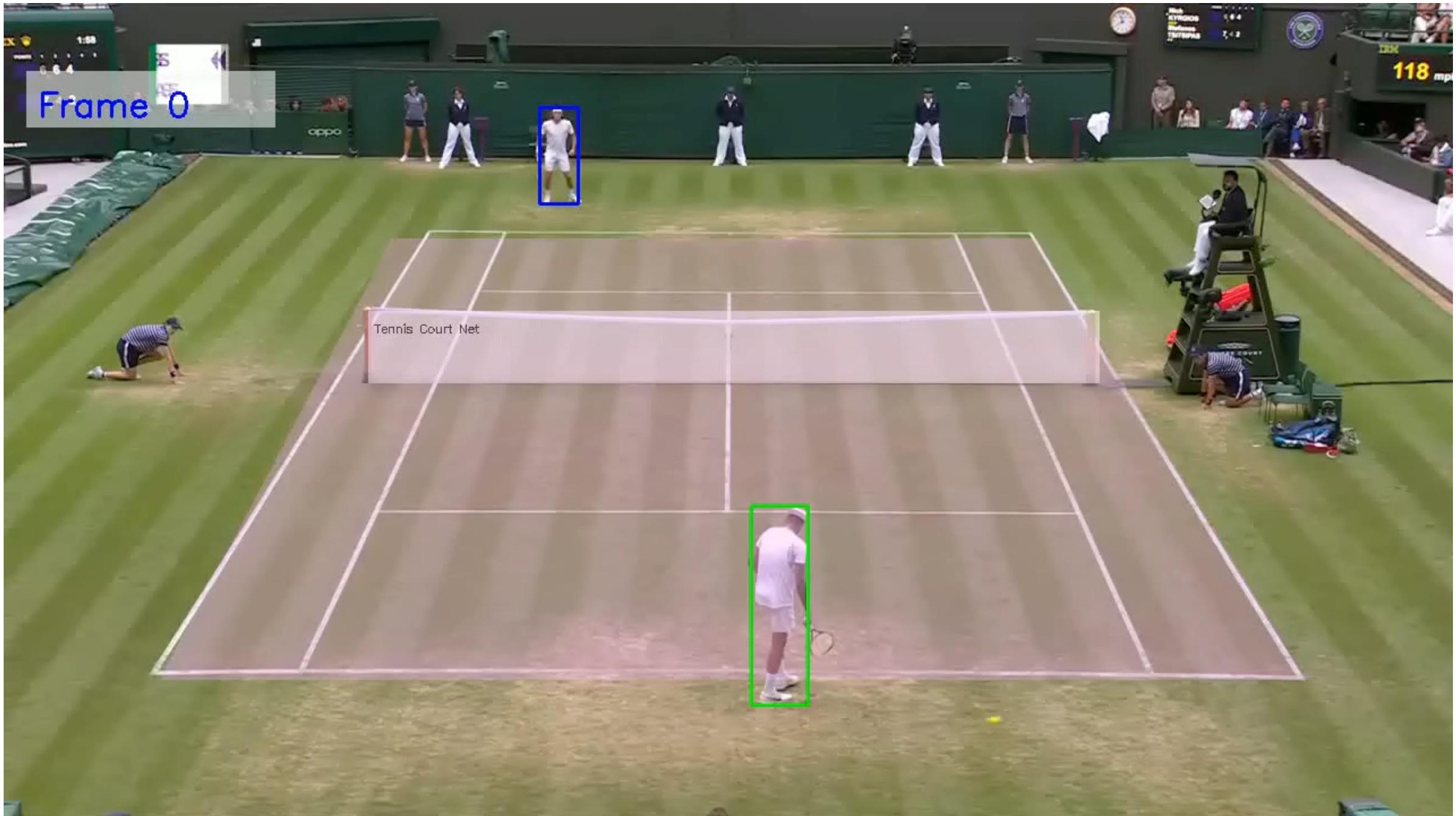


Shot Directionality



Action  
Recognition

Serve, Forehand, Volley, ...



# Metadata from other sensors and modalities



Score information



Ball speed (Radar)



Player information



Rally Length



Audio and vision excitement scores

DJOKOVIC	6-1 6-7 6-3	ALMAGRO
12	ACES	10
0	DOUBLE FAULTS	2
71	1st SERVE IN %	57
78	1st SERVE PTS WON %	70
58	2nd SERVE PTS WON %	53
40	WINNERS	29
21	UNFORCED ERRORS	32
6/7	NET POINTS WON	6/11
3/3	BREAK POINTS WON	0/1

Match analysis



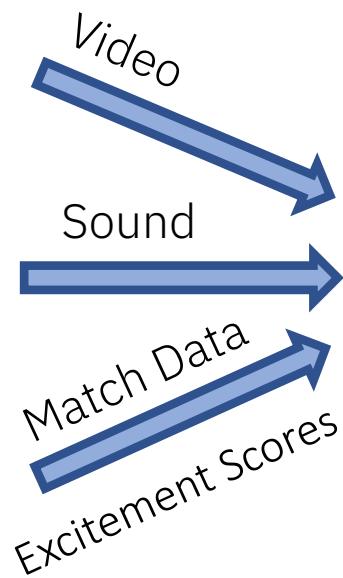
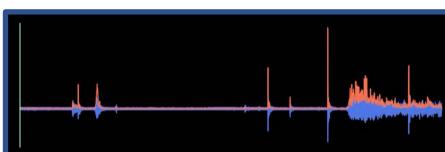
Winner vs Error

# Metadata: Computer vision + other modalities (input to large language model)

[{"id": "1127\_10\_db", "p1": "R. Peniston", "p2": "H. Laaksonen", "p1b": "", "p2b": "", "server": "R. Peniston", "set\_num": 1, "point\_number": 10, "points": "30-15", "game": "1-0", "description": "Peniston loses the point with a forehand volley unforced error.", "excitement\_overall\_score": 0.19817530137293593, "excitement\_crowd\_score": 0.2138548544686001, "excitement\_analytics\_score": 0, "excitement\_gesture\_score": 0.015000396221818044, "winner\_of\_point": 2, "high\_excitement": 0, "long\_rally": 0, "very\_long\_rally": 0, "winner": 0, "break\_point": 0, "break\_point\_won": 0, "forced\_error": 0, "ace": 0, "replay": 0, "volley": 1, "smash": 0, "net\_approach": 1, "serve\_number": 1, "file\_name": "1127\_10\_db.mp4", "event": "wimbledon", "year": "2022", "server\_id": 1, "event\_type": "tennis", "day\_msg": "Day 2: Tue, Jun 29", "court": "Court 12", "p1\_fn": "Ryan Peniston", "p2\_fn": "Henri Laaksonen", "p1b\_fn": "", "p2b\_fn": "", "p1\_nat": "GBR", "p2\_nat": "SUI", "pos": "bottom", "stf": 105, "DoubleFault": 0, "UnforcedError": 1, "NetPoint": 1, "BreakPoint": 0, "BreakPointWon": 0, "ServeDepth": "NCTL", "ReturnDepth": "ND", "rally\_count": 4, "ServeWidth": "BW", "WinnerType": "O", "WinnerShotType": "O", "shots": {"top": "H. Laaksonen", "bottom": "R. Peniston", "seq": [{"who": "R. Peniston", "pos": "bottom", "stf": 58, "eff": 87, "act": "Serve", "where": "out wide"}, {"who": "H. Laaksonen", "pos": "top", "stf": 95, "eff": 126, "act": "Backhand", "where": "up the middle"}, {"who": "R. Peniston", "pos": "bottom", "stf": 142, "eff": 162, "act": "Backhand", "where": "out wide"}, {"who": "R. Peniston", "pos": "bottom", "stf": 199, "eff": 254, "act": "Volley", "where": "up the middle"}, {"who": "R. Peniston", "pos": "bottom", "stf": 275, "eff": 292, "act": "NotAShot", "where": ""}], "pbp\_file\_path": "/dcnstor/leonidka1/sonia\_space/resultsOfMassiveProcessing/videoCVgeneratedMetadata/1127\_10\_db\_metadata.json"}, {"id": "1137\_150\_db", "p1": "A. De Minaur", "p2": "H. Dellien", "p1b": "", "p2b": "", "server": "A. De Minaur", "set\_num": 3, "point\_number": 150, "points": "15-30", "game": "4-4", "description": "De Minaur loses the point with a forehand forced error.", "excitement\_overall\_score": 0.40437613307910225, "excitement\_crowd\_score": 0.6252508636363142, "excitement\_analytics\_score": 0.1, "excitement\_gesture\_score": 0.015000289342755619, "winner\_of\_point": 2, "high\_excitement": 0, "long\_rally": 0, "very\_long\_rally": 0, "winner": 0, "break\_point": 0, "forced\_error": 1, "ace": 0, "replay": 0, "volley": 0, "smash": 0, "net\_approach": 0, "serve\_number": 1, "file\_name": "1137\_150\_db.mp4", "event": "wimbledon", "year": "2022", "server\_id": 2, "event\_type": "tennis", "day\_msg": "Day 2: Tue, Jun 29", "court": "No.1 Court", "p1\_fn": "Alex De Minaur", "p2\_fn": "Hugo Dellien", "p1b\_fn": "", "p2b\_fn": "", "p1\_nat": "AUS", "p2\_nat": "BOL", "pos": "bottom", "stf": 118, "DoubleFault": 0, "UnforcedError": 0, "NetPoint": 0, "BreakPoint": 0, "BreakPointWon": 0, "ServeDepth": "", "ReturnDepth": "", "rally\_count": 1, "ServeWidth": "C", "WinnerType": "O", "WinnerShotType": "O", "shots": {"top": "H. Dellien", "bottom": "A. De Minaur", "seq": [{"who": "H. Dellien", "pos": "top", "stf": 45, "act": "Forehand", "where": "up the middle"}, {"who": "H. Dellien", "pos": "bottom", "stf": 84, "act": "NotAShot", "where": ""}], "pbp\_file\_path": "/dcnstor/leonidka1/sonia\_space/resultsOfMassiveProcessing/videoCVgeneratedMetadata/1137\_150\_db\_metadata.json"}, {"id": "1221\_97\_db", "p1": "D. Shapovalov", "p2": "B. Nakashima", "p1b": "", "p2b": "", "server": "B. Nakashima", "set\_num": 3, "point\_number": 97, "points": "15-0", "game": "0-0", "description": "Shapovalov wins the point with a forehand winner.", "excitement\_overall\_score": 0.6488774868315741, "excitement\_crowd\_score": 0.7745454437580231, "excitement\_analytics\_score": 0.55, "excitement\_gesture\_score": 0.1, "excitement\_analytics\_score": 0.015000289342755619, "winner\_of\_point": 2, "high\_excitement": 0, "long\_rally": 0, "very\_long\_rally": 0, "winner": 0, "break\_point": 0, "forced\_error": 1, "ace": 0, "replay": 0, "volley": 0, "smash": 0, "net\_approach": 0, "serve\_number": 2, "file\_name": "1221\_97\_db.mp4", "event": "wimbledon", "year": "2022", "server\_id": 2, "event\_type": "tennis", "day\_msg": "Day 4: Thur, Jul 1", "court": "Court 12", "p1\_fn": "Denis Shapovalov", "p2\_fn": "Brandon Nakashima", "p1b\_fn": "", "p2b\_fn": "", "p1\_nat": "CAN", "p2\_nat": "USA", "pos": "bottom", "stf": 103, "DoubleFault": 0, "UnforcedError": 0, "NetPoint": 2, "BreakPoint": 0, "BreakPointWon": 0, "ServeDepth": "NCTL", "ReturnDepth": "ND", "rally\_count": 6, "ServeWidth": "C", "WinnerType": "O", "WinnerShotType": "F", "shots": {"top": "D. Shapovalov", "bottom": "B. Nakashima", "seq": [{"who": "B. Nakashima", "pos": "bottom", "stf": 2, "act": "Serve", "where": "up the middle"}, {"who": "D. Shapovalov", "pos": "top", "stf": 48, "eff": 70, "act": "Backhand", "where": "up the middle"}, {"who": "B. Nakashima", "pos": "bottom", "stf": 74, "eff": 95, "act": "Forehand", "where": "out wide"}, {"who": "D. Shapovalov", "pos": "top", "stf": 109, "eff": 120, "act": "UnknownShot", "where": ""}, {"who": "B. Nakashima", "pos": "bottom", "stf": 132, "eff": 154, "act": "Volley", "where": "up the middle"}, {"who": "D. Shapovalov", "pos": "top", "stf": 171, "eff": 193, "act": "UnknownShot", "where": ""}, {"who": "B. Nakashima", "pos": "bottom", "stf": 192, "eff": 203, "act": "Volley", "where": ""}], "pbp\_file\_path": "/dcnstor/leonidka1/sonia\_space/resultsOfMassiveProcessing/videoCVgeneratedMetadata/1221\_97\_db\_metadata.json"}, {"id": "1204\_229\_db", "p1": "T. Van Rijthoven", "p2": "R. Opelka", "p1b": "", "p2b": "", "server": "T. Van Rijthoven", "set\_num": 4, "point\_number": 229, "points": "40-40", "game": "0-1", "description": "Opelka loses the point with a forehand forced error.", "excitement\_overall\_score": 0.2833576388458572, "excitement\_crowd\_score": 0.47660311447231984, "excitement\_analytics\_score": 0.1, "excitement\_gesture\_score": 0.015000647568195426, "winner\_of\_point": 1, "high\_excitement": 1, "long\_rally": 0, "very\_long\_rally": 0, "winner": 0, "break\_point": 0, "forced\_error": 1, "ace": 0, "replay": 0, "volley": 0, "smash": 0, "net\_approach": 0, "serve\_number": 1, "file\_name": "1204\_229\_db.mp4", "event": "wimbledon", "year": "2022", "server\_id": 2, "event\_type": "tennis", "day\_msg": "Day 4: Thur, Jul 1", "court": "Court 12", "p1\_fn": "Denis Shapovalov", "p2\_fn": "Brandon Nakashima", "p1b\_fn": "", "p2b\_fn": "", "p1\_nat": "CAN", "p2\_nat": "USA", "pos": "bottom", "stf": 12, "eff": 41, "act": "Serve", "where": "up the middle"}, {"who": "T. Van Rijthoven", "pos": "bottom", "stf": 42, "eff": 82, "act": "Forehand", "where": "up the middle"}, {"who": "T. Van Rijthoven", "pos": "bottom", "stf": 87, "eff": 102, "act": "Volley", "where": "up the middle"}, {"who": "T. Van Rijthoven", "pos": "bottom", "stf": 130, "eff": 184, "act": "Backhand", "where": "out wide"}, {"who": "T. Van Rijthoven", "pos": "bottom", "stf": 196, "eff": 203, "act": "Backhand", "where": "down the line"}, {"who": "T. Van Rijthoven", "pos": "top", "stf": 202, "eff": 214, "act": "NotAShot", "where": ""}], "pbp\_file\_path": "/dcnstor/leonidka1/sonia\_space/resultsOfMassiveProcessing/videoCVgeneratedMetadata/1204\_229\_db\_metadata.json"}, {"id": "1158\_219\_db", "p1": "J. Sousa", "p2": "R. Gasquet", "p1b": "", "p2b": "", "server": "R. Gasquet", "set\_num": 4, "point\_number": 219, "points": "40-AD", "game": "1-1", "description": "Gasquet loses the point with a double fault.", "excitement\_overall\_score": 0.12421195070626616, "excitement\_crowd\_score": 0.01360409128807183, "excitement\_analytics\_score": 0, "excitement\_gesture\_score": 0.01502846247610608, "winner\_of\_point": 1, "high\_excitement": 0, "long\_rally": 0, "very\_long\_rally": 0, "winner": 0, "break\_point": 0, "forced\_error": 0, "ace": 0, "replay": 0, "volley": 0, "smash": 0, "net\_approach": 0, "serve\_number": 0, "file\_name": "1158\_219\_db.mp4", "event": "wimbledon", "year": "2022", "server\_id": 1, "event\_type": "tennis", "day\_msg": "Day 3: Wed, Jun 30", "court": "Court 4", "p1\_fn": "Tim Van Rijthoven", "p2\_fn": "Reilly Opelka", "p1b\_fn": "", "p2b\_fn": "", "p1\_nat": "NED", "p2\_nat": "USA", "pos": "bottom", "stf": 12, "eff": 41, "act": "Serve", "where": "up the middle"}, {"who": "R. Gasquet", "pos": "bottom", "stf": 42, "eff": 82, "act": "Forehand", "where": "up the middle"}, {"who": "R. Gasquet", "pos": "bottom", "stf": 87, "eff": 102, "act": "Volley", "where": "up the middle"}, {"who": "R. Gasquet", "pos": "bottom", "stf": 130, "eff": 184, "act": "Backhand", "where": "out wide"}, {"who": "R. Gasquet", "pos": "top", "stf": 196, "eff": 203, "act": "Backhand", "where": "down the line"}, {"who": "R. Gasquet", "pos": "top", "stf": 202, "eff": 214, "act": "NotAShot", "where": ""}], "pbp\_file\_path": "/dcnstor/leonidka1/sonia\_space/resultsOfMassiveProcessing/videoCVgeneratedMetadata/1158\_219\_db\_metadata.json"}, {"id": "1151\_134\_db", "p1": "D. Novak", "p2": "F. Bagnis", "p1b": "", "p2b": "", "server": "D. Novak", "set\_num": 2, "point\_number": 134, "points": "40-40", "game": "2-3", "description": "Bagnis wins the point with a backhand volley winner.", "excitement\_overall\_score": 0.4132240626662806, "excitement\_crowd\_score": 0.40765417486424604, "excitement\_analytics\_score": 0.3, "excitement\_gesture\_score": 0, "winner\_of\_point": 2, "high\_excitement": 0, "long\_rally": 0, "very\_long\_rally": 0, "winner": 0, "break\_point": 0, "forced\_error": 0, "ace": 0, "replay": 0, "volley": 0, "smash": 0, "net\_approach": 0, "serve\_number": 0, "file\_name": "1151\_134\_db.mp4", "event": "wimbledon", "year": "2022", "server\_id": 2, "event\_type": "tennis", "day\_msg": "Day 4: Thur, Jul 1", "court": "Court 12", "p1\_fn": "Joao Sousa", "p2\_fn": "Richard Gasquet", "p1b\_fn": "", "p2b\_fn": "", "p1\_nat": "POR", "p2\_nat": "FRA", "pos": "bottom", "stf": 2, "DoubleFault": 2, "UnforcedError": 2, "NetPoint": 0, "BreakPoint": 0, "ServeDepth": "CTL", "ReturnDepth": "", "rally\_count": 0, "ServeWidth": "BW", "WinnerType": "O", "WinnerShotType": "O", "shots": {"top": "R. Gasquet", "pos": "top", "stf": 35, "eff": 62, "act": "Serve", "where": "up the middle"}, {"who": "R. Gasquet", "pos": "top", "stf": 75, "eff": 82, "act": "UnknownShot", "where": ""}, {"who": "R. Gasquet", "pos": "top", "stf": 199, "eff": 271, "act": "Serve", "where": ""}, {"who": "R. Gasquet", "pos": "bottom", "stf": 279, "eff": 291, "act": "Backhand", "where": "up the middle"}, {"who": "R. Gasquet", "pos": "top", "stf": 298, "eff": 320, "act": "NotAShot", "where": ""}, {"who": "R. Gasquet", "pos": "top", "stf": 413, "eff": 434, "act": "NotAShot", "where": ""}], "pbp\_file\_path": "/dcnstor/leonidka1/sonia\_space/resultsOfMassiveProcessing/videoCVgeneratedMetadata/1151\_134\_db\_metadata.json"}, {"id": "1105\_210\_db", "p1": "N. Basilashvili", "p2": "L. Rosol", "p1b": "", "p2b": "", "server": "N. Basilashvili", "set\_num": 4, "point\_number": 210, "points": "0-0", "game": "null-null", "description": "Rosol loses the point with a forehand forced error.", "excitement\_overall\_score": 0.16482244279582572, "excitement\_crowd\_score": 0.25398933260169507, "excitement\_analytics\_score": 0.1, "excitement\_gesture\_score": 0.015000608387808913, "winner\_of\_point": 1, "high\_excitement": 0, "long\_rally": 0, "very\_long\_rally": 0, "winner": 0, "break\_point": 0, "forced\_error": 1, "ace": 0, "replay": 0, "volley": 0, "smash": 0, "net\_approach": 0, "serve\_number": 2, "file\_name": "1105\_210\_db.mp4", "event": "wimbledon", "year": "2022", "server\_id": 1, "event\_type": "tennis", "day\_msg": "Day 1: Mon, Jun 28", "court": "Court 15", "p1\_fn": "Nikoloz Basilashvili", "p2\_fn": "Lukas Rosol", "p1b\_fn": "", "p2b\_fn": "", "p1\_nat": "GEO", "p2\_nat": "CZE", "pos": "bottom", "stf": 153, "eff": 155, "act": "UnknownShot", "where": ""}, {"who": "N. Basilashvili", "pos": "top", "stf": 132, "eff": 135, "act": "Volley", "where": "out wide"}, {"who": "N. Basilashvili", "pos": "bottom", "stf": 134, "eff": 146, "act": "Volley", "where": "out wide"}, {"who": "N. Basilashvili", "pos": "top", "stf": 146, "eff": 154, "act": "NotAShot", "where": ""}, {"who": "D. Novak", "pos": "top", "stf": 153, "eff": 167, "act": "NotAShot", "where": ""}], "pbp\_file\_path": "/dcnstor/leonidka1/sonia\_space/resultsOfMassiveProcessing/videoCVgeneratedMetadata/1105\_210\_db\_metadata.json"}, {"id": "1105\_210\_db", "p1": "N. Basilashvili", "p2": "L. Rosol", "p1b": "", "p2b": "", "server": "N. Basilashvili", "set\_num": 4, "point\_number": 210, "points": "0-0", "game": "null-null", "description": "Rosol loses the point with a forehand forced error.", "excitement\_overall\_score": 0.16482244279582572, "excitement\_crowd\_score": 0.25398933260169507, "excitement\_analytics\_score": 0.1, "excitement\_gesture\_score": 0.015000608387808913, "winner\_of\_point": 1, "high\_excitement": 0, "long\_rally": 0, "very\_long\_rally": 0, "winner": 0, "break\_point": 0, "forced\_error": 1, "ace": 0, "replay": 0, "volley": 0, "smash": 0, "net\_approach": 0, "serve\_number": 2, "file\_name": "1105\_210\_db.mp4", "event": "wimbledon", "year": "2022", "server\_id": 1, "event\_type": "tennis", "day\_msg": "Day 1: Mon, Jun 28", "court": "Court 15", "p1\_fn": "Nikoloz Basilashvili", "p2\_fn": "Lukas Rosol", "p1b\_fn": "", "p2b\_fn": "", "p1\_nat": "GEO", "p2\_nat": "CZE", "pos": "bottom", "stf": 153, "eff": 167, "act": "NotAShot", "where": ""}], {"id": "1115\_171\_db", "p1": "T. Griekspoor", "p2": "F. Fognini", "p1b": "", "p2b": "", "server": "T. Griekspoor", "set\_num": 3, "point\_number": 171, "points": "15-40", "game": "5-2", "description": "Fognini wins the point with an ace.", "excitement\_overall\_score": 0.31960342629821145, "excitement\_crowd\_score": 0.5567825983088074, "excitement\_analytics\_score": 0.1, "excitement\_gesture\_score": 0.015079165274028783, "winner\_of\_point": 2, "high\_excitement": 0, "long\_rally": 0, "very\_long\_rally": 0, "winner": 0, "break\_point": 0, "forced\_error": 0, "ace": 2, "replay": 0, "volley": 0, "smash": 0, "net\_approach": 0, "serve\_number": 0, "file\_name": "1115\_171\_db.mp4", "event": "wimbledon", "year": "2022", "server\_id": 2, "event\_type": "tennis", "day\_msg": "Day 1: Mon, Jun 28", "court": "Court 16", "p1\_fn": "Tallon Griekspoor", "p2\_fn": "Fabio Fognini", "p1b\_fn": "", "p2b\_fn": "", "p1\_nat": "NED", "p2\_nat": "ITA", "pos": "bottom", "stf": 29, "eff": 61, "act": "Backhand", "where": "up the middle"}, {"who": "T. Griekspoor", "pos": "bottom", "stf": 265, "eff": 275, "act": "Forehand", "where": "up the middle"}, {"who": "T. Griekspoor", "pos": "bottom", "stf": 290, "eff": 296, "act": "NotAShot", "where": ""}], "pbp\_file\_path": "/dcnstor/leonidka1/sonia\_space/resultsOfMassiveProcessing/videoCVgeneratedMetadata/1115\_171\_db\_metadata.json"}, {"id": "1115\_171\_db", "p1": "T. Griekspoor", "p2": "F. Fognini", "p1b": "", "p2b": "", "server": "T. Griekspoor", "set\_num": 3, "point\_number": 171, "points": "15-40", "game": "5-2", "description": "Fognini wins the point with an ace.", "excitement\_overall\_score": 0.31960342629821145, "excitement\_crowd\_score": 0.5567825983088074, "excitement\_analytics\_score": 0.1, "excitement\_gesture\_score": 0.015079165274028783, "winner\_of\_point": 2, "high\_excitement": 0, "long\_rally": 0, "very\_long\_rally": 0, "winner": 0, "break\_point": 0, "forced\_error": 0, "ace": 2, "replay": 0, "volley": 0, "smash": 0, "net\_approach": 0, "serve\_number": 0, "file\_name": "1115\_171\_db.mp4", "event": "wimbledon", "year": "2022", "server\_id": 2, "event\_type": "tennis", "day\_msg": "Day 1: Mon, Jun 28", "court": "Court 16", "p1\_fn": "T. Griekspoor", "p2\_fn": "Fabio Fognini", "p1b\_fn": "", "p2b\_fn": "", "p1\_nat": "NED", "p2\_nat": "ITA", "pos": "bottom", "stf": 19, "eff": 45, "act": "Serve", "where": "up the middle"}, {"who": "T. Griekspoor", "pos": "top", "stf": 55, "eff": 92, "act": "Backhand", "where": "up the middle"}, {"who": "T. Griekspoor", "pos": "bottom", "stf": 127, "eff": 239, "act": "Serve", "where": "up the middle"}, {"who": "T. Griekspoor", "pos": "bottom", "stf": 130, "eff": 140, "act": "Volley", "where": "out wide"}, {"who": "T. Griekspoor", "pos": "bottom", "stf": 153, "eff": 167, "act": "NotAShot", "where": ""}], "pbp\_file\_path": "/dcnstor/leonidka1/sonia\_space/resultsOfMassiveProcessing/videoCVgeneratedMetadata/1115\_171\_db\_metadata.json"}, {"id": "1144\_109\_db", "p1": "A. Balazs", "p2": "R. Bautista Agut", "p1b": "", "p2b": "", "server": "A. Balazs", "set\_num": 3, "point\_number": 109, "points": "15-15", "game": "2-1", "description": "Balazs loses the point with a backhand forced error.", "excitement\_overall\_score": 0.181513479450402357, "excitement\_crowd\_score": 0.23575143355335748, "excitement\_analytics\_score": 0.1, "excitement\_gesture\_score": 0.015001179158329013, "winner\_of\_point": 2, "high\_excitement": 0, "long\_rally": 0, "very\_long\_rally": 0, "winner": 0, "break\_point": 0, "forced\_error": 1, "ace": 0, "replay": 0, "volley": 0, "smash": 0, "net\_approach": 0, "serve\_number": 1, "file\_name": "1144\_109\_db.mp4", "event": "wimbledon", "year": "2022", "server\_id": 2, "event\_type": "tennis", "day\_msg": "Day 2: Tue, Jun 29", "court": "Court 9", "p1\_fn": "Attila Balazs", "p2\_fn": "Roberto Bautista Agut", "p1b\_fn": "", "p2b\_fn": "", "p1\_nat": "HUN", "p2\_nat": "ESP", "pos": "bottom", "stf": 29, "eff": 61, "act": "Backhand", "where": "up the middle"}, {"who": "A. Balazs", "pos": "bottom", "stf": 290, "eff": 296, "act": "NotAShot", "where": ""}], "pbp\_file\_path": "/dcnstor/leonidka1/sonia\_space/resultsOfMassiveProcessing/videoCVgeneratedMetadata/1144\_109\_db\_metadata.json"}, {"id": "1144\_109\_db", "p1": "A. Balazs", "p2": "R. Bautista Agut", "p1b": "", "p2b": "", "server": "R. Bautista Agut", "set\_num": 3, "point\_number": 109, "points": "15-15", "game": "2-1", "description": "Balazs loses the point with a backhand forced error.", "excitement\_overall\_score": 0.181513479450402357, "excitement\_crowd\_score": 0.23575143355335748, "excitement\_analytics\_score": 0.1, "excitement\_gesture\_score": 0.015001179158329013, "winner\_of\_point": 2, "high\_excitement": 0, "long\_rally": 0, "very\_long\_rally": 0, "winner": 0, "break\_point": 0, "forced\_error": 1, "ace": 0, "replay": 0, "volley": 0, "smash": 0, "net\_approach": 0, "serve\_number": 1, "file\_name": "1144\_109\_db.mp4", "event": "wimbledon", "year": "2022", "server\_id": 2, "event\_type": "tennis", "day\_msg": "Day 2: Tue, Jun 29", "court": "Court 9", "p1\_fn": "Attila Balazs", "p2\_fn": "Roberto Bautista Agut", "p1b\_fn": "", "p2b\_fn": "", "p1\_nat": "HUN", "p2\_nat": "ESP", "pos": "bottom", "stf": 29, "eff": 61, "act": "Backhand", "where": "up the middle"}, {"who": "A. Balazs", "pos": "bottom", "stf": 290, "eff": 296, "act": "NotAShot", "where": ""}], "pbp\_file\_path": "/dcnstor/leonidka1/sonia\_space/resultsOfMassiveProcessing/videoCVgeneratedMetadata/1144\_109\_db\_metadata.json"}, {"id": "1219\_21\_db", "p1": "A. De Minaur", "p2": "J. Draper", "p1b": "", "p2b": "", "server": "J. Draper", "set\_num": 1, "point\_number": 21, "points": "15-40", "game": "1-2", "description": "Draper wins the point with an ace.", "excitement\_overall\_score": 0.3127906036660247, "excitement\_crowd\_score": 0.2544923735988407, "excitement\_analytics\_score": 0.1, "excitement\_gesture\_score": 0.01513316058898463, "winner\_of\_point": 2, "high\_excitement": 0, "long\_rally": 0, "very\_long\_rally": 0, "winner": 2, "break\_point": 0, "forced\_error": 0, "ace": 2, "replay": 0, "volley": 0, "smash": 0, "net\_approach": 0, "serve\_number": 1, "file\_name": "1219\_21\_db.mp4", "event": "wimbledon", "year": "2022", "server\_id": 2, "event\_type": "tennis", "day\_msg": "Day 4: Thur, Jul 1", "court": "No.1 Court", "p1\_fn": "Alex De Minaur", "p2\_fn": "Jack Draper", "p1b\_fn": "", "p2b\_fn": "", "p1\_nat": "AUS", "p2\_nat": "GBR", "pos": "bottom", "stf": 131, "DoubleFault": 0, "UnforcedError": 0, "NetPoint": 0, "BreakPoint": 0, "BreakPointWon": 0, "ServeDepth": "CTL", "ReturnDepth": "", "rally\_count": 1, "ServeWidth": "C", "WinnerType": "O", "WinnerShotType": "O", "shots": {"top": "A. De Minaur", "bottom": "J. Draper", "seq": [{"who": "A. De Minaur", "pos": "bottom", "stf": 25, "act": "Serve", "where": "up the middle"}, {"who": "A. De Minaur", "pos": "top", "stf": 37, "eff": 63, "act": "Replay", "where": "up the middle"}, {"who": "A. De Minaur", "pos": "bottom", "stf": 40, "eff": 47, "act": "NotAShot", "where": ""}], "pbp\_file\_path": "/dcnstor/leonidka1/sonia\_space/resultsOfMassiveProcessing/videoCVgeneratedMetadata/1219\_21\_db\_metadata.json"}, {"id": "1219\_21\_db", "p1": "A. De Minaur", "p2": "J. Draper", "p1b": "", "p2b": "", "server": "J. Draper", "set\_num": 1, "point\_number": 21, "points": "15-40", "game": "1-2", "description": "Draper wins the point with an ace.", "excitement\_overall\_score": 0.3127906036660247, "excitement\_crowd\_score": 0.2544923735988407, "excitement\_analytics\_score": 0.1, "excitement\_gesture\_score": 0.01513316058898463, "winner\_of\_point": 2, "high\_excitement": 0, "long\_rally": 0, "very\_long\_rally": 0, "winner": 2, "break\_point": 0, "forced\_error": 0, "ace": 2, "replay": 0, "volley": 0, "smash": 0, "net\_approach": 0, "serve\_number": 1, "file\_name": "1219\_21\_db.mp4", "event": "wimbledon", "year": "2022", "server\_id": 2, "event\_type": "tennis", "day\_msg": "Day 4: Thur, Jul 1", "court": "No.1 Court", "p1\_fn": "Alex De Minaur", "p2\_fn": "Jack Draper", "p1b\_fn": "", "p2b\_fn": "", "p1\_nat": "AUS", "p2\_nat": "GBR", "pos": "bottom", "stf": 131, "DoubleFault": 0, "UnforcedError": 0, "NetPoint": 0, "BreakPoint": 0, "BreakPointWon": 0, "ServeDepth": "CTL", "ReturnDepth": "", "rally\_count": 1, "ServeWidth": "C", "WinnerType": "O", "WinnerShotType": "O", "shots": {"top": "A. De Minaur", "bottom": "J. Draper", "seq": [{"who": "A. De Minaur", "pos": "bottom", "stf": 25, "act": "Serve", "where": "up the middle"}, {"who": "A. De Minaur", "pos": "top", "stf": 37, "eff": 63, "act": "Replay", "where": "up the middle"}, {"who": "A. De Minaur", "pos": "bottom", "stf": 40, "eff": 47, "act": "NotAShot", "where": ""}], "pbp\_file\_path": "/dcnstor/leonidka1/sonia\_space/resultsOfMassiveProcessing/videoCVgeneratedMetadata/1219\_21\_db\_metadata.json"}, {"id": "1107\_106\_db", "p1": "T. Van Rijthoven", "p2": "F. Delbonis", "p1b": "", "p2b": "", "server": "F. Delbonis", "set\_num": 2, "point\_number": 106, "points": "15-40", "game": "1-2", "description": "Delbonis wins the point with an ace.", "excitement\_overall\_score": 0.181513479450402357, "excitement\_crowd\_score": 0.23575143355335748, "excitement\_analytics\_score": 0.1, "excitement\_gesture\_score": 0.015001179158329013, "winner\_of\_point": 2, "high\_excitement": 0, "long\_rally": 0, "very\_long\_rally": 0, "winner": 2, "break\_point": 0, "forced\_error": 0, "ace": 2, "replay": 0, "volley": 0, "smash": 0, "net\_approach": 0, "serve\_number": 1, "file\_name": "1107\_106\_db.mp4", "event": "wimbledon", "year": "2022", "server\_id": 2, "event\_type": "tennis", "day\_msg": "Day 4: Thur, Jul 1", "court": "No.1 Court", "p1\_fn": "T. Van Rijthoven", "p2\_fn": "F. Delbonis", "p1b\_fn": "", "p2b\_fn": "", "p1\_nat": "NED", "p2\_nat": "ESP", "pos": "bottom", "stf": 131, "DoubleFault": 0, "UnforcedError": 0, "NetPoint": 0, "BreakPoint": 0, "BreakPointWon": 0, "ServeDepth": "CTL", "ReturnDepth": "", "rally\_count": 1, "ServeWidth": "C", "WinnerType": "O", "WinnerShotType": "O", "shots": {"top": "T. Van Rijthoven", "bottom": "F. Delbonis", "seq": [{"who": "T. Van Rijthoven", "pos": "bottom", "stf": 25, "act": "Serve", "where": "up the middle"}, {"who": "T. Van Rijthoven", "pos": "top", "stf": 37, "eff": 63, "act": "Replay", "where": "up the middle"}, {"who": "T. Van Rijthoven", "pos": "bottom", "stf": 40, "eff": 47, "act": "NotAShot", "where": ""}], "pbp\_file\_path": "/dcnstor/leonidka1/sonia\_space/resultsOfMassiveProcessing/videoCVgeneratedMetadata/1107\_106\_db\_metadata.json"}]

# Turning multi-sensory data into fluent commentary using Large Language Models

After a wide out serve by Peniston, Laaksonen returns the service with a short backhand. Peniston then approaches the net, hits a forehand, but Laaksonen strikes back with a magnificent backhand. What a beautiful shot! Peniston makes all efforts to save the point but hits the net with a volley unforced error. The game is now tied: 30-30



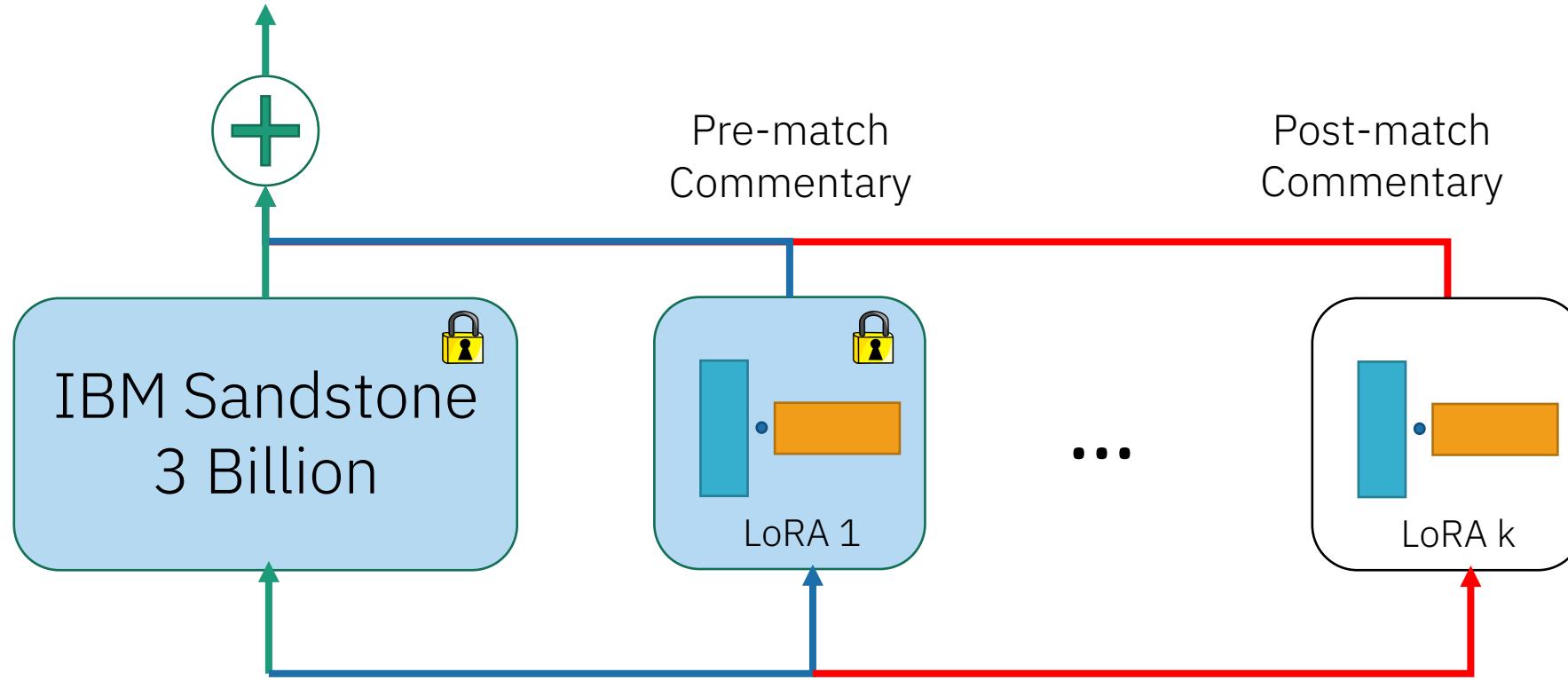
Projection into a unified embedding space  
(v1: Data to Text)

## Billion-scale IBM Language Model

- IBM Sandstone.3B model
- Pretrained on petabytes of data
- Fine-tuned on real and synthetic commentary data

# Our Layered LoRA Architecture (LaLo)

[James Smith et al, CVPR 2023]



LaLo can support task modularity & efficient recall of past models

Efficient adaptation: < 3% of all parameters are adapted !

# Wimbledon Multimodal Perception and Commentary

IBM

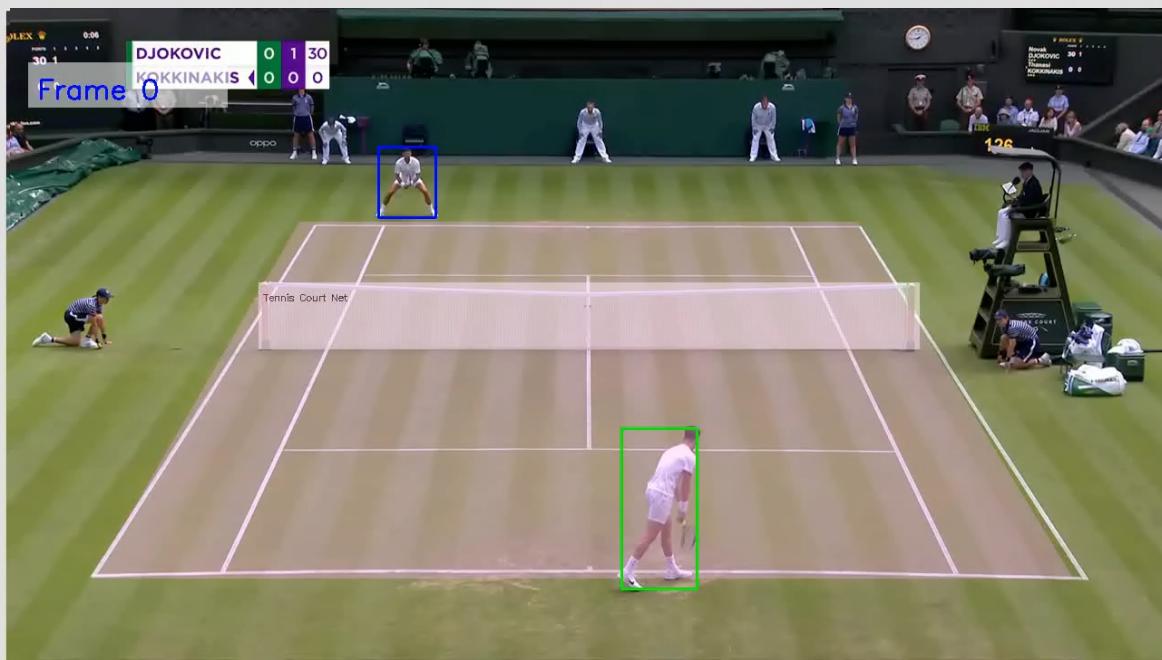
All Events All Rounds All Statistics Day 3 Wed 3 July

Type Player Name

COVERAGE  
13169 48  
Points Hours

HIGHLIGHTS  
14 0  
Produced Published

TOP 5 MAIN EVENTS HIGHLIGHTS  
0 0 900  
Published Secs Published Secs Remaining



## Djokovic vs Kokkinakis

Set 1: 1-0; 15-30

### Auto-Generated Commentary:

After a Kokkinakis serve up the middle, Djokovic hits a backhand. After a Kokkinakis forehand up the middle, the athletes jump into a sequence of backhand to backhand. Unlucky miss by Djokovic as he directs a backhand unforced error into the net to lose the point.

Most Excitement | Most Recent



0.90

Wednesday, 3 July 2019, 13:08

Novak Djokovic vs. Thanasi Kokkinakis

Set 1: 15-30 : Djokovic loses the point with a backhand unforced error.



0.15

Wednesday, 3 July 2019, 13:08

Serena Williams vs. Alize Cornet

Set 3: 15-40 : Break Point; Williams wins the point with a forehand winner.



0.27

Wednesday, 3 July 2019, 13:07

Novak Djokovic vs. Thanasi Kokkinakis

Set 1: 0-30 : Kokkinakis loses the point with a forehand unforced error.



0.86

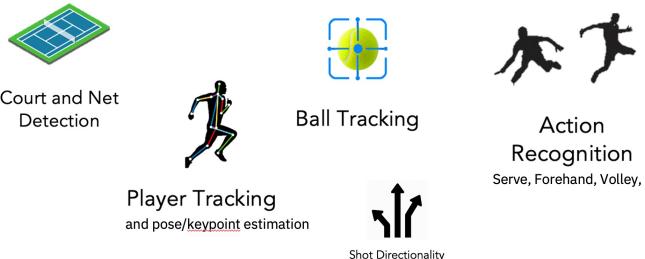
Wednesday, 3 July 2019, 13:06

Jeremy Chardy vs David Goffin

Set 1: 15-40 : Break Point; Goffin wins the point with a forehand winner.

# AI Storytelling

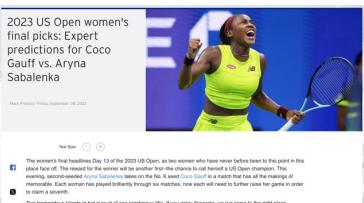
## Computer Vision Metadata



## Metadata from other Sensors and Modalities



## Previous News Articles and Interview Transcripts



### US OPEN

September 9, 2023

Coco Gauff

New York, New York, USA

Press Conference

C. GAUFF/A. Sabalenka

2-6, 6-3, 6-2

MATCH CHAMP

COCO GAUFF: Thank you.

THE MODERATOR: You were hot coming into this tournament. You're obviouslyizzling right now. Talk about how you feel physically and mentally coming into this tournament.

COCO GAUFF: I mean, I didn't get more dramatic than that to be honest. Yeah, I went to the bathroom and washed my face, washed my hands, put some water on my face, and I was like, Okay, I've got to just relax and relax it.

Today I went into this match like it was any other match. Honestly, wasn't nervous going in. She was just playing great, and I was just trying to stay focused and try to play my best tennis and be happy with the results.

THE MODERATOR: Questions?

Q: Coco, you said on court after the French Open that you felt like the world was going to end because you lost to Aryna Sabalenka. Do you feel like you're still dealing with that?

COCO GAUFF: Honestly, I feel like probably this French Open is honestly where it changed. Because I felt

guilty to have lost to her, and I think that was a rough, rough loss, because I thought I was never

going to lose.

This tremendous victory is in pursuit of one prestigious title. If you enjoy tennis, you've come to the right place.

## Coco Gauff's Dream Comes True: Defeats Aryna Sabalenka in Thrilling US Open Final

In one of the most thrilling matches of the US Open, Coco Gauff defeated Aryna Sabalenka in three sets: 2-6, 6-3, 6-2. The final match was played at Arthur Ashe Stadium on Saturday, September 09, 2023. Gauff, who is just 19 years old, proved that she is a force to be reckoned with in the world of women's tennis.

The first set was a tightly contested affair, with Sabalenka emerging victorious 2-6. However, Gauff bounced back in the second set, winning 6-3. The third set was a nail-biting affair, with Gauff ultimately emerging victorious 6-2.

The match was filled with dramatic moments, including a set point for Gauff in the first set that was thwarted by an unforced forehand error. Sabalenka, who is 25 years old and from Belarus, capitalized on the mistake to take the set. However, Gauff refused to give up, and her determination paid off in the end.



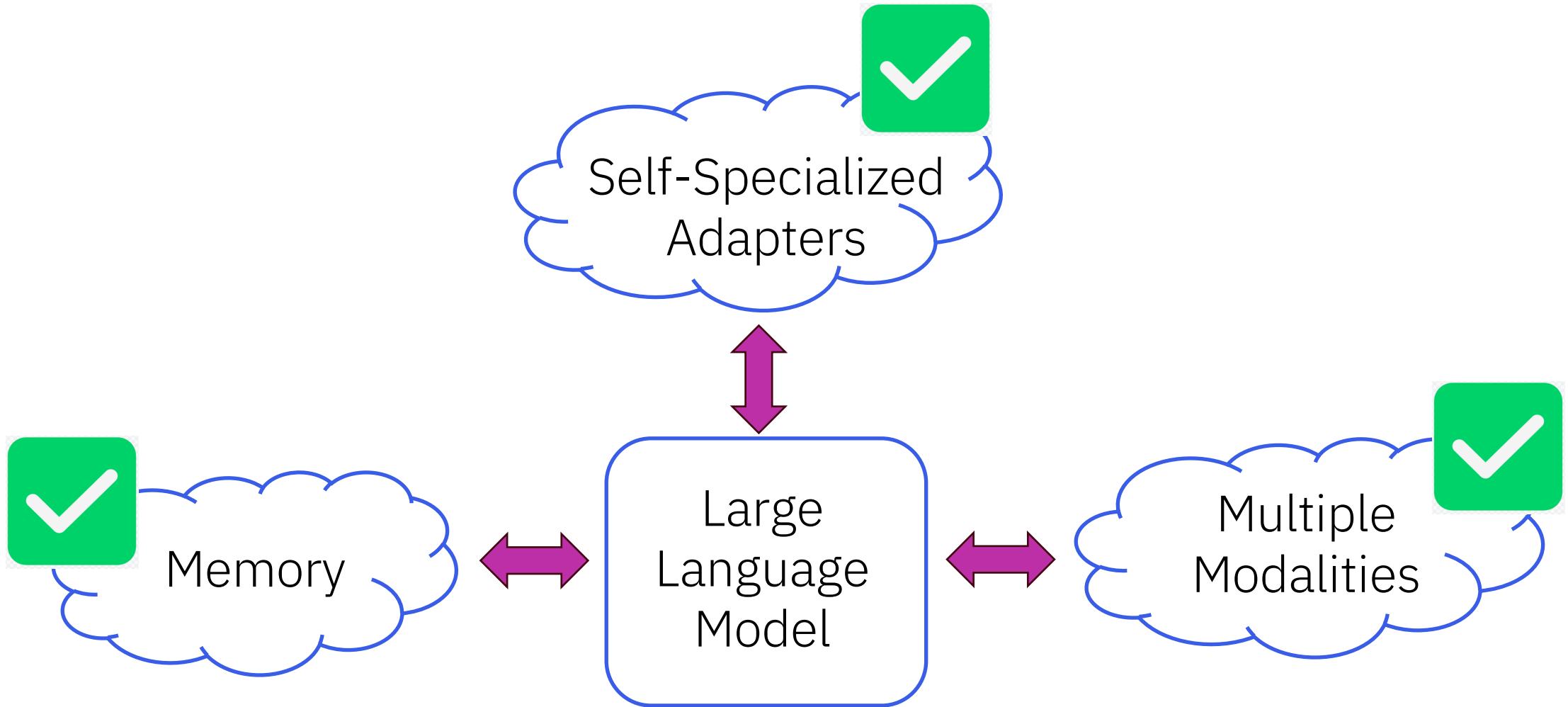
In an interview after the match, Gauff struggled to put into words the significance of her victory. "I don't know, I think the three words I would put it in is 'Dreams come true,' and that this is crazy. I still have no words," she said.

Gauff also reflected on the lessons she learned from her previous defeat at the French Open. "Honestly, I feel like probably this French Open is honestly where it changed. Because I felt pressure to back up the final, and I obviously didn't," she said.

Despite the challenges she faced, Gauff's determination and resilience ultimately paid off, as she emerged victorious in the final set. The crowd erupted in cheers as Gauff was crowned the champion, and she was overcome with emotion.

The victory marks a significant milestone in Gauff's career, as she moves up to rank 3 in the world. Sabalenka, who was ranked number 1 before the match, will drop to number 2.

The US Open final match between Gauff and Sabalenka will be remembered as one of the most thrilling and dramatic matches of the tournament. Gauff's determination and resilience proved that she is a force to be reckoned with in the world of women's tennis.



- Ongoing work: Self-specialization and long content modeling for multimodal input sequences

# Thank you!

See more at <http://rogerioferis.org>