**Software Requirements Specification**

**for**

Box.com/eDefender Integration

**Version 1.1 approved**

**Prepared by** **Joshua Cabrera, Marco De La Torre, Raul Gallegos, Daniel Guevara-Dominguez, Chuang Huang, Dang Le, Jesica Lopez De Leon, Shaocheng Shi, Sergio Tapia, Luke Williams**.

Santa Barbara Public Defender Office

**October 7, 2021**

**Table of Contents**

[**Revision History**](#_sxceh2dujzh2) **3**

[**1. Introduction**](#_evu4vla744lg) **4**

[1.1 Purpose](#_q3conn35w42q) 4

[1.2 Intended Audience and Reading Suggestions](#_x3la3zkb1c5n) 4

[1.3 Product Scope](#_zc6d0ki3een7) 4

[1.4 Definitions, Acronyms, and Abbreviations](#_5bfxx99a30c5) 5

[1.5 References](#_ye42ljm9sha4) 5

[**2. Overall Description**](#_byen67io4mvi) **6**

[2.1 System Analysis](#_stkvyxvyb5h) 6

[2.2 Product Perspective](#_ojnkz4bugjg3) 6

[2.3 Product Functions](#_asxj4pgu4emb) 6

[2.4 User Classes and Characteristics](#_uc8ktoax6smw) 6

[2.2 Operating Environment](#_wzidvsb4mu15) 6

[2.6 Design and Implementation Constraints](#_nzx0gwclm8ku) 6

[2.7 User Documentation](#_yjd41ryy5svf) 7

[2.8 Assumptions and Dependencies](#_aoycjjepyzk8) 7

[2.9 Apportioning of Requirements](#_7tkyweo9rw80) 7

[**3. External Interface Requirements**](#_lrrohnxf3371) **8**

[3.1 User Interfaces](#_6nb9bmyhh1d9) 8

[3.2 Hardware Interfaces](#_eoddzj1z7qat) 9

[3.3 Software Interfaces](#_3k6lyvhqe2um) 9

[3.4 Communications Interfaces](#_1ilqktim1d0m) 9

[**4. Requirements Specification**](#_ekz9zo8ijjcy) **10**

[**4.1 Functional Requirements**](#_o4evd4l7rkt0) **12**

[4.2 External Interface Requirements](#_z884ftbtodng) 14

[4.3 Logical Database Requirements](#_7mxo32fof621) 14

[4.4 Design Constraints](#_czbteyt8l4w6) 14

[**5. Other Nonfunctional Requirements**](#_2qlzqlnqm8uv) **15**

[5.1 Performance Requirements](#_sequ2tkv86zi) 15

[5.2 Safety Requirements](#_537mxb80sxp5) 15

[5.3 Security Requirements](#_108xpnsjrcqp) 15

[5.4 Software Quality Attributes](#_l993xe839x76) 15

[5.5 Business Rules](#_ou26xt2n5r1u) 15

[**6. Legal and Ethical Considerations**](#_cbo5l5aq9hom) **16**

[**Appendix A: Glossary**](#_z3xw0pvhv4ap) **17**

[**Appendix B: Analysis Models**](#_67o6f78ugv04) **17**

[**Appendix C: To Be Determined List**](#_5p7n2jlgizzk) **17**

# Revision History

| **Name** | **Date** | **Reason For Changes** | **Version** |
| --- | --- | --- | --- |
| Team | 10/7/21 | First Draft | 1.0 |
| Team | 12/7/21 | Update Requirements, Project Scope | 1.1 |
|  |  |  |  |
|  |  |  |  |

# 1. Introduction

The purpose of this Box.com and Edefender integration software is to serve as an intermediary between Box.com and E-Defender in order to provide automatic transcription services and other advanced machine learning functionality.

The second software is the document tag parser app. Its purpose is to extract pdf files tag number and then rename the file with the tag number included.

### 1.1 Purpose

The purpose of this document is the following:

1. Identify the requirements for Box.com and eDefender integration (Version 1.1).
2. Identify the needed software and how it is being used.
3. State how the user is meant to interact and use the product.
4. How the product will interact with any external service.
5. Define and describe concepts that will clarify the functions of the product.

This software borrows heavily from previous implementations that provided similar functionality for the City of Los Angeles. This document covers the features that were implemented on behalf of the City of Santa Barbara by California State University, Los Angeles which includes facial recognition, voice transcription, pdf text transcription and tagging, and translation for supported languages.

### 1.2 Intended Audience and Reading Suggestions

This document is intended for developers, testers, project managers, and users to understand the purpose, function, and requirements of the product. Suggested reading sequence for each is the following:

* Developers - Understand the purpose and functions of the product to help during development; Recommended Reading: REST + API Documentation
* Testers - Understand the purpose and requirements of the product to help during testing to guarantee intended function; Recommended Reading: REST + API Documentation
* Users - Understand the purpose of the product and its functions; Recommended Reading: Section(s) 1 - 5

### 1.3 Product Scope

Box.com and E-Defender integration will allow for reduced workloads on lawyers and public defenders by providing transcription services for digital audio and/or video evidence. The transcription will include timestamps for the dialogue as well as faces of parties involved.

The document tag parser app will help speed up the process for public defenders by eliminating the process of manually renaming each file one by one, saving a lot of time.

### 1.4 Definitions, Acronyms, and Abbreviations

* See Appendix A: Glossary

### 1.5 References

* REST+API.doc
* user\_manual\_box\_indexer.pdf
* https://journaltech.com/
* https://box.com
* <https://aws.amazon.com/lambda/>

# 2. Overall Description

Box Skills makes use of third-party AI/ML to get metadata within media files, it is not a product on its own and requires the use of other APIs. Box skills node is a middleware and needs to be hosted on a serverless function.

### 2.1 System Analysis

* The goal is to provide automatic transcription services for audio and/or video files through Box.
* File extension support may vary, and errors or debugging is not covered.
* Our design right now is to have AWS Lambda as our serverless function and Azure Video Analyzer to analyze .
* Pdf files will contain tag numbers that will be extracted from the pdf and added to the name of the pdf file.

### 2.2 Product Perspective

Box Skills acts as middleware between third-party AI/ML services and other API technologies. As a result, Box Skills needs to be hosted on a serverless function.

Many other solutions exist, including but not limited to, Google and Microsoft transcription services. Prices will differ and Box was chosen by sponsor for their skills technology.

### 2.3 Product Functions

* Transcription of audio and/or video files into English, Spanish.
* Facial recognition and indexing, providing timestamps
* Transcription of pdf tag and renaming of pdf files based on tag.

### 2.4 User Classes and Characteristics

This software is designed with Public Defenders as the main users/recipients.

### 2.2 Operating Environment

This software will operate on serverless technologies and cloud-based networks.

### 2.6 Design and Implementation Constraints

There are associated costs for each video being transcribed of $0.0000166667 per GB-s as of December 2021. Additionally, not all languages are as readily transcribable. For languages that are supported, accuracy of transcription services is not guaranteed to be 100% accurate. Access to an internet connection is required for accessing Box and E-Defender.

### 2.7 User Documentation

* REST+API.doc
* user\_manual\_box\_indexer.pdf

### 2.8 Assumptions and Dependencies

* AWS Lambda retrieves a file from Box and analyzes it before the metadata is sent back to Box from AWS.
* The Box Skill Application will need to be hosted on the Serverless Function AWS Lambda.

### 2.9 Apportioning of Requirements

* Transcribing multiple languages
* Sending a notification of completion to desired parties involved after transcription

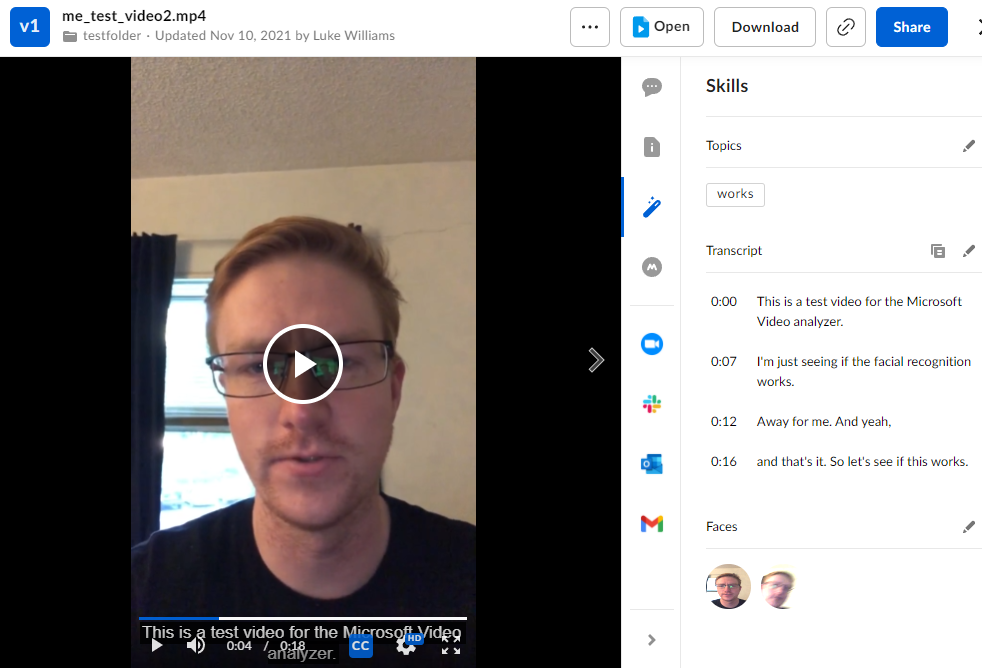
# 3. External Interface Requirements

All aspects of the Box skill will be running and controlled through the cloud and webapps. Once deployed, the Box skill will run automatically once a file is uploaded to Box. The document tag parser app will be a desktop program installed on the machine of the user.

### 3.1 User Interfaces

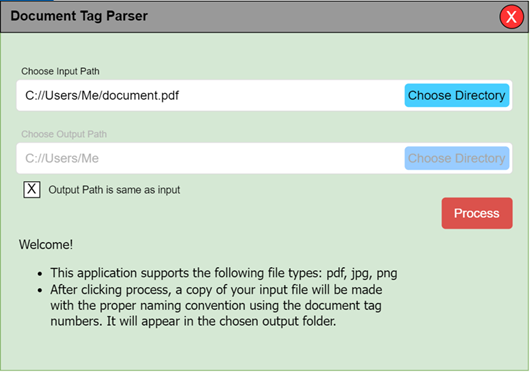
**BOX/e-Defender Integration:**

The user must have a Box account given to them by the Santa Barbara County Public Defender Office. With the account, the user may login and upload a video to its respective skill enabled folder. The user will receive an eDefender notification of the action, and now only has to wait for a confirmation of the file being successfully uploaded and transcribed. Once the confirmation is given, the metadata will be displayed on the side as a skill.



**Document Tag Parser App:**

The user will interact with a desktop program that is installed on the user's machine. The program lets the user choose the file directory they would like to read from, and the directory to place the text file into.



### 3.2 Hardware Interfaces

N/A

### 3.3 Software Interfaces

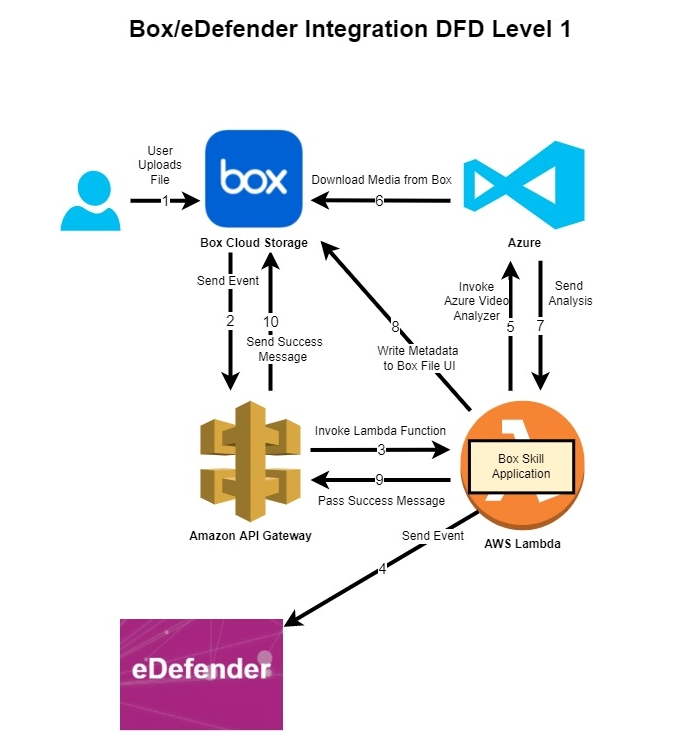
N/A

### 3.4 Communications Interfaces

Box invokes AWS, which then invokes the lambda function and then finally the rest of the program. For this to happen, these portions of the program will need to be interconnected via security keys, invocation URLs, and access tokens.

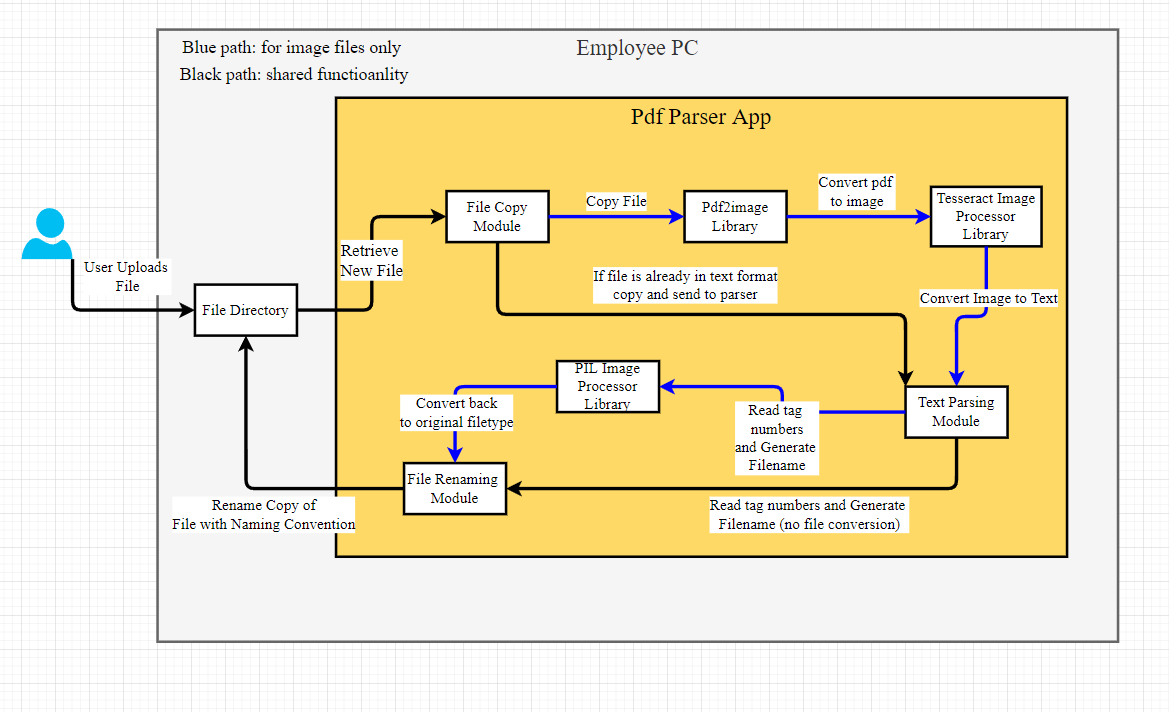
# 4. Requirements Specification

**BOX/e-Defender Integration:**



Users should be able to upload video and audio to Box and Box custom skill should trigger upon uploading files. Box skill should send access tokens to serverless hosting sites(AWS), and later pass to ML service(Azure Video Indexer) for file retrieval. ML service should produce metadata including transcription, facial recognition, and topic detection and the hosting site should write that metadata back to the video in Box. A notification should be sent to e-Defender once the video is ready to view with all the metadata embedded to it.

**Document Tag Parser(DTP):**

****

This is a local executable program which prompts users to parse any pdfs and images they want into a text file. Furthermore, the program will accept the disc number and find the beginning and ending tag numbers within the file and use that information to rename the file using SBPD’s naming convention. Most importantly, the original file will not be modified during the whole process.

# 4.1 Functional Requirements

**BOX/e-Defender Integration:**

| **1.1:Box** | |
| --- | --- |
| 4.1-1.1-1 | Box shall allow users upload and download files |
| 4.1-1.1-2 | Box shall trigger BCS upon file upload to pre-configured folder |
| 4.1-1.1-3 | Box shall send access token to hosting site(AWS) |
| 4.1-1.1-4 | Box shall embed metadata to original video/audio |
| 4.1-1.1-5 | Box shall visualize metadata as transcript, face, and topic card |

| **1.2:AWS** | |
| --- | --- |
| 4.1-1.2-1 | AWS shall host the Box skill application |
| 4.1-1.2-2 | AWS shall capture and parse all information of an event send by Box |
| 4.1-1.2-3 | AWS shall send an event to ML service(Azure Video Analyzer) |
| 4.1-1.2-4 | AWS shall send access token to ML service |
| 4.1-1.2-5 | AWS shall receive response from ML service |
| 4.1-1.2-6 | AWS shall retrieve metadata from ML service |
| 4.1-1.2-7 | AWS shall write metadata back to Box |

| **1.3:ML service(AVA)** | |
| --- | --- |
| 4.1-1.3-1 | ML service shall listen for AWS call to process video/audio |
| 4.1-1.3-2 | ML service shall use access token to retrieve file from Box |
| 4.1-1.3-3 | ML service shall produce transcription of a file |
| 4.1-1.3-4 | ML service shall perform language auto detection |
| 4.1-1.3-5 | ML service shall recognize faces and keywords |
| 4.1-1.3-6 | ML service shall provide timestamp for analyzed result |

**Document Tag Parser(DTP):**

| **1.1:Document Tag Parser(DTP)** | |
| --- | --- |
| 4.1-1.1-1 | DTP shall convert pdfs and images into text files for parsing |
| 4.1-1.1-2 | DTP shall read the tag numbers on the first and last page of each document |
| 4.1-1.1-3 | DTP shall rename the file based on the SBPD’S naming convention |
| 4.1-1.1-4 | DTP shall add the disc number based on the user’s choice within the application |
| 4.1-1.1-5 | DTP shall not modify the original file |
| 4.1-1.1-6 | DTP shall create new copies with the appropriate file naming convention |

### 4.2 External Interface Requirements

**BOX/e-Defender Integration:**

* Files upload to Box restricted to video or audio only
  + Acceptable video formats: aac, aif, aifc, aiff, amr, au, flac, m4a, mp3, ogg, ra, wav, wma.
  + Acceptable audio format: 3g2, 3gp, avi, m2v, m2ts, m4v, mkv, mov, mp4, mpeg, mpg, ogg, mts, qt, ts, wmv.

**Document Tag Parser(DTP):**

* Pdfs, images.

### 4.3 Logical Database Requirements

N/A

### 4.4 Design Constraints

**BOX/e-Defender Integration:**

* ML services do not provide 100% accurate transcription.
* Inaccuracy increases with multiple languages, voices, and background noises.
* Language auto detection happens when language is not specified, thus increasing inaccuracy and wrong language when generating transcripts.
* Faces recognition not 100% working in picking up faces.
* Topic detection could be irrelevant.

**Document Tag Parser(DTP):**

* Restriction to the materials could be parsed by the program; for example, pictures, diagrams, and others that are not text.

# 5. Other Nonfunctional Requirements

### 5.1 Performance Requirements

* Media metadata extraction time should be at least 1:1 with video length.
* Accuracy should be up to standard with current transcription services.
* Media file will be handled box skills enterprise plan uploaded and handled can be up to 150 gb.

### 5.2 Safety Requirements

N/A

### 5.3 Security Requirements

Files distributed outside of Box should be deleted after metadata is extracted, as they are often important evidence of ongoing cases.

### 5.4 Software Quality Attributes

Box Skill Application is flexible and can be hosted in a number of ways. On a server, on serverless functions, or on a cloud server. AI/ML services can also be replaced. This process is simple as it only needs to be directed by AWS Lambda to make its call to a specific AI/ML.

### 5.5 Business Rules

N/A

# 6. Legal and Ethical Considerations

N/A

# Appendix A: Glossary

* SRS: Software Requirements Specification
* AI/ML: Artificial Intelligence/Machine Learning
* Box: Cloud Storage provider
* Box Skills: A framework used to provide customization for folder/files on Box.
* AWS: Amazon Web Services
* AWS Lambda: A serverless, event-driven compute service
* E-Defender: supplies case management software to courts, prosecutors, public defenders, probation and other justice agencies.
* BCS: Box Custom Skill
* AVA: Azure Video Indexer
* DTP: Document Tag Parser

# Appendix B: Analysis Models

* Refer to flow diagram in Section 4

# Appendix C: To Be Determined List

* document parser app to extract tags from documents and append it to file names