



VESTA

Improving education through automation

Canadian education has advanced significantly beyond lectures delivered before a chalkboard. Technology has changed the educational landscape with videos, tutorials, and learning games, creating numerous new ways to train our next generation of doctors, engineers, teachers, and historians. These technologies afford reach to many more people in many more places, all the while reducing the overall cost of education.

Clearly, computer-assisted education has many advantages. But without the presence of a human instructor to ensure that students understand the content behind the video or online tutorial, how do we know if the instruction is effective?

The efficiency of technology-based education can be measured by examining videos of students interacting with a system. For most studies, this results in thousands of hours of video that needs to be transcribed and analyzed for student interactions – a process that can take ten times longer than the length of the video captured.

Automated video analysis

Researchers and experts at the Computer Research Institute of Montreal (CRIM), together with the Learning Environments Across Disciplines research group (LEADS), have created a host of valuable tools for educational researchers that allows them to assess the quality of the educational experience and to fine-tune educational technology for different learning styles. These tools create a consistent way to measure and improve computer-based education, and form the Research Software Platform known as VESTA (Video Evaluation System for Task Analysis).

VESTA distinguishes when the student is reading the screen or thinking aloud (which they are encouraged to do), and when they are paying attention or are distracted. These cues, along with the text transcription of the student's interaction and automatic annotation of notable events (video transitions, facial changes, new individuals

VESTA allows researchers to assess the quality of the educational experience and to fine-tune educational technology for different learning styles.

entering the scene) create catalogued and searchable results for each student video, making research on educational technology significantly easier.

Developing better doctors

Although VESTA is applicable to many types of research requiring video analysis, two current examples are in the medical field. One project helps train prospective doctors in assessing urgent care cases and the other helps them become better at delivering bad news. In both cases, student doctors are filmed as they work through a diagnosis or interact with patients. The resulting analysis not only helps the students become more efficient and empathetic doctors but also helps identify where the training material itself may be improved.

Contributions to other researchers

In addition to the VESTA Platform itself, a number of Software Services used within it have been made available to other researchers via the CANARIE Software Registry. These include tools for annotation storage, facial analysis, video transition detection, load balancing, multimedia file storage, speech segmentation, speaker discrimination, speech-to-text, and text-to-audio matching.

These new Services are available for reuse by researchers across many scientific disciplines, in the same way that VESTA made use of Services provided by other researchers through the CANARIE Software Registry.

Technical Details

Platform: VESTA - Video Annotation Processing System

Description	VESTA focuses on developing and deploying innovative web based tools for analyzing and annotating audio and video recordings.
Creator(s)	Computer Research Institute of Montreal (CRIM)
Collaborator(s)	Learning Environments Across Disciplines (LEADS)
Research Subject	Audio and video processing, Cognitive science - fundamental
Managed Versionⁱ	Yes – available via the CRIM website
Self-deployed Versionⁱⁱ	No
Cloud Support	Yes
Host OS	N/A
Licence	Available for VESTA users
URL	canarie.ca/software/vestasystem

Contributed Services:

	Detection Service for Transitions in Video Content	Face Analysis Service	Load Balancing and Access to Automated Annotation Service	Multimedia File Storage Service
Description	RESTful service providing transition detection in video. Transition detection is the process of detecting when a continuous shot ends and the next one starts.	RESTful service providing Facial Analysis of a video recording. This service will provide a list of detected human faces within a given recording, and information about when the faces appear in the recording.	RESTful service providing a simple load balancer for the VESTA platform. This includes: a gateway to queue based distributed processing framework, and an asynchronous API to query long tasks.	RESTful service providing a multimedia file storage system. This API provides storage, retrieval and transcoding functionality for multimedia files.
Category	Data Manipulation	Data Manipulation	Resource / Cloud Management	Data Storage and Retrieval
Research Subject	Computer vision	Speech Analysis	Dynamic horizontal scaling for services	High performance file retrieval
Managed Versionⁱ	Yes – RESTful web service – available via VESTA platform	Yes – RESTful web service – available via VESTA platform	Yes – RESTful web service – available via VESTA platform	Yes – RESTful web service – available via VESTA platform
Self-deployed Versionⁱⁱ	No	No	No	No
Cloud Support	Yes	Yes	Yes	Yes
Host OS	N/A	N/A	N/A	N/A
Licence	Available for VESTA users	Available for VESTA users	Available for VESTA users	Available for VESTA users
URL	canarie.ca/software/detection	canarie.ca/software/faceanalysis	canarie.ca/software/loadbalancing	canarie.ca/software/filestorage

ⁱ Managed version: Creators host a live instance of the software on their infrastructure, available for use by others

ⁱⁱ Self-deployed version: Users host a private instance of the software on their own infrastructure

Contributed Services:

	Speech Diarisation Service	Speech to Text Service	Text/Audio Matching Service
Description	RESTful service providing speaker diarisation for video content. Speaker diarisation is the process of partitioning an input audio stream into segments according to the identity of the person speaking. This service provides facilities to determine at which times each person is speaking.	RESTful service providing Speech to Text transcription. This service takes an audio file and produces a JSON document containing the text of the audio file, and metadata regarding the confidence of the transcription.	RESTful service providing text to audio matching. Using an audio track containing a phrase and text document, the service will report where the audio phrase occurs within the text document.
Category	Data Manipulation	Data Manipulation	Data Manipulation
Research Subject	Speech analysis	Speech analysis	Speech analysis
Managed Versionⁱ	Yes – RESTful web service – available via VESTA platform	Yes – RESTful web service – available via VESTA platform	Yes – RESTful web service – available via VESTA platform
Self-deployed Versionⁱⁱ	No	No	No
Cloud Support	Yes	Yes	Yes
Host OS	N/A	N/A	N/A
Licence	Available for VESTA users	Available for VESTA users	Available for VESTA users
URL	canarie.ca/software/diarisation	canarie.ca/software/speechtotext	canarie.ca/software/matchingservice

Funding for the development of VESTA was provided through CANARIE's Research Software Program.

ⁱ Managed version: Creators host a live instance of the software on their infrastructure, available for use by others

ⁱⁱ Self-deployed version: Users host a private instance of the software on their own infrastructure