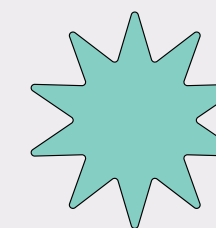
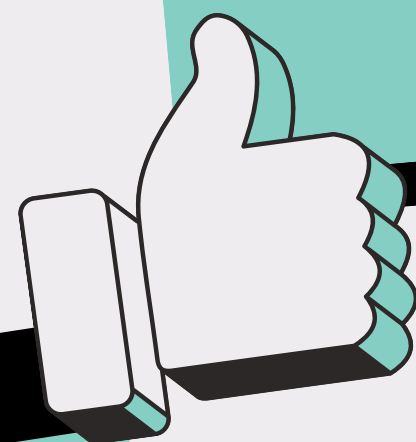
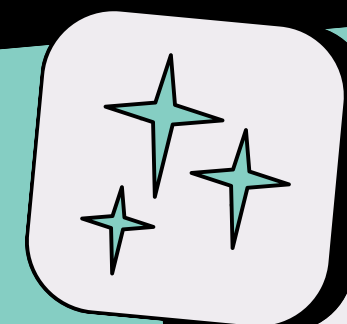


APLICAÇÃO DAS FERRAMENTAS

MULTIMODAIS



CONTEÚDO



- **Tópicos:**

- Entenda o que são ferramentas multimodais;
- Descubra como funcionam os MLLMs;
- Veja aplicações práticas na vida real;
- Conheça os principais modelos atuais.

CONCEITOS



- **Ferramentas Multimodais**
 - Integram texto, imagem, vídeo e/ou áudio
 - Imitação da percepção humana
 - Tornam a IA mais útil, contextual e interativa

CONCEITOS

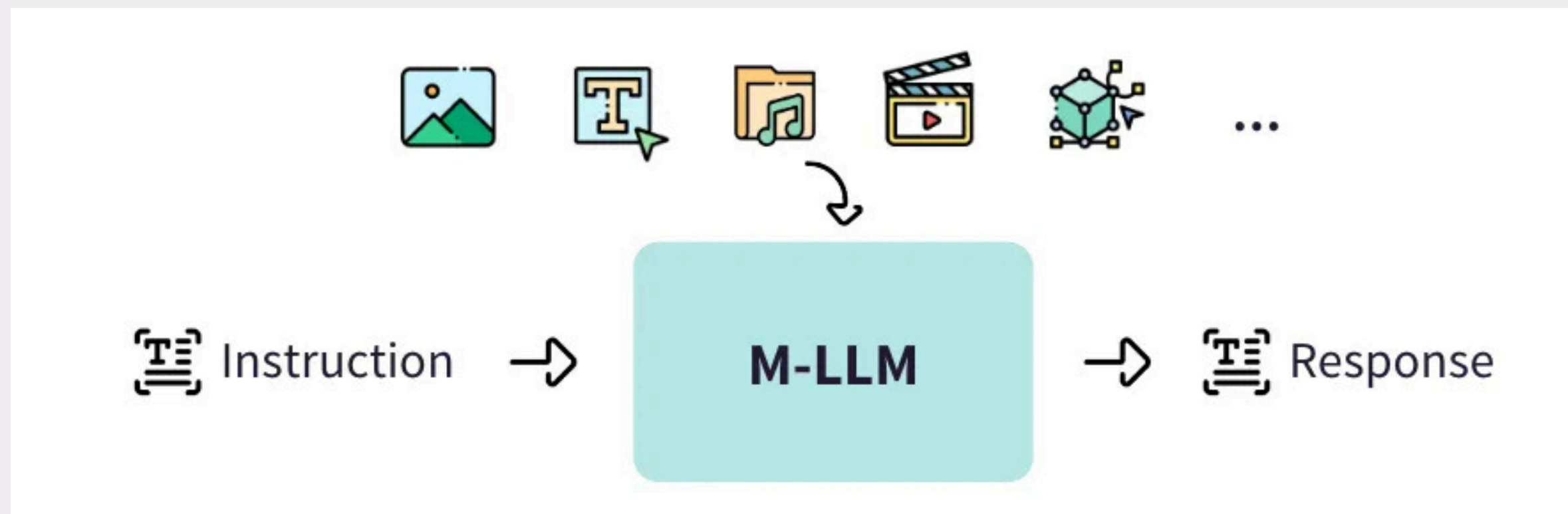
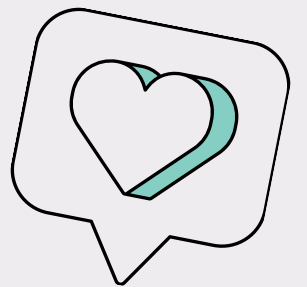


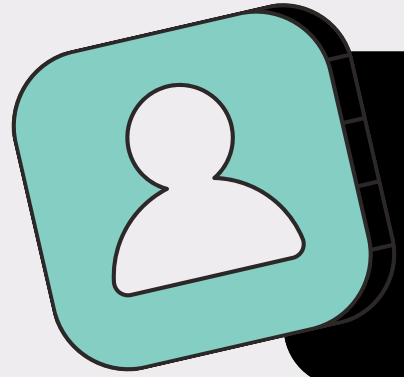
- **MLLMs: Modelos de Linguagem Multimodal**
 - Modelos que entendem múltiplas formas de dados
 - Usam texto, imagem, som ou vídeo como entrada
 - Geram respostas combinando diferentes formatos



COMO OS MLLMS FUNCIONAM?

- **Entrada:** múltiplos formatos (ex.: texto + imagem)
- **Processamento:** redes neurais profundas
- **Saída:** resposta coerente e integrada

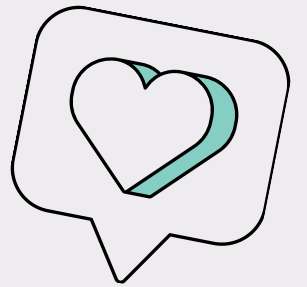




ESTRATÉGIAS DE FUSÃO MULTIMODAL

- **Fusão Precoce (Early Fusion):**

- Combina os dados logo na entrada (ex.: imagem + texto processados juntos desde o início)
- Exige formatos compatíveis



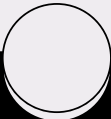
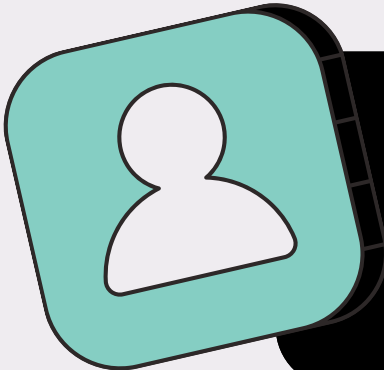
- **Fusão Intermediária (Intermediate Fusion):**

- Cada tipo de dado é processado separadamente antes de ser integrado
- Balanceia especialização e contexto

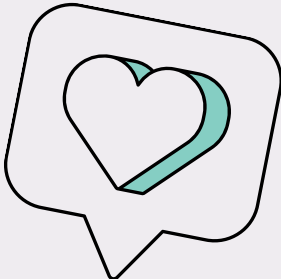
- **Fusão Tardia (Late Fusion):**

- Cada modalidade é processada de forma independente até a decisão final
- Mais simples, mas pode perder nuances contextuais



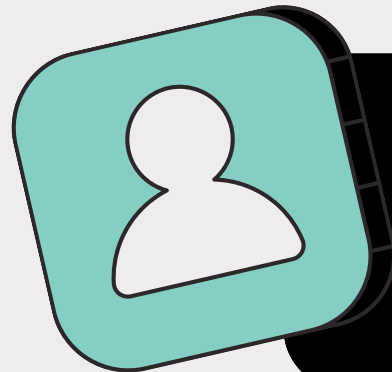


LLM X MLLM

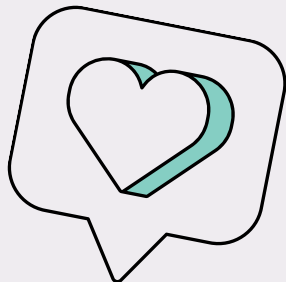
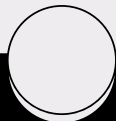


Característica	LLM	MLLM
Modalidade de entrada	Texto	Texto + Imagem (ou outras mídias)
Exemplos	GPT-3, GPT-4 (texto), Claude	GPT-4o, Gemini, Claude 3.5, LLaVA, Qwen VL
Capacidade	Linguagem natural	Linguagem + conteúdo visual (ou sonoro)
Uso típico	Chatbots, resumos, respostas	Leitura de imagens, assistência visual



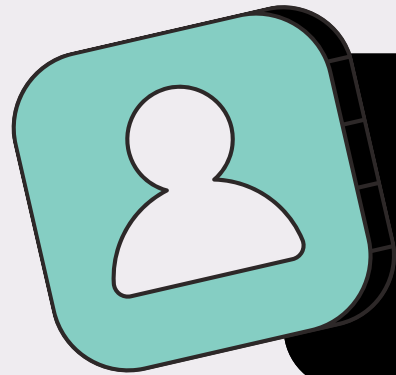


EXEMPLOS DE APLICAÇÕES

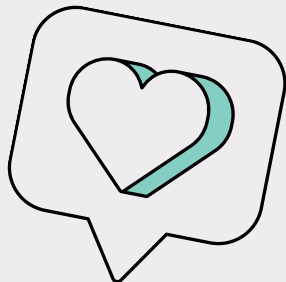
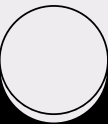


Área	Aplicação Prática
Assistência Médica	IA que interpreta exames e dá sugestões
Educação	Explicação visual de problemas matemáticos
Entretenimento	Criação de personagens e cenas com IA
Atendimento	Análise de imagem enviada por cliente
Segurança	Detecção automática de objetos em vídeo
Comércio Eletrônico	IA que gera imagens a partir de descrições





PRINCIPAIS FERRAMENTAS MLLMS



Ferramenta	Destaque Principal
GPT-4o	Texto, imagem e áudio combinados em tempo real
Gemini	Multimodal do Google com integração total
InternVL	Entendimento visual com contexto textual
LLaVA	Open-source para Q&A com imagem
Claude 3.5	Forte em leitura multimodal e contexto detalhado
Qwen VL	Alto desempenho em tarefas complexas com imagem





Quais os nomes citados no vídeo?

Os nomes que aparecem no vídeo são:

- Padre João
- Chicó
- João Grilo
- Antônio Moraes
- Dona Dorinha
- Rosinha
- Eurico



crie um desenho estilizado do barco da foto



Com certeza!





```
[
  {
    "nome do livro": "As 100 Regras do Dinheiro: Os Segredos Mais Importantes...",
    "autor": "SAMUEL RIVER",
    "número de estrelas": "4,4",
    "preço": "R$ 0,61"
  },
  {
    "nome do livro": "VIDAS IMORTAIS: EPITÁFIOS",
    "autor": "Bruno Ezagui",
    "número de estrelas": "3,0",
    "preço": "R$ 0,01"
  },
  {
    "nome do livro": "Ninguém Pode Ferir Você Sem a Sua Permissão:...",
    "autor": "Ranjot Singh Chahal",
    "número de estrelas": "4,7",
    "preço": "R$ 0,69"
  },
  {
    "nome do livro": "FAMÍLIA FRANKSTEIN: CRIANDO UMA EDUCAÇÃO...",
    "autor": "Bruno Ezagui",
    "número de estrelas": "4,5",
    "preço": "R$ 0,01"
  },
  {
    "nome do livro": "As Armas da Persuasão: Estratégias Poderosas para...",
    "autor": "SERGIO T.A",
    "número de estrelas": "4,5",
    "preço": "R$ 2,99"
  }
]
```

objetos JSON, na qual cada um

CONCLUSÃO



- MLLMs ampliam as capacidades da IA
- Transformam setores como saúde, educação e criação
- Dominar essas ferramentas é diferencial estratégico

