**Documentation for the Entire Process**

**Create CSV File**

**Website Visit:** visiting a specific site, such as sec.gov.bd.

**PDF Download:**

>Navigate to the Enforcement Action field.

> Download the PDF file to a designated folder.

**PDF Conversion:**

> Download the PDF image for conversion.

> Convert the PDF image to text online.

**Data Organization:**

>Save the text in Google with the existing .csv format (id, content, Name, page, and URL)

> Content**:** paste this content on the existing excel file

> **Page Number:** Add page number from the content field (Content Serial number is page number)

> **Name:** name field will use pdf file name

> **URL:** https://jupyter.tappware.com/files/ /bsec\_pdf\_4/ Mahabub\_Alam\_09.05\_.2023\_.pdf

**https://jupyter.tappware.com/files:** (By default, we use this format for every file).

**(bsec\_pdf\_4>:** only pdf number will be change.

**Mahabub\_Alam\_09.05\_.2023\_.pdf:** it will be name of the pdf.

**Database Operations:**

Share the PDF and CSV files based on the infra team's requirements.

Log in to the database using PHPMyAdmin

URL: https://nothi-web.tappware.com/pmad/

Username: [Your Username]

Password: [Your Password]

Enter (bsec\_db).

>import

>go

Import the CSV file into the bsec\_db.

**Server Tasks:**

Log in to the server (logstash server: 172.17.130.77 - mycourt).

>Open CMD.

>SSH root@172.17.130.77.

**Backup:**

Index:

Menu>

>dev tools

>Console run this Command:-

POST \_reindex

{

source: {

index: "bsec\_indexing"

},

dest: {

index: "bsec\_indexing\_bk\_date"

}

}

>>run

**Configuration on Test Server:**

>/etc/logstash/conf.d

>nano bsec.conf

Statement: add table name such as (bsec\_3)

Output:

Index name will be (index => "bsec\_indexing\_3")

>save and exit (ctrl+x),y and enter.

Run config file by this command:

>logstash -f bsec.conf

15. Verify the backup by checking the index location and authentication:

https://search-nothi-test-kibana.tappware.com/

- User: [Your Username]

- Pass: [Your Password]

-**View Index file:**

>menu > stack management> data and click (index management)> then view all index file and search index file from the searching menu.

> View Pdf data from the Discover menu.

**Live Configuration:**

For live configuration, revert changes in bsec.conf and set the index to "bsec\_indexing".

>/etc/logstash/conf.d

>nano bsec.conf

Statement: add table name such as (bsec\_3)

Output:

Index name will be (index => "bsec\_indexing")

>save and exit (ctrl+x),y and enter.

> run config file by this command

> command for run:

logstash -f bsec.conf

**Upload PDF to Server:**

17. Upload the PDF file to the server (172.17.130.12) at

> /var/www/html/bsec-datawatch/public/assets/pdf/bsec.

Index Check:

18. After completing the server part, check the index at

https://bsec-smart-search.bokhshi.com/.

- File Content: Use any text from the content file and search in the provided field.