Converting PDF at Scale

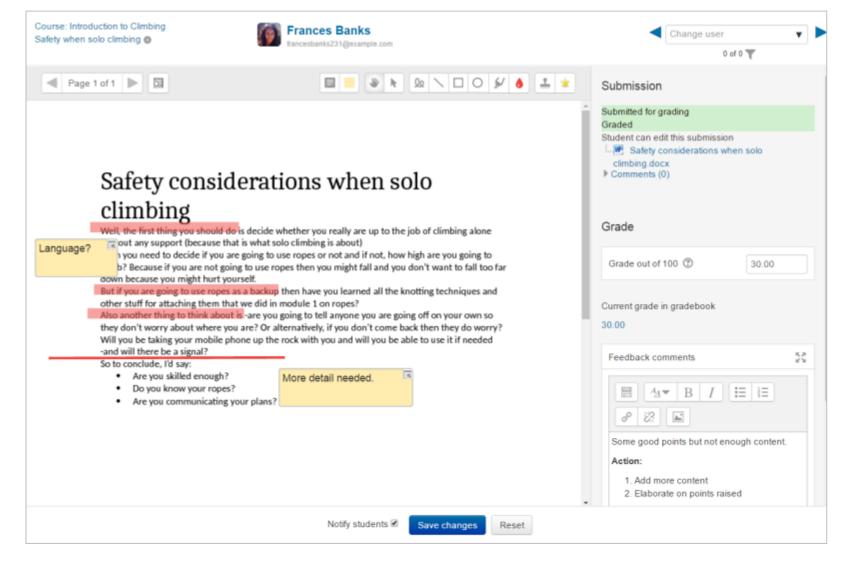
Presented By Matt Porritt



@MatthewPorritt









PDF Conversion

- Non-PDF Documents Must Be Converted
- Conversion Occurs Outside Moodle
- Conversion Process:
 - 1. Student Submits Non-PDF Document
 - 2. Document Converted By Moodle
 - 3. Teacher Annotates PDF





Core Converters

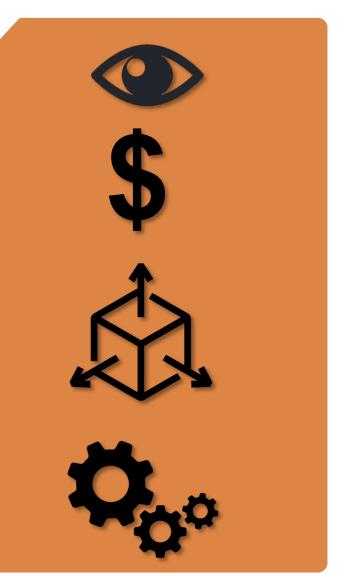
- Unoconv
 - LibreOffice
 - Need To Setup And Manage
- Google Drive
 - Managed Service





Replacement Requirements

- Privacy
- Cost
- Scalability & Performance
- Management & Set Up

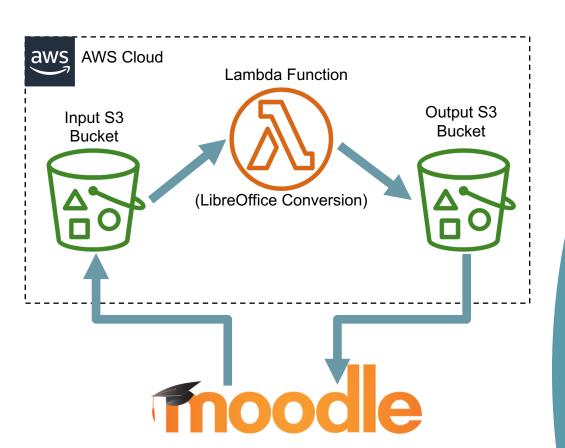


LibreLambda Converter Plugin

- Converts Documents To PDF
- Conversion Via LibreOffice
- Uses AWS Lambda



LibreLambda Converter Architecture





Cost Profiling

- Transfer Internet to S3
 - Put object request: \$0.0055 per 1,000 requests
 - All data transfer in \$0.00 per GB
- Transfer to lambda
 - File transfer: free
 - Get Request \$0.00044 per 1,000 requests
- Lambda conversion
 - Invocations:
 - First 1M requests per month are free.
 - \$0.20 per 1M requests thereafter
 - Execution cost:
 - 400,000 GB-Seconds per month free
 - \$0.0001667 every GB-Second used thereafter
- Transfer from Lambda to S3
 - File transfer: free
 - Put object request: \$0.0055 per 1,000 requests
- Per Bucket S3 Storage Cost
 - Storage Cost: \$0.025 per GB
- Transfer S3 to Internet
 - Get Request \$0.00044 per 1,000 requests
 - Up to 1 GB / Month \$0.00 per GB
 - Next 9.999 TB / Month \$0.114 per GB

(Prices in AUD)



Cost Profiling

- Converted 100,000 Documents
- Total Cost:
 - \$7.31AUD
- Cost per Document:
 - \$0.0000731AUD



Scalability and Performance

- Converted 100,000 Documents
- Total Conversion Time: ~120 minutes
- Time Per Document: ~7 seconds
- Running Multiple Threads
- No Need To Manually Scale



Management & Setup

- Plugin Includes:
 - Setup Script
 - Test Script
- Moodle Settings Automatically Set



Source Conversion Formats

- Plugin Converts:
 - doc

• ppt

• txt² • png²

docx

- pptx
- html¹

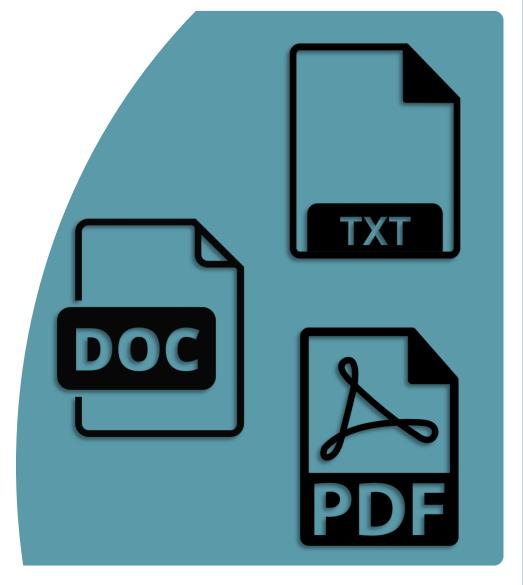
x|s

x|sx

rtf

- odt
- ods
- jpg²
 - gif²

- HTML Markup To Text (Not Rendered)
- Thanks to University College London



Next Steps

- Client Rollout
- More Development Cycles
 - More file formats to PDF
 - Convert to other formats
- Other Platform Support



Links

- https://moodle.org/plugins/fileconverter_librelambda
- https://github.com/catalyst/moodle-fileconverter_librelambda
- https://docs.moodle.org/36/en/Using_Assignment#Annotating_submissions

