```
Start coding or generate with AI.
```

Day 2: LLM/NLP Subteam Documentation

1 Shortlist of Pre-trained LLMs

LLM Name	Size	Pretrained?	Link	Notes
Local 7–8B Instruct	7–8B	Yes	[Link]	Can run locally; scores transcript + visual cues for consistency
LLaMA 7B	7B	Yes	Hugging Face	Instruction-tuned; can process text prompts locally
GPT-4 / GPT-3.5	175B	Yes	OpenAI	Cloud API; can be used for semantic/prosody alignment checks

---

2 Basic LLM Integration Pipeline

Goal: Use LLM to check **text + visual cue consistency** as a post-processing step for AV deepfake detection.

Pipeline Steps:

1. **Input preparation**
   - Extract transcripts using ASR (e.g., Whisper)
   - Optionally include **visual cue descriptions** from VSR model (e.g., AV-HuBERT embeddings or lip motion features)
2. **LLM processing**
   - Prompt the LLM to evaluate semantic/prosody plausibility
   - Example prompt:

     ```
     Given the transcript and visual cues, rate the plausibility of lip-sync and natural speech rhythm on a scale of 0–1.
     ```
3. **Output handling**
   - LLM returns a **consistency score**
   - Score can be passed to a **light classifier** along with AV sync features
4. **Integration with final decision**
   - Fuse LLM score with **SyncNet / AV-HuBERT features**
   - Compute final **real/fake label** and (optionally) temporal localization

Notes:

- This step is purely **text/documentation** now — no coding required until datasets are available.
- Keep prompts and scoring rubric flexible for later experiments.