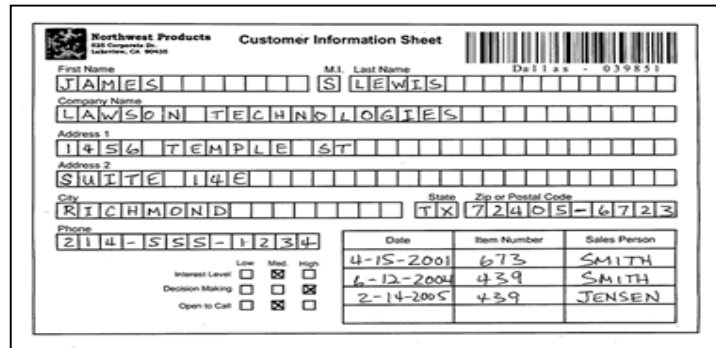# Classification

What is Classification?

- The ability to confidently recognize a document type by its layout, content, or both

KOFAX

Recap - Structured, Semi-structured, Unstructured

- Kofax TotalAgility 7 includes modules and activities that work with Kofax Transformation Designer to provide advanced document classification, separation, and extraction for structured, semi-structured and unstructured documents
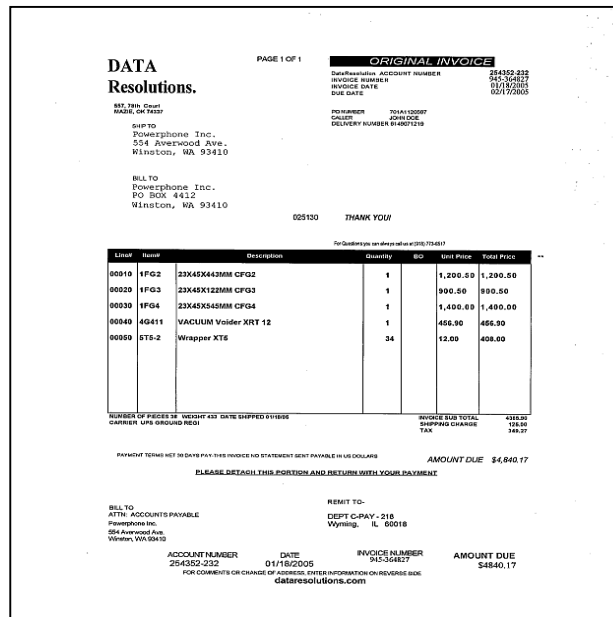

Structured

Semi-structured

Unstructured

Analyze Documents

- Classification
    - Get a large stack of paper and try to sort it

    - Have you recognized it by a specific layout?
        - ➔ Layout classification

    - Have you read the whole document?
        - ➔ Content classification (Adaptive Feature classifier)

    - Have you read just a phrase or a word?
        - ➔ Content classification (Instruction)

KOFAX

Classification

- Transformation Designer supports following classification methods
  - Layout classification
  - Content classification "Adaptive feature classifier"
  - Content classification "Instructions"

KOFAX

Classification

- Select **Layout**, **Content**, or **Both** methods for Classification
- Provide **Training Sample**s for each Document Type
- **Train** the Project
- **Test** Classification at Design Time using **Test Document Sets**
- Examine **Classification Confidence**
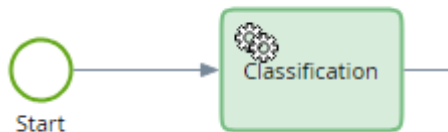- Test Classification at **Run Time** by Creating a Job

In the Training Example we will use Layout Classification. (Content Classification is covered in other courses)

KOFAX

Configuring Classification

- Open the **Classification Group** in the Transformation Designer
- Configure **Classification** at the Project Level
  - Layout only (in this example)
- Obtain good quality **samples** (e.g. at least 1, 300 dpi, black and white)
  - Open the Samples Document Set
  - Select the Class (Document Type)
  - Add the sample(s) to the Classification Set
  - Repeat for each Class
- **Train** the Project
- **Test** on a Document Test Set

KOFAX

Classification Activity

- Automated activity that executes Classification project information
- Requires an Input Folder
- Requires Classification Group (which can be inherited from the Process)
- Separation Profile only required if separation will take place during classification rather than on scan
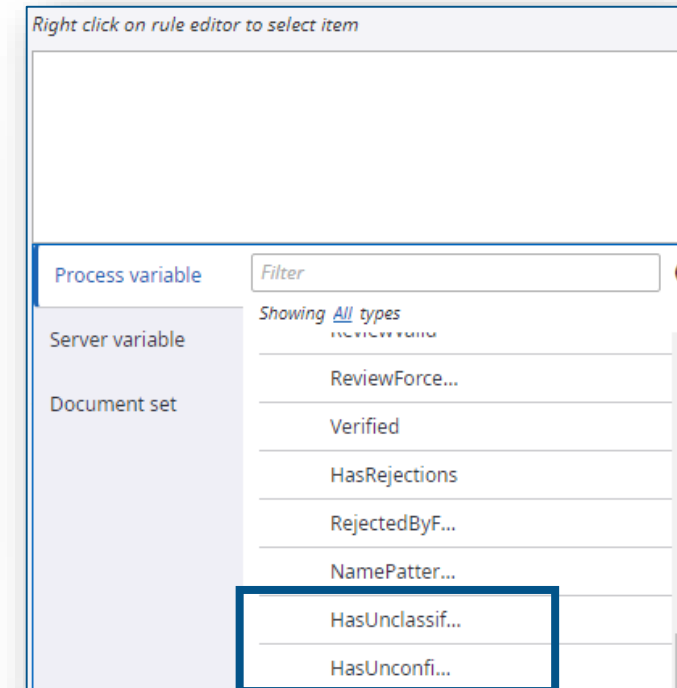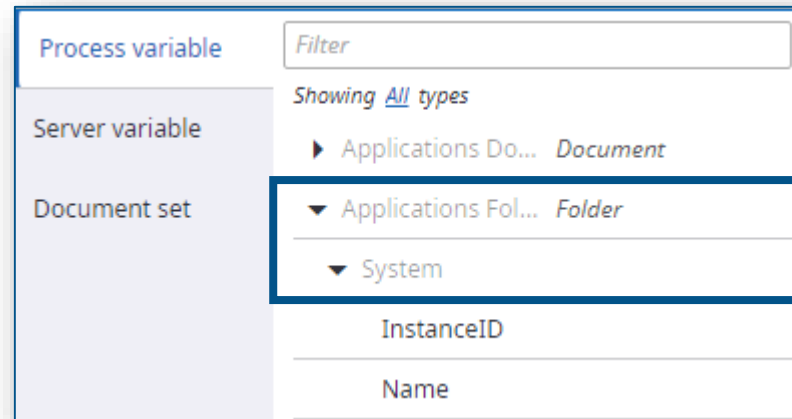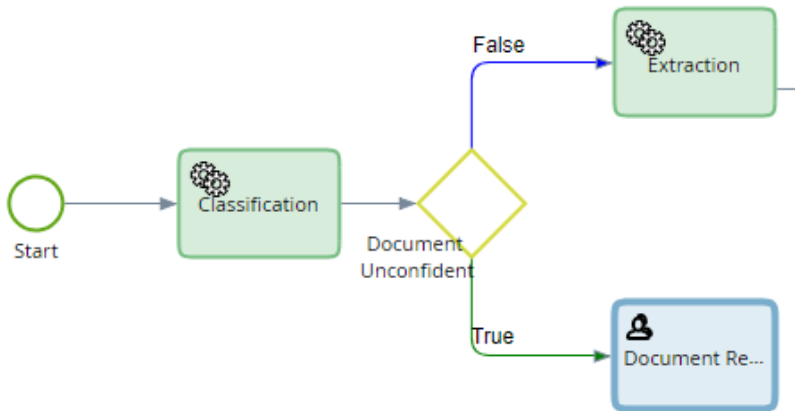
What is Document Review?

- Optional manual activity to **review classification** results

- Document Review can be bypassed by checking the folder system fields: HasUnconfidentlyClassifiedDocument returns a Boolean (True/False) value

- HasUnconfidentlyClassifiedDocument returns a Boolean (True/False) value

- You can use a decision activity to test its value after classification and bypass Document Review if required

KOFAX

Test Classification at Run Time

- You can test classification at run time by creating a job in the TotalAgility Workspace
- The **Document Review** manual activity allows users to review (and modify) classification results

Classification Online Learning

Classification Online Learning

- When a new layout is encountered, the document reviewer can manually change its type. The next time this document is encountered it will automatically be classified
  - Online Learning must be enabled in the Classification project
  - Monitor this class (document type) for Classification Online Learning must be enabled in the Extraction project
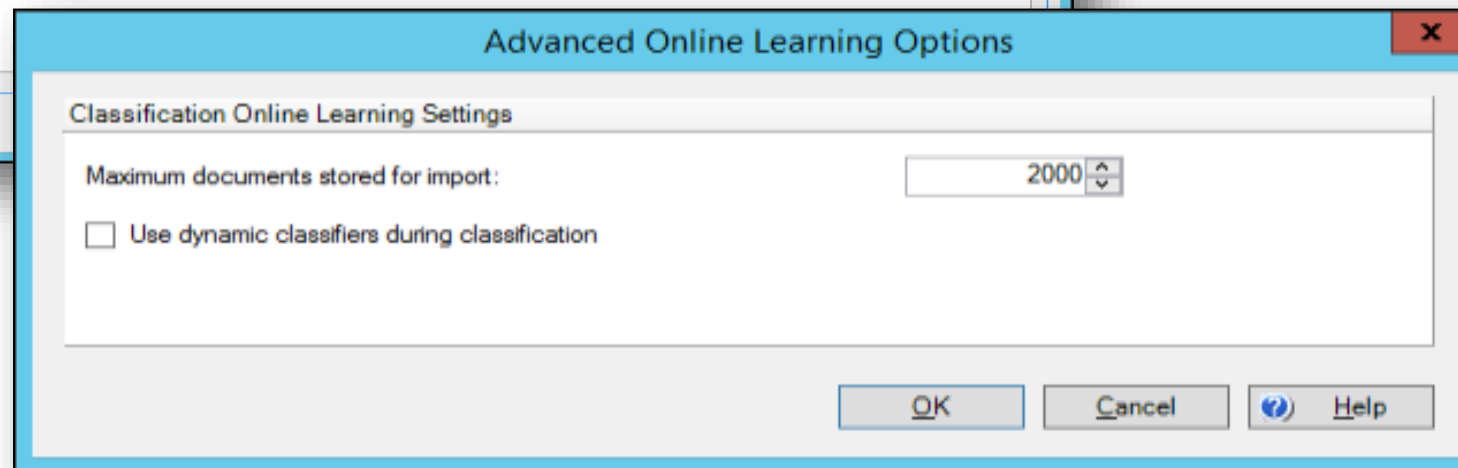 - An operator can also type the first few characters of a document type to filter the document type list (Auto Complete)

Classification Online Learning
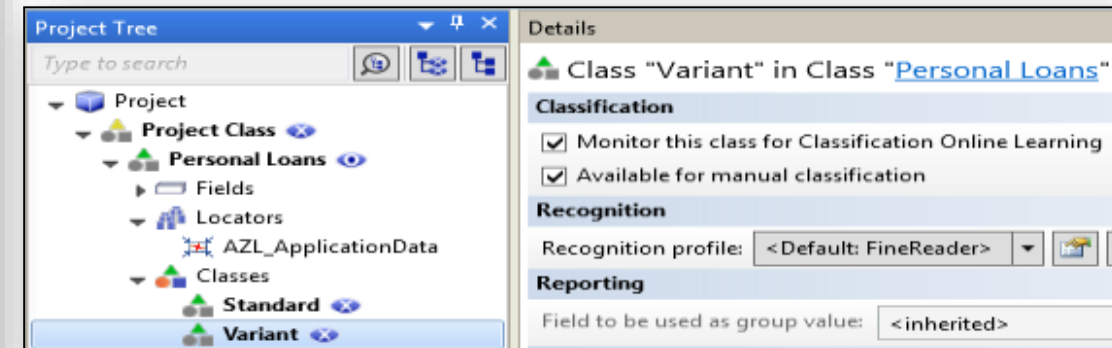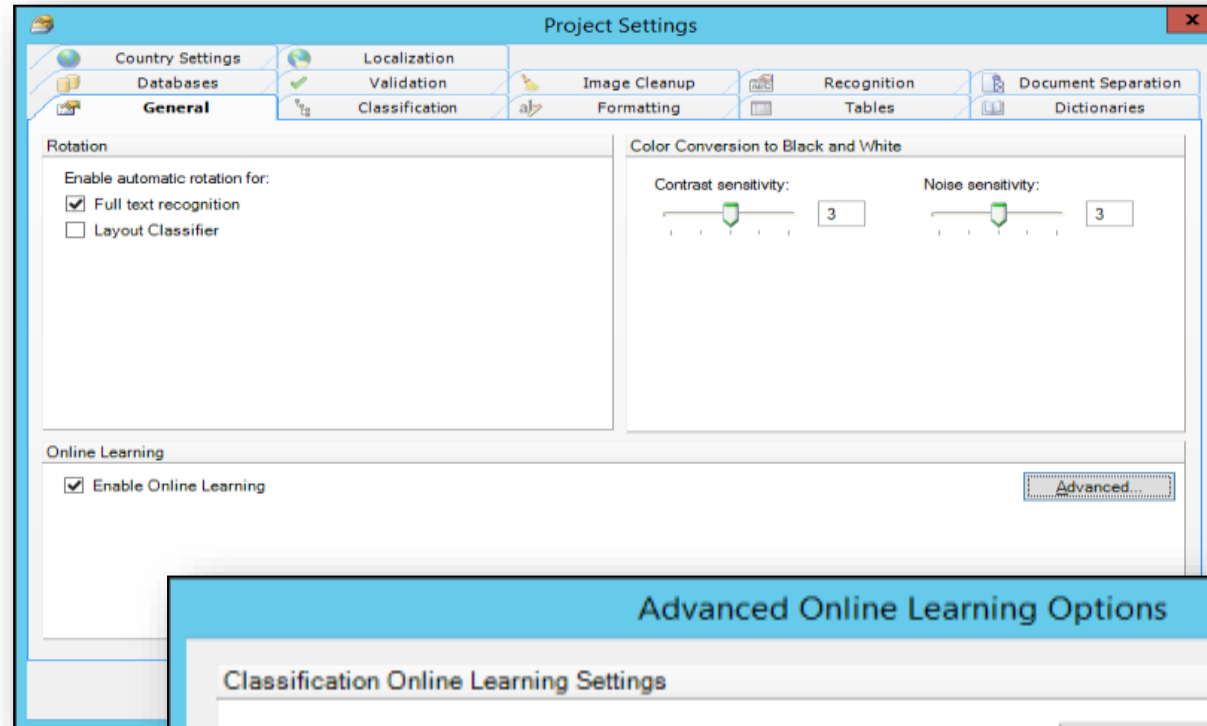
Dynamic Classifiers

- Automatically classify after new samples have been received
  - A maximum of 2000 (default) samples can be stored

- You can turn this option off if you wish to control which samples are used
  - Samples can be downloaded and deleted from the server
  - Downloaded samples can be added to the classification set

Advanced Online Learning Options

Classification Online Learning Settings

Maximum documents stored for import:　2000

☐ Use dynamic classifiers during classification

OK　Cancel　? Help

KOFAX