

Team 7: TechSmith

Technical Specification / Schedule Cloud Powered Media Searching

Team 7: TechSmith
CSE 498, Collaborative Design

Derek Gebhard
Keith Barber
Scott Lamparski
Caitlin Nelson

Department of Computer Science and Engineering
Michigan State University

FALL 2008



Team 7: TechSmith

Project Overview

- Cloud Powered Media Search Service API
- Amazon Cloud Computing
- Video Splitting
- Optical Character Recognition
- Server Controller
- Website

2



Functional Specifications

- Cloud Powered Media Search Service API
 - Provide interface to upload videos to server
 - Return XML Document

Example:

```
<Video>
  • <Name>TestRecording.avi</Name>
  • <Frames>
    <Frame number="2">
      <Term xLoc="50" yLoc="100">ScreenCapture</Term>
      <Term xLoc="100" yLoc="30">Dialog</Term>
      <Term xLoc="1" yLoc="10">Help</Term>
    </Frame>
    <Frame number="32">
      <Term xLoc="10" yLoc="13">Times</Term>
      <Term xLoc="62" yLoc="301">Snag</Term>
      <Term xLoc="93" yLoc="87">File</Term>
    </Frame>
  </Frames>
</Video>
```

Team 7: TechSmith

3



Functional Specifications

- Amazon Cloud Computing
 - Amazon EC2
 - Run video splitting code
 - Run OCR code on image
 - Amazon S3
 - Store Videos
 - Store Images
 - Amazon SQS
 - Job list

Team 7: TechSmith

4

S

Functional Specifications

Team 7: TechSmith

- FFMPEG Module
 - Input: Video
 - Output: All frames of the video
- OCR Module
 - Input: Image
 - Output: Text contained within Image
- Server Controller
 - Monitors videos/frames to be processed
 - Increases/Decreases EC2s based on work load

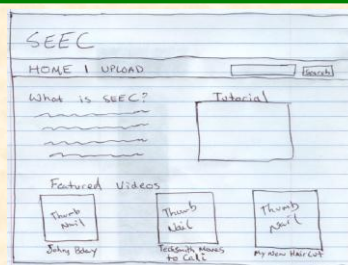
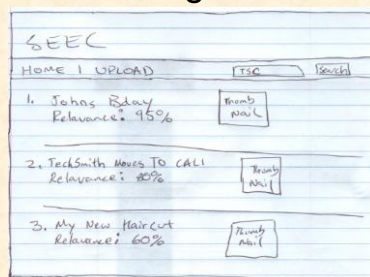
5

S

Functional Specifications

Team 7: TechSmith

- WebSite
 - Home Page
 - Upload Page
 - Search Page
 - Video Page



6

System Components

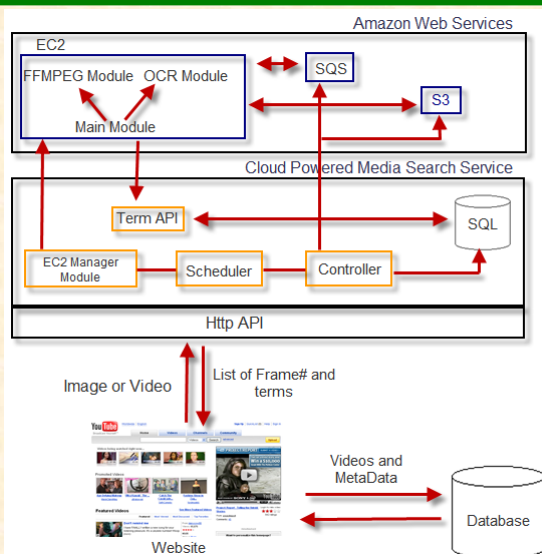


- Hardware Platforms
 - Linux Debian 4.0 Server
 - Amazon S3
 - Amazon SQS
 - Amazon EC2
- Software Platforms / Technologies
 - C++
 - Python
 - Postgres
 - PHP
 - Silverlight
 - FFMPEG
 - OCROPUS

Team 7: TechSmith

7

Architecture Illustrated



Team 7: TechSmith

8

Risks




Team 7: TechSmith

- Performance
 - Processing OCR might be take a while
 - Developing parallel strategy
- OCR
 - No experience and no idea how well it will work
 - Early development
- Language and Library interoperability
 - Multiple libraries and languages will interact
 - Modular design and research
- Flash / Silverlight
 - Client side web code needed for custom player
 - Research and previous TSC project

9

Project Schedule



Team 7: TechSmith

1. EC2 Setup
 - a) Configure Amazon Linux image
 - b) Date: 9/11/2008
2. Upload Interface Code
 - a) Being able to upload videos to S3 via code
 - b) Date: 9/11/2008
3. OCR Sample Program
 - a) Simple program that uses an OCR API to get list of terms on an image
 - b) Date: 9/17/2008
4. FFMPEG Sample Program
 - a) Simple program that uses FFMPEG API to split video
 - b) Date: 9/17/2008

10

Project Schedule

Team 7: TechSmith

5. Create Schema for Service Database
 - a) Description: Database layout for Cloud search service
 - b) Date: 9/17/2008
6. EC2 Manager
 - a) Description: Handles the starting, stopping, and monitoring of EC2 images
 - b) Date: 9/19/2008
7. OCR Module
 - a) Make simple OCR code modular so other code can use it
 - b) Date: 9/22/2008
8. FFMPEG Module
 - a) Make simple FFMPEG code modular so other code can use it
 - b) Date: 9/22/2008

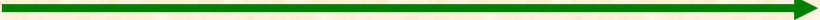
11

Project Schedule

Team 7: TechSmith

9. Create SQL Commands
 - a) Have SQL commands complete for service database
 - b) Date: 9/24/2008
10. OCR Testing
 - a) Complete testing on OCR Module
 - b) Date: 9/26/2008
11. FFMPEG Testing
 - a) Complete testing on FFMPEG Module
 - b) Date: 9/26/2008
12. Scheduler
 - a) Handles when the Amazon's Cloud should process videos
 - b) Date: 9/26/2008

12



Project Schedule

S

Team 7: TechSmith

13. Http API into Service Database
 - a) Provide a way to insert data into service database through Http interface
 - b) Date: 9/29/2008
14. Controller
 - a) Main component of the service
 - b) Date: 10/03/2008
15. EC2 Main Module
 - a) Runs on the EC2. Fetches jobs, calls OCR and FFMPEG
 - b) Date: 10/06/2008
16. Alpha Demo
 - a) Class Demo
 - b) Date: 9/29/2008

13



Project Schedule

S

Team 7: TechSmith

17. Website – Home Page
 - a) Date: 10/9/2008
18. Website – Banner and Nav Bar
 - a) Date: 10/9/2008
19. Website – Database Schema
 - a) Date: 10/9/2008
20. Website – Upload Page
 - a) Date: 10/16/2008
21. Website – Search Page
 - a) Date: 10/16/2008
22. Website – Video Page
 - a) Date: 10/24/2008
23. Website – Silverlight Player
 - a) Date: 10/28/2008

14



Project Schedule

24. Beta Demo

a) Date: 11/3/2008

25. Video Term Highlighting

a) Date: 11/19/2008

26. Cloud video processing optimization

a) Date: 11/24/2008

Team 7: TechSmith

15