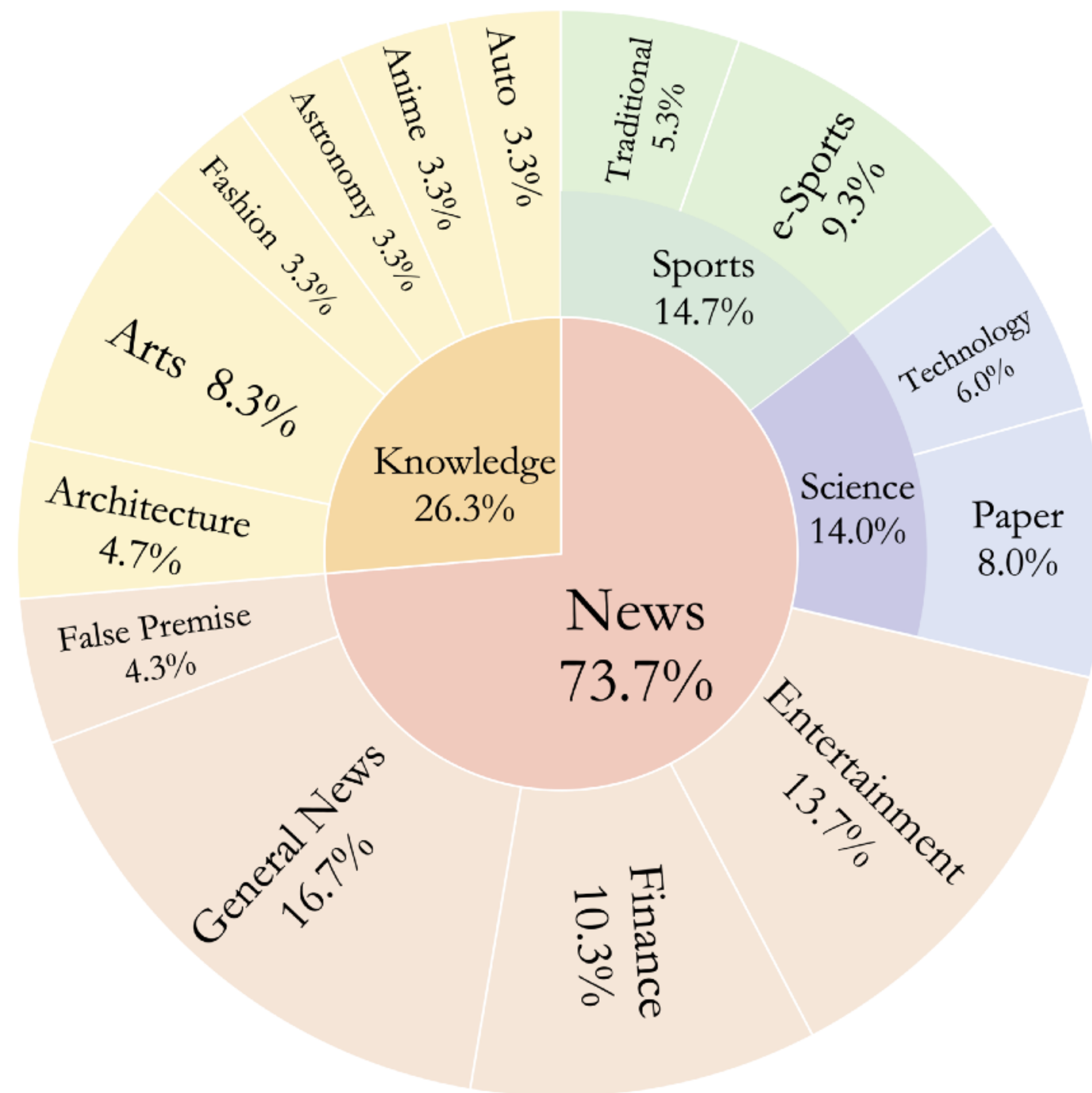


MMSearch

Benchmarking the Potential of Large Models as Multi-modal Search Engines


- MMSearch, a comprehensive evaluation benchmark to assess the multimodal search performance of LMMs
- MMSearch-Engine, to empower any LMMs with multimodal search capabilities




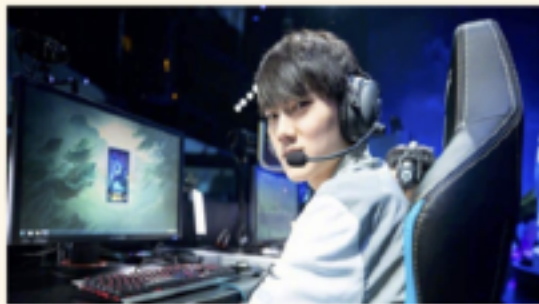


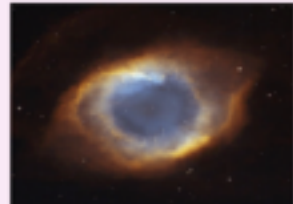
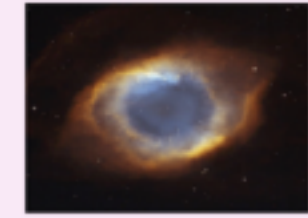




MMSearch

Background, Motivation

- Most existing benchmark for LLM search doesn't include image input.
- How to evaluate and analyze MLLM's performance when using them as search engine?



MMSearch

<p>News - Finance</p>  <p>Question: Is the closing price (at 4:00 PM) of the company shown in the image on August 2nd higher than the closing price on the day shown in the image?</p> <p>→ Answer: No</p>	<p>News - Sports</p>  <p>Question: How many times did this player play Kalista in LPL 2024 spring season?</p> <p>→ Answer: 9</p>	<p>News - Science</p>  <p>Question: What format of data does this work convert 3D point clouds to for segmentation?</p> <p>→ Answer: Videos</p>	<p>News - Entertainment</p>  <p>Question: How the number of concurrent players on the first day of this game on Steam rank in history?</p> <p>→ Answer: 2</p>
<p>News - General</p> <p>Question: How many cabins on the Ferris wheel were caught fire at a music festival in Germany in August 2024?</p> <p>→ Answer: 2</p>			<p>News - False Premise</p> <p>Question: When did Roger Federer win his first Olympic gold medal in singles tennis?</p> <p>→ Answer: Invalid question</p>
<p>Knowledge - Astronomy</p>  <p>Question: What is the radius of the nebula shown in the image?</p> <p>→ Answer: 2.5 light-years</p>			<p>Knowledge - Auto</p>  <p>Question: What is the drag coefficient of this car?</p> <p>→ Answer: 0.198Cd</p>
<p>Knowledge - Fashion</p>  <p>Question: When did this clothes release?</p> <p>→ Answer: 2024-08-31</p>	<p>Knowledge - Art</p>  <p>Question: Which library houses the book collection that includes this painting?</p> <p>→ Answer: Heidelberg University Library</p>	<p>Knowledge - Architecture</p>  <p>Question: What kind of people can go to this building without buying ticket?</p> <p>→ Answer: Children (18 and Under)</p>	<p>Knowledge - Anime</p>  <p>Question: How old is the user of this stand?</p> <p>→ Answer: 19</p>