Scanning

- Creates an Electronic bit map image of forms
Content Recognition Modules:

- **Form-ID Module**
  - Form Identification
  - Automatically identifies a form using Empty Form Images (EFI) which are defined during the design stage

- **Manual Form-ID Module**
  - When the Form-ID Module cannot identify the form due to poor scanning or other problems, the user can manually match a scanned image with the proper Empty Form Image (EFI)

- **Processing/Recognition Module**
  - Character recognition using OCR/ICR/OMR Recognition engines
Processing / Recognition

- Most important station
- Automatically Recognises data from image
- Automatic recognition is not 100% accurate
- Inaccurate data requires human intervention
- Positive effect on throughput as long as characters incorrectly identified are controlled

OCR/ICR/OMR engines used to recognize characters.
Key from Image

Certain types of data do not lend themselves to automated extraction via a recognition engine of one kind or another.

Cursive handwriting in a form, is an example.

Such data can be extracted efficiently from a form using key from image techniques.

Software displays zoomed predefined areas of an image—or zones—to a human operator for keying. As soon as the contents of the area is keyed, the software automatically displays the next zoomed area, and so on.
Content Validation

- Content Validation includes 3 modules:
  - Tile
  - Completion
  - Exception

Tile

- The unique TILE module optimize data accuracy with a systemized display of characters grouped together to allow easy identification

- Possible to identify which characters are correct and which are not and allows to mark as reject.

- Makes the completion more accurate
Completion

The Completion module
- Receives data from the Tile Station
- Enables the operator to correct unrecognized characters

The data is completed manually with the help of
- Data dictionaries
- Rules
- Validation functions etc.

Modes of display - Page and Field mode

Page Mode Completion

The operator is shown all or part of the form and is given a text box in which to type all or part of the form data, as appropriate.

Field Mode Completion

- The same fields from several pages are Displayed on the screen
- This type of display improves the speed of the Completion process
### COMPLETION STATION [Field mode]

**Image Based Forms Processing**

- **House List**:
  - 016
  - 017
  - 018
  - 019
  - 020
  - 021
  - 022
  - 023
  - 024
  - 025
  - 026
  - 027
  - 028
  - 029
  - 030

- **Enterprise Schedule**

- **Enterprise Abstract**

**Extracted data**

### COMPLETION STATION [Page mode]

**Image Based Forms Processing**

- **Address Slip**

<table>
<thead>
<tr>
<th>ES_No</th>
<th>ES_PageNo</th>
<th>SMo</th>
<th>No_Persons</th>
<th>Rem_Hired</th>
<th>Year_Start</th>
<th>NIC_Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>01</td>
<td>09</td>
<td>0000000</td>
<td>000</td>
<td>1782</td>
<td></td>
</tr>
</tbody>
</table>

**Address 1**

- Name: Name_1
- Address: Address_1

**Address 2**

- PostCode
- STDCode
- TelephoneNo
- STDCode
- FaxNo
- Email
Exception

- When the operator is unable to accurately identify a specific data item, the form is sent to the Exception Module/Station.

- The Exception Module handles forms that require a supervisor's attention and allows the operator to advance to the next form.

**Exception Station**

Form Image

Exception cases
CONTENT DELIVERY

- EXPORT the final data in ASCII or Database file format.