

Videotto Engineering Internship Exercise

Context

Videotto is a video automation platform that turns long-form videos into short, shareable clips.

You have been shortlisted for an internship exercise to determine if you have what it takes to join Videotto and build the future with us.

This exercise mirrors a simplified version of the real problems we work on.

We care more about **how you think and make decisions** than perfect code or complex algorithms.

Your Task

Build a small system that:

1. Takes a **local video file** as input (provided via Dropbox link)
2. Analyzes the video and identifies the **top 3 clips**
3. For each clip, outputs:
 - **Start time**
 - **End time**
 - A short explanation of **why this clip was selected**
4. Exposes a way to:
 - Track processing **status/progress**
 - View the results on a web based UI

Core Requirements (Must-Have)

1) Video Input

- Use the **provided local video file** as the primary input.

2) Clip Selection

- Select **3 best clips** from the video.
- There is **no fixed definition of “best”** and this is intentional.

You may use any approach you think is reasonable, for example:

- Transcript-based scoring
- Audio energy / speech detection
- Scene changes
- Heuristics or simple rules
- AI-assisted methods

Important: What matters most is **how you explain and justify your ranking**.

3) Results Output

Return results via UI output that includes:

- Start/end timestamps for each clip
- A short explanation per clip

4) Status / Progress

- Implement a simple way to **print step statuses** during processing.
- Suggested statuses: processing, completed, failed

5) Deliverables

- Web app link to access and test (AWS backend, Frontend of your choice)
- Source code (GitHub rep)
- A clear **README** explaining:
 - How your clip-ranking works
 - Tradeoffs and decisions you made
 - What you would improve with more time

Bonus points (Optional)

- Support **YouTube** input
 - Support **Dropbox link** input (download + process)
 - Support **Google Drive picker**
 - Export and encode the actual clips in mp4 format
 - Convert landscape videos to vertical (center crop or face tracking)
 - Basic editing features (trim/crop/merge, etc.)
-

Technical Requirements

- **AWS, Python, React** are required
-

Time Expectations

- Expected time: **5 hours**
 - Please do not over-engineer the solution
-

Use of AI Tools

- You may use AI tools.
 - You should understand and be able to explain your solution.
-

How We Evaluate

We evaluate based on:

- Decision-making and tradeoffs
- Code clarity and structure
- Ability to explain and justify choices
- Practicality and completeness

Dropbox Link [Here](#)