

Automated Document Classification

Golden Goose AI's Textual Understanding Framework (TUF) is an Intelligent Automation (IA) Natural Language Understanding (NLU) framework that uses Personas.

A Persona is a collection of custom no-code textual content identification criteria based on patterns and Natural Language Processing (NLP). Personas support conditional hierarchical identification criteria. Transactions are invoked on content found or not found.

Transactions are various sorts of activities, such as sending emails, creating a report, calling a 3rd party microservice, document classification, and others.

A custom no-code Persona can quickly be created for a specified purpose, a particular department/team, or a company.

This whitepaper is focused on Document Classification. Document Classification is needed in a variety of industries and verticals, where there is a need to sort, categorize, and analyze large volumes of documents.

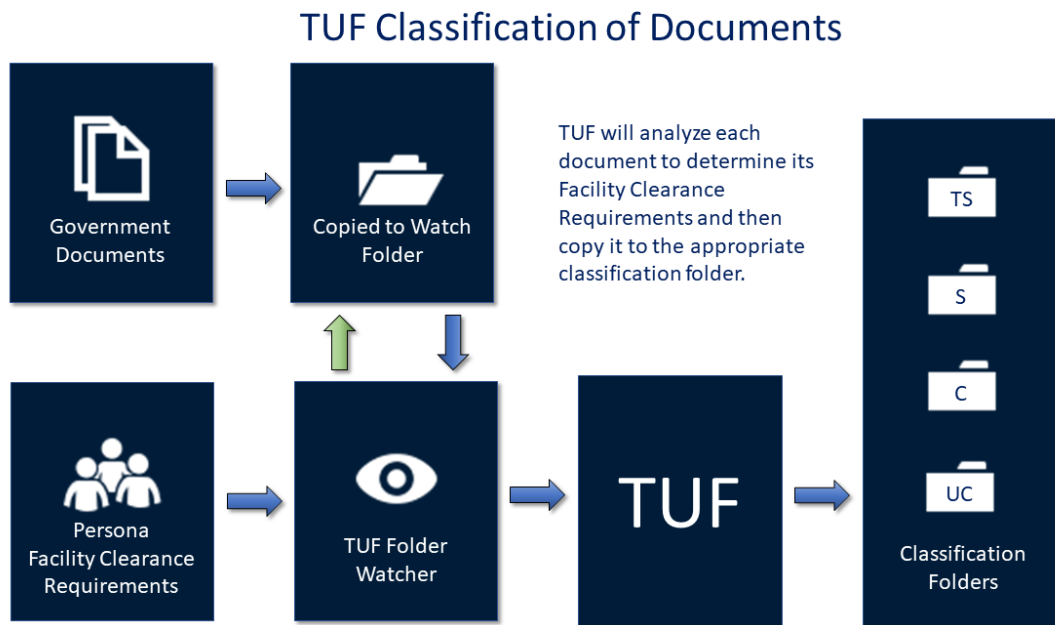
Some examples of verticals benefiting from Document Classification include:

- Legal
- Healthcare
- Financial
- Government
- Education
- Marketing

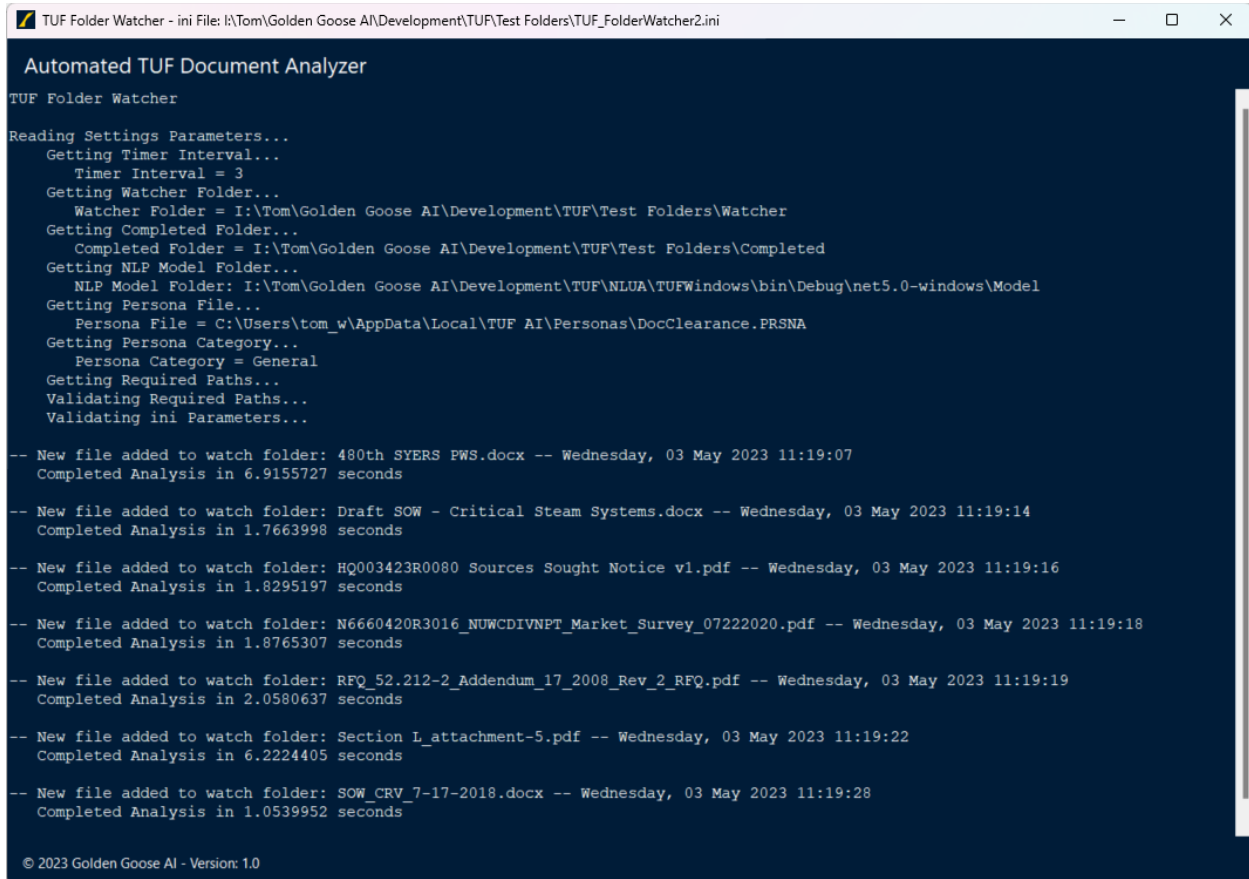
The Document Classification example we will be discussing is identifying Facility Clearance Requirements in Government documents, such as Request for Proposals (RFP). We will be using the TUF Folder Watcher to automate this process.

Process:

- TUF Folder Watcher looks for new document files in a specified Watch Folder.
- When a new document has been copied into the Watch Folder, the TUF Folder Watcher calls the TUF to analyze the document’s content and seeks its Facility Clearance Requirement, based on a Persona.
- Once the TUF identifies the document’s Facility Clearance Requirement, it moves the document to its associated classification folders, e.g., TS = Top Secret, S = Secret, C = Classified, and UC = Unclassified.



The below image is a screen shot of the TUF Folder Watcher running.



```
TUF Folder Watcher
Reading Settings Parameters...
Getting Timer Interval...
  Timer Interval = 3
Getting Watcher Folder...
  Watcher Folder = I:\Tom\Golden Goose AI\Development\TUF\Test Folders\Watcher
Getting Completed Folder...
  Completed Folder = I:\Tom\Golden Goose AI\Development\TUF\Test Folders\Completed
Getting NLP Model Folder...
  NLP Model Folder: I:\Tom\Golden Goose AI\Development\TUF\NLUA\TUFWindows\bin\Debug\net5.0-windows\Model
Getting Persona File...
  Persona File = C:\Users\tom_w\AppData\Local\TUF AI\Personas\DocClearance.PRSNA
Getting Persona Category...
  Persona Category = General
Getting Required Paths...
Validating Required Paths...
Validating ini Parameters...

-- New file added to watch folder: 480th SYERS PWS.docx -- Wednesday, 03 May 2023 11:19:07
Completed Analysis in 6.9155727 seconds

-- New file added to watch folder: Draft SOW - Critical Steam Systems.docx -- Wednesday, 03 May 2023 11:19:14
Completed Analysis in 1.7663998 seconds

-- New file added to watch folder: HQ003423R0080 Sources Sought Notice v1.pdf -- Wednesday, 03 May 2023 11:19:16
Completed Analysis in 1.8295197 seconds

-- New file added to watch folder: N6660420R3016_NUWCDIVNPT_Market_Survey_07222020.pdf -- Wednesday, 03 May 2023 11:19:18
Completed Analysis in 1.8765307 seconds

-- New file added to watch folder: RFQ_52.212-2_Addendum_17_2008_Rev_2_RFQ.pdf -- Wednesday, 03 May 2023 11:19:19
Completed Analysis in 2.0580637 seconds

-- New file added to watch folder: Section L_attachment-5.pdf -- Wednesday, 03 May 2023 11:19:22
Completed Analysis in 6.2224405 seconds

-- New file added to watch folder: SOW_CRV_7-17-2018.docx -- Wednesday, 03 May 2023 11:19:28
Completed Analysis in 1.0539952 seconds

© 2023 Golden Goose AI - Version: 1.0
```

Stats:



7 – Documents



71 – Pages



21.45 – Seconds

This entire automated document classified process was developed without writing code.

After the documents have been classified and automatically copied into their classification folder, additional analysis and transactions can occur. A Persona can be defined for each classification for specialized analysis and transactions.

TUF Transactions from Classified Folders

