**UI PATH PDF AUTOMATION**

In the ui path automation the datas are stored by a pdf.it has been working many steps.

The pdf activities are classified into two types

* Extracting large texts
* Extracting specific elements

**EXTRACTING LARGE TEXTS**

PDF datas extraction is performed through native pdf text and OCR technology.

1. **Read pdf text**: for files with text only.
2. **Read pdf with ocr**: for files with text and image.

**Steps in read pdf text**

* Create a blank project
* Check the pdf packages are installed or not.
* Search sequence then drag and drop
* Search flow chart drag and drop
* Sequence is putit into the flowchart
* Change the sequence name
* Drag and drop read pdf text,insert the pdf path
* Drag and drop write text,the insert the path.click the right button then set the output
* Drag and drop message box
* Click read pdf activity insert the input range value and without find the image in the pdf file
* click debug to see the message are extract in the message box

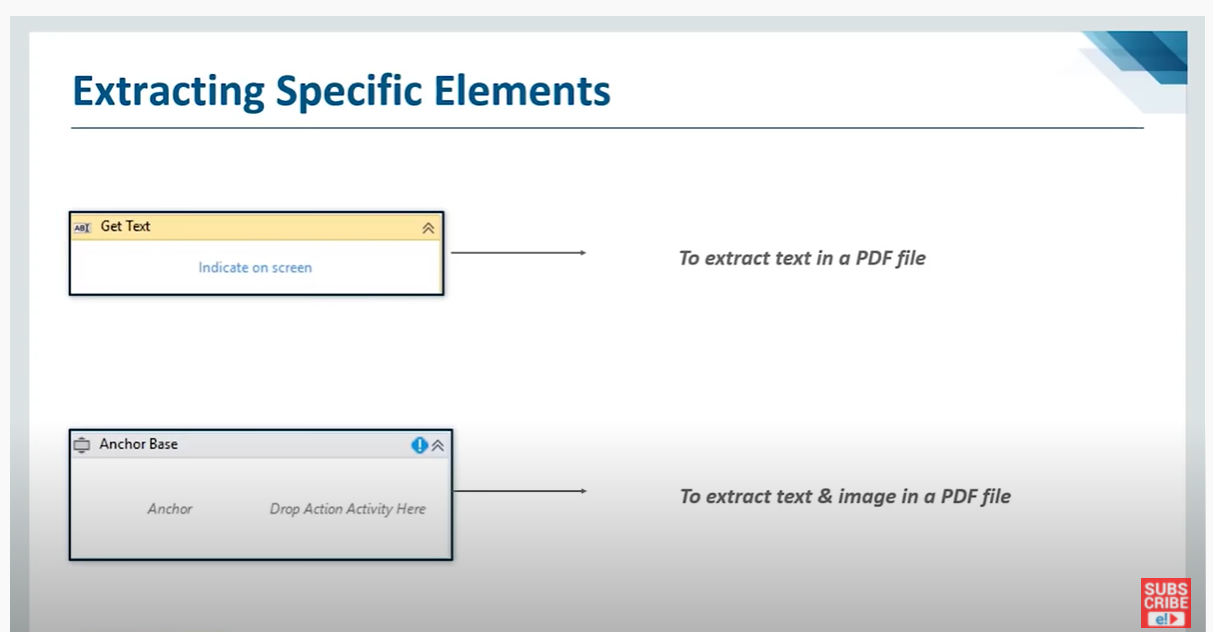
**steps in pdf with ocr**

* Go to the flowchart activity and drag and drop down the sequence.Mention the name of the sequence
* Double click to that sequence and drag and drop **read pdf with ocr** activity. mention the file path
* Search the ocr engine and check the ocr engines are installed or not.then Drag and drop **google ocr**
* search **meassage box** drag and drop down and drag and drop **write text file** and mention the output value.
* Fill the google ocr and message box.execute the project and see the complete pdf is to be extracted

**EXTARCTING SPECIFIC ELEMENTS**

Extracting specific activity can be classified into

* Get Text: To extract text in a PDF file
* Anchor base: To extract text & image in a PDF file

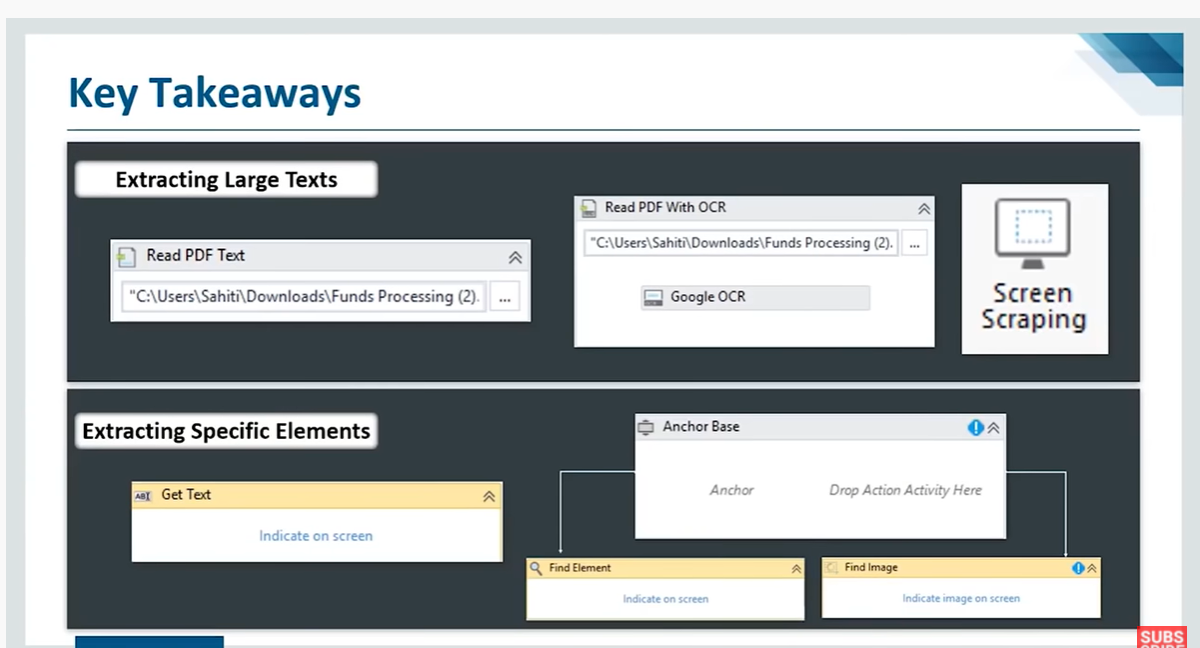
****

**Anchor base**

Anchor base elements activity can be classified into two types

* Find Element: used to find an element
* Find Image: used to find an image

**KEY TAKEAWAYS**

****