



Harvard John A. Paulson
School of Engineering
and Applied Sciences

PAVVY: Learning Tools for Lecture Content

Pavlos' Perceptron Pals (Ben, Caleb, Chase, Cyrus, Warren)



Overview

1. Background & Motivation
2. User Case Scenario
3. Challenges & Scaling
4. Demo



Harvard John A. Paulson
School of Engineering
and Applied Sciences

WHERE
SCIENCE
AND
ENGINEERING
CONVERGE

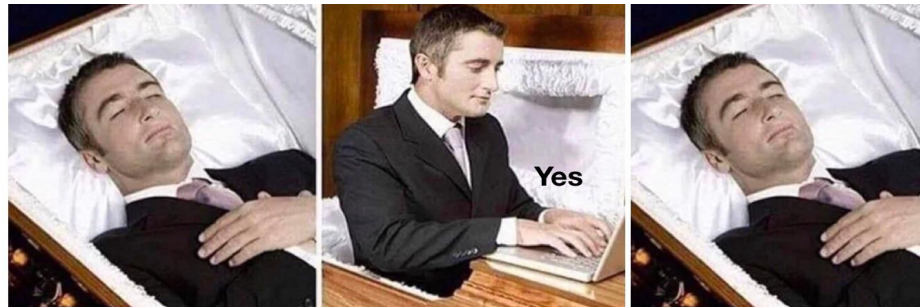
Why?

Motivation

- Since the pandemic, lecture recording & virtual learning have become extremely commonplace
- BUT, **Zoom sessions, recordings, and transcripts are often very non-engaging**

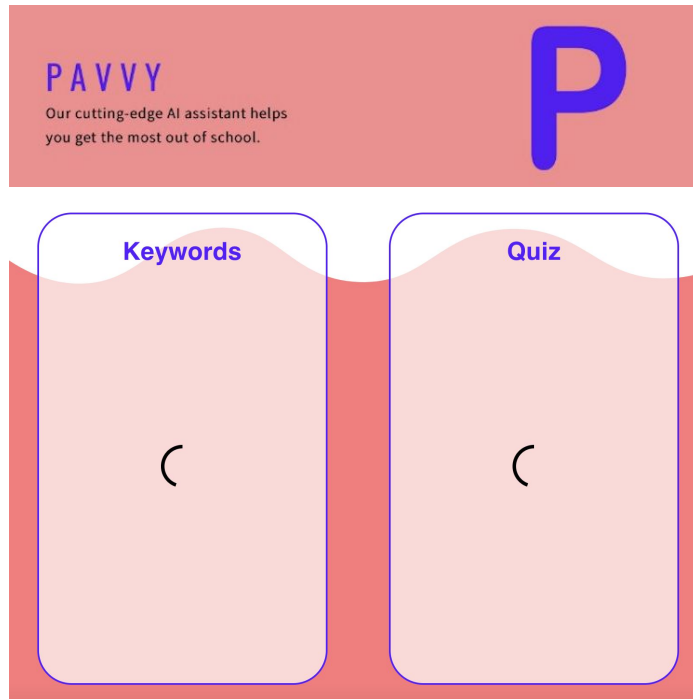
Teacher: Did everyone understand the lesson clearly?

Students during online class:



Solution

- **Develop better learning tools for lecture content!**
- **Keyphrase Extraction:** APA notes that identifying main topics improves studying
- **Quiz Generation:** Harvard's ARC recommends self-testing for deeper learning





Harvard John A. Paulson
School of Engineering
and Applied Sciences

WHERE
SCIENCE
AND
ENGINEERING
CONVERGE

Who?

User Case Scenario

- **Typical Users:** Students, teachers, but, really, all learners!

Imagine you are a student about to start preparing for an exam...



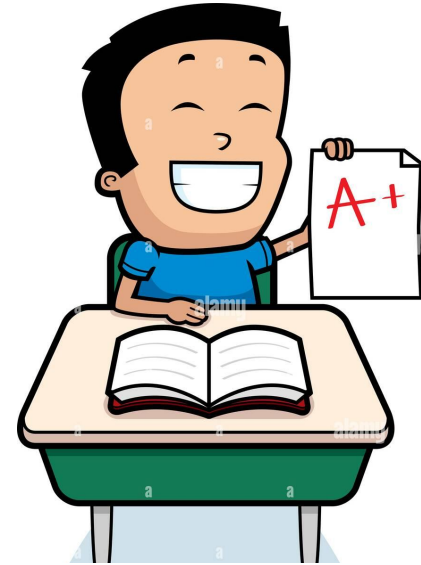
But, once you start, you become overwhelmed, having forgot the big ideas from each lecture and unsure how to assess yourself...



Benefits Summary

- **User Needs Fulfilled:** better way to review and engage with lecture content
- **Key Functionalities**
 - Keyphrase Extraction: isolate big ideas from content
 - Quiz Generation: generate sample questions to simulate exams and reinforce important concepts

With PAVVY, you can review effectively and ace your exam!!





Harvard John A. Paulson
School of Engineering
and Applied Sciences

WHERE
SCIENCE
AND
ENGINEERING
CONVERGE

How?

Cutting-edge Models

- **Lecture Transcription** → OpenAI Whisper JAX
- **Keyword Extraction** → Distilbert
- **Quiz Generation** → ChatGPT 3.5 Turbo



ChatGPT



Harvard John A. Paulson
School of Engineering
and Applied Sciences

Challenges & Scaling

01

Speeding up lecture transcription

- Dask - parallelized lecture preprocessing
- Whisper Jax - 5x faster than standard Whisper

02

Extracting keywords from jargon-intensive content

- Distillation of BERT - high speed and quality
- Inspec - trained our model on labelled scientific abstracts
- Processing of incomplete/misspelled keywords with GPT

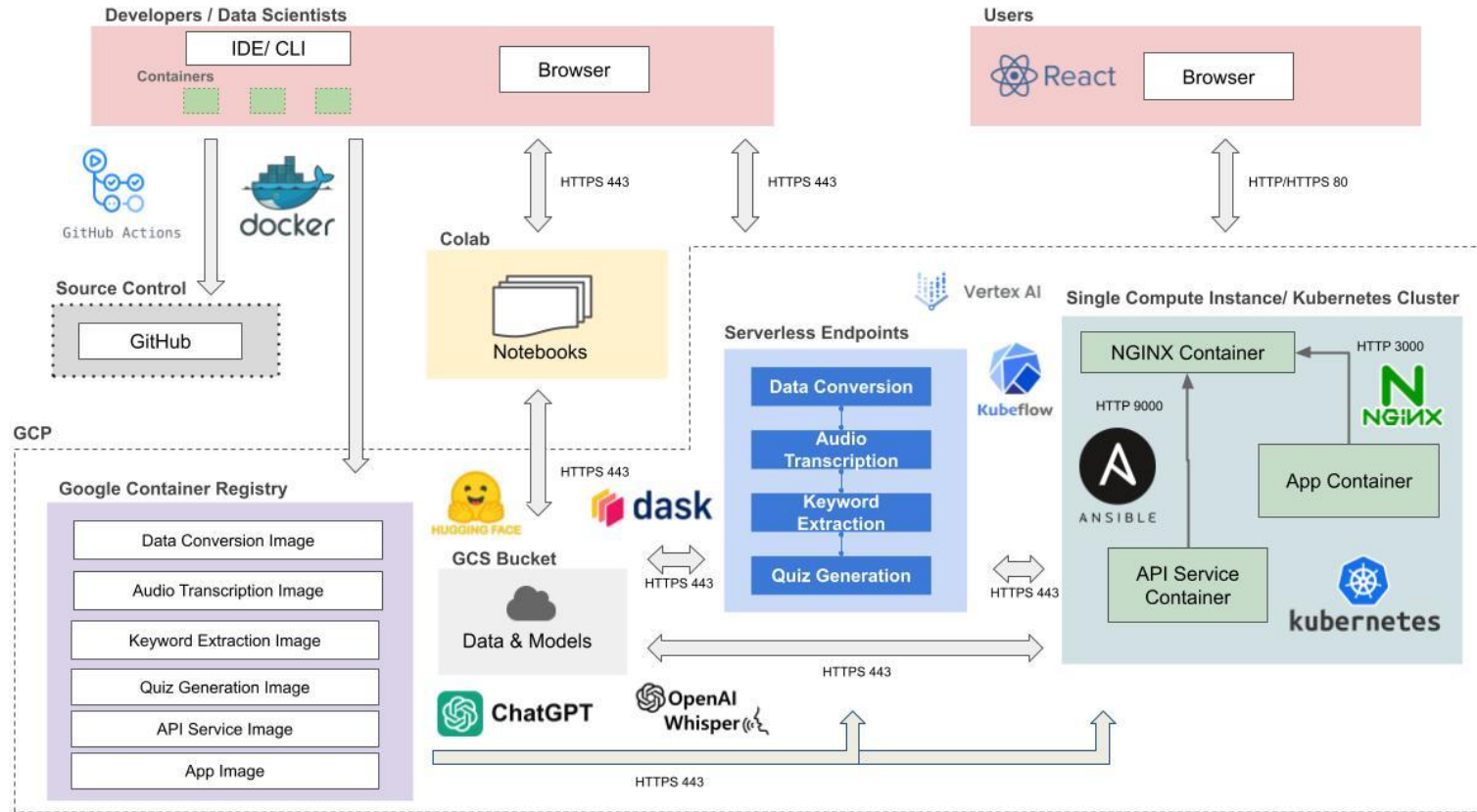
03

Serverless deployment of backend APIs and models

- Model endpoint, Cloud Run, and cloud functions
- Automated deployment to Kubernetes Cluster with Ansible
- Automated CI/CD via GitHub Actions



Technical Architecture





Harvard John A. Paulson
School of Engineering
and Applied Sciences

WHERE
SCIENCE
AND
ENGINEERING
CONVERGE

Demo

References

- <https://www.apa.org/gradpsych/2011/11/study-smart>
- <https://academicresourcecenter.harvard.edu/learning-practices-backed-research>

