Enhancing Video Content Metadata Using AI-Generated Summaries

Context

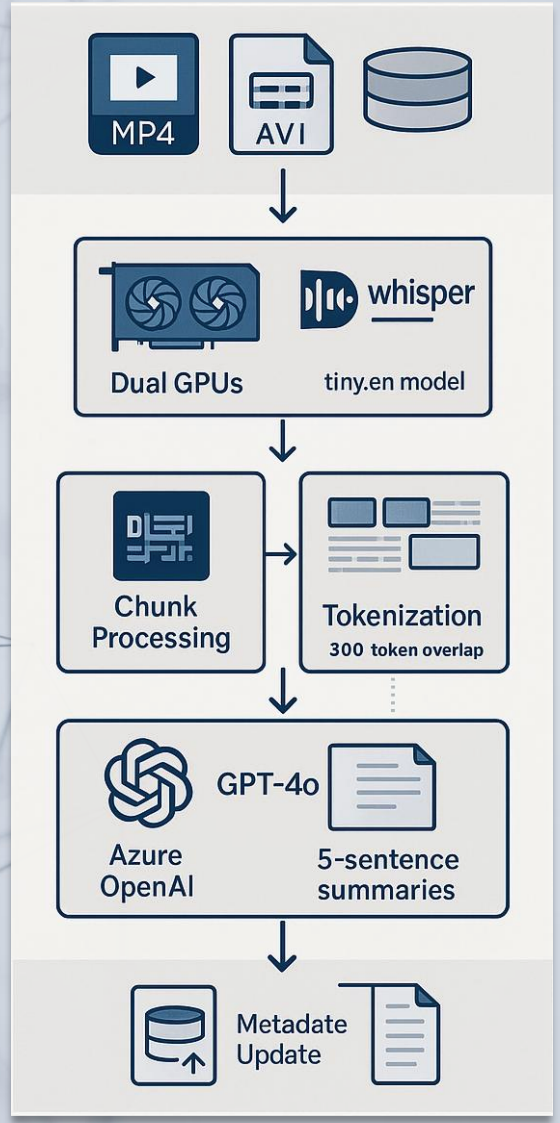
Accurate metadata is essential for organizing and finding digital content effectively. In our organization, Preservica has ingested terabytes of video files with minimal descriptive metadata, making it challenging to search for and use these assets. However, the benefits of improved metadata extend beyond access and discovery, it enables us to understand the context, provenance, and relationships of each asset, which is crucial for effective preservation. By introducing an automated solution that generates metadata for these videos, we strengthen our ability to preserve, manage, and trust our digital assets for the long term.

Solution

To address this issue, a Python script was developed utilizing the following technical resources:

- **OpenAI Whisper Model** for transcription of audio from video files.
- **Tiktoken** to facilitate tokenization of transcribed text.
- **ChatGPT-4o** for summarization of transcriptions.
- **pyPreservica** library employed to update metadata within Preservica via API.

The solution adopts a structured, multi-step workflow for summarizing video files and updating their metadata in Preservica. Given the sensitivity of the data, deployment occurs entirely on-premises using OpenAI Whisper models and a secure, private instance of ChatGPT-4o.



Performance Features

- **Dual GPU Parallel Processing** - Utilizes 2x NVIDIA L40-16Q GPUs simultaneously for 2x throughput.
- **Dynamic Memory Allocation** - Adaptively uses 35% GPU memory (5.6GB) with automatic adjustment.
- **Batch Processing** - Processes multiple videos concurrently with configurable batch size.
- **Optimized Model Selection** - Uses Whisper tiny.en (150MB) for maximum efficiency and reliability.
- **Real-time Progress Monitoring** - Live GPU statistics, processing metrics, and completion notifications.

Error Handling & Recovery

- **Automatic GPU-to-CPU Fallback** - Seamlessly switches to CPU mode after 3 consecutive CUDA errors.
- **Smart Rate Limiting** - Exponential backoff with automatic retry for API 429 errors.
- **Memory Overflow Protection** - Clears GPU cache and reduces allocation on Out of Memory errors.
- **Content Filter Resilience** - Sanitizes problematic text and retries failed API calls.
- **Checkpoint Recovery** - Tracks processed assets to resume after interruptions without duplication.

Elvis Valdes Ramirez (WBG Archives) elvisvr@worldbankgroup.org www.worldbank.org/en/archive/aboutus/Enhancing-Video-Content-Metadata-Using-AI-Generated-Summaries