USTD0H - Capture et traitement de l'information

Mis à jour le 13-02-2025

Code : USTD0H

Unité spécifique de type cours

6 crédits

Responsabilité nationale : EPN15 - Stratégies / Thomas GRIGNON

Présentation

Prérequis

Être inscrit au Master « Chef de projet en stratégie et gouvernance de l'information ».

Objectifs pédagogiques

- Acquisition de données :
  - Savoir choisir entre méthodes de capture manuelle (saisie assistée), semi-automatisée (OCR avec Abbyy FineReader) ou *full-automation* (RPA avec UiPath) ;
  - Intégrer des sources hétérogènes (PDF, emails, formulaires web) dans un système centralisé ;
  - Usage des outils documentaires de type GED.
- Transformation et valorisation :
  - Appliquer des techniques de nettoyage, enrichissement (ajout de métadonnées) et classification (taxonomies, thésaurus) ;
  - Mobiliser des outils d'analyse sémantique (Talend, Elasticsearch) pour structurer l'information

Programme

Contenu

- Technologies de capture :
  - Atelier OCR : paramétrage de zones de capture, gestion des erreurs (ex : reconnaissance de caractères manuscrits) ;
  - GED et outils de stockages documentaires.
- Traitement avancé :
  - Workflows BPMN (Business Process Model and Notation) : modélisation d'un processus de validation multicanal ;
  - IA et Machine Learning : utilisation de modèles NLP (*Natural Language Processing*) pour l'extraction d'entités nommées.
- Qualité et gouvernance :
  - Méthodologies DCAM (*Data/Document Management Capability Assessment Model*) pour évaluer la maturité des processus ;
  - Étude de cas : gestion des données sensibles (ex : anonymisation RGPD avec Talend).

Modalités de validation

- Contrôle continu

Description des modalités de validation

- Etude de cas en groupe
- QCM individuel