

# GESTION ELECTRONIQUE DES DOCUMENTS (GED)

---

**2023/2024**

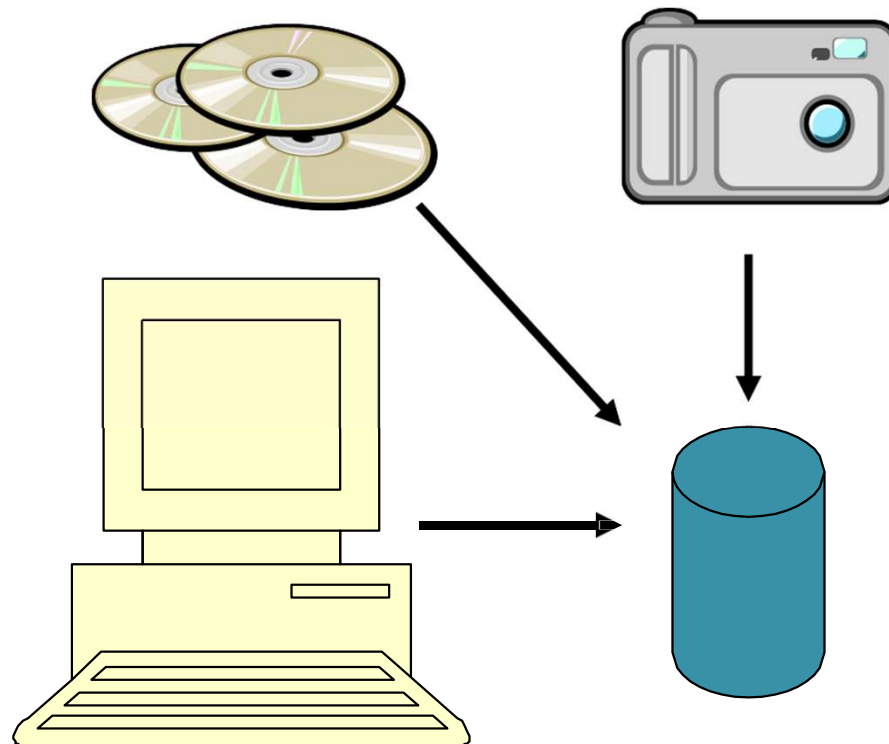
Anass MAMOUNY  
[amamouny@esi.ac.ma](mailto:amamouny@esi.ac.ma)

# GED : Fonctionnalités

- Intégration des documents nativement électroniques dans le système de GED.
- Acquisition des documents dans n'importe quel format.
- Organiser, indexer et relier l'information.
- Stocker les documents.
- Fournir un accès rapide à tout moment, n'importe où.
- Gérer le processus de modification du document, en veillant à l'intégrité de l'information.
- Communiquer les modifications à toutes les parties concernées.
- Réguler l'accès en fonction des autorisations.

# Acquisition des documents

Intégration des documents dans  
le système GED



# Acquisition des documents

Acquérir des documents :

- par numérisation des documents sur format papier
- en intégrant des formats électroniques existants (format bureautique, PDF...)
- par production de documents électroniques (intégration avec une application tierce : bureautique)
- par échange de documents électroniques avec des organismes extérieurs

# Acquisition des documents

- La lecture automatique de documents (**LAD**) est un ensemble de technologies qui permet de segmenter et d'extraire des informations textuelles sur des documents numérisés de type formulaires structurés ou semi-structurés.
- La reconnaissance automatique de documents (**RAD**) est une technique permettant de distinguer un type de document d'un autre à partir de l'image numérique du document.

# Acquisition des documents

- La reconnaissance optique de caractères (**ROC**), en anglais optical character recognition (OCR), ou océrisation, désigne les procédés informatiques pour la traduction d'images de textes imprimés ou dactylographiés en fichiers de texte.
- La reconnaissance intelligente de caractères (abrégée **ICR** pour Intelligent Character Recognition) est une technique avancée de reconnaissance optique de caractères (dit OCR). Comme l'OCR, l'ICR part d'une image contenant du texte pour en extraire le contenu textuel.

# Traitement des documents

Classer les documents :

- **L'indexation par type de document** : les métadonnées servent à indexer les documents (type, auteur, titre, source, date, etc.) Pour que ceux-ci puissent être recherchés.

- **L'indexation manuelle** : l'ajout de mots-clés permettent de faciliter les opérations de recherches.

- Versionning

# Traitement des documents

## **L'indexation par type de document**

1.une fois le document numérisé, un composant logiciel réalise le nommage des fichiers selon une norme unique, classement dans un plan, préparation pour le stockage dans une base de données de la GED.

**2.Récupération d'informations** : informations issues du système de gestion de fichiers. Certaines peuvent être pertinentes comme des métadonnées : le nom du fichier, les dates de création et de modification, sa taille, son emplacement.

**3.Extraction de données structurées** : lorsque ces structures sont connues et documentées, il est possible d'utiliser des automates pour extraire les informations pertinentes directement depuis le fichier et les utiliser pour alimenter une fiche d'indexation.

**4.Induction de métadonnées** : certaines solutions, les plus avancées, permettent de déterminer automatiquement des métadonnées pertinentes pour indexer un document.

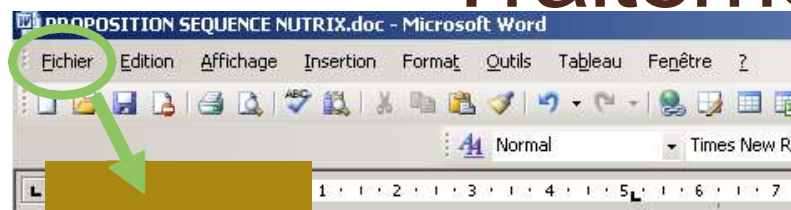


# Traitement des documents

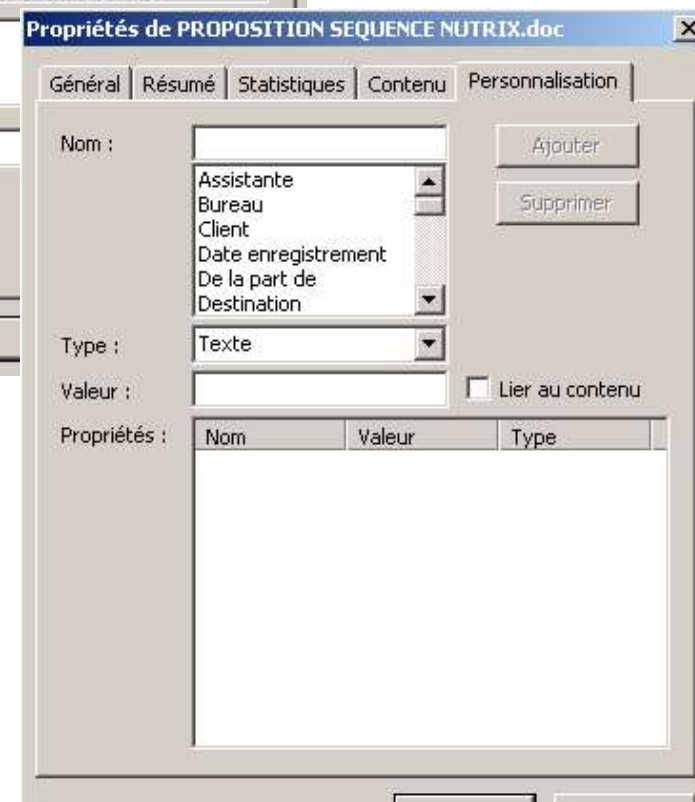
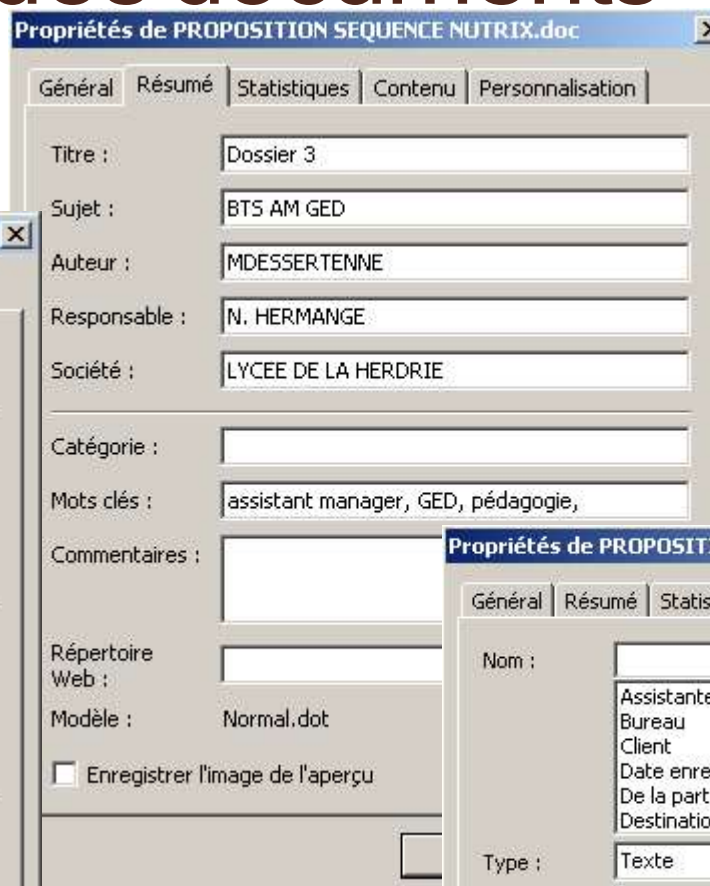
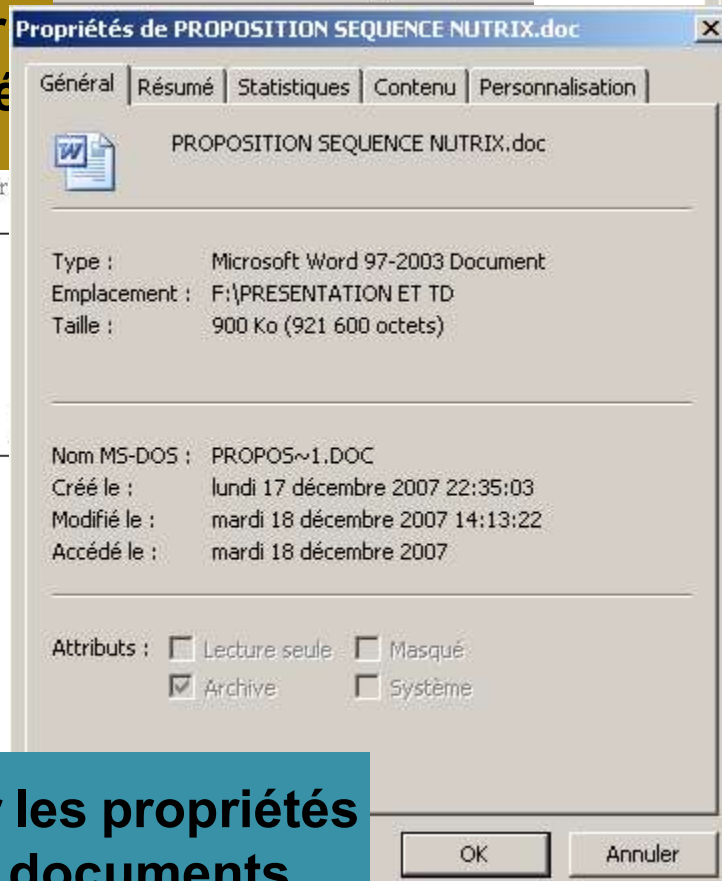
## Techniques (souvent combinées) :

- ❖ **Des outils de reconnaissance** : permettent de retrouver dans un document les chaînes de caractères présentes dans un référentiel pouvant être ajoutées à la fiche d'indexation.
- ❖ **Des outils statistiques** : analysent les chaînes de caractères les plus présentes et donc potentiellement les plus représentatives.
- ❖ **Des outils sémantiques** : capables d'extraire automatiquement les mots et expressions les plus pertinents, voire de reconnaître s'il s'agit de mot-clé, de date, du titre...

# Traitement des documents



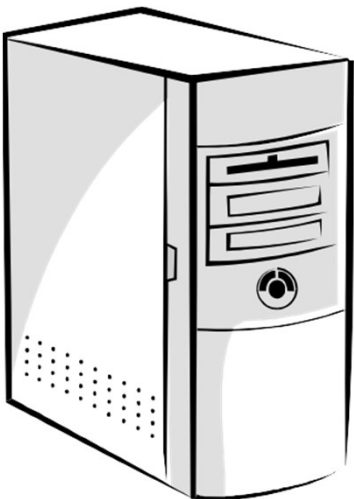
**Fichier  
Propriété**



**Définir les propriétés  
des documents  
bureautiques pour  
faciliter  
leur recherche**

# Stockage des documents

- Mettre en place une solution de stockage adaptée à ses besoins pour le bon fonctionnement des outils de gestion de documents.
- Pour le choix du support de stockage, Il faut tenir compte des éléments suivants :
  - volume de stockage.
  - la fréquence de consultation.
  - l'importance des données.
  - l'organisation du stockage.
  - la durée de conservation.



# Stockage des documents

Stocker des documents :

- en fonction du contenu des documents, de leurs provenances, états, types, etc.
- archivage des documents anciens
- sauvegarde et restauration en cas de destruction
- destruction programmée

# Diffusion des documents

Diffuser des documents :

via internet, un intranet, un client léger

- On distingue les deux modes suivant :
  - **Le mode pull** : la mise à disposition des documents selon les droits des utilisateurs qui consultent le référentiel de documents après connexion. Les documents sont visualisables après recherche et sélection.
  - **Le mode push**: la distribution et l'envoi des documents à des destinataires désignés préalablement.

Quels sont les composants  
matériels et logiciels nécessaires  
pour mettre en place un système  
GED ?

# Composants matériels

- Les Scanners de Production
- Les supports de stockage
- Les imprimantes

# Composants matériels

## Les Scanners de Production

- Sont des périphériques d'acquisition de documents en vue de leur transformation sous forme d'images numérisées.
- **Les critères principaux**
  - Le format de numérisation des documents : A4, A3, A2, ...
  - La problématique des volumes et des flux;
  - Vitesse de numérisation en nombre de pages/minute;
  - La qualité de l'image : résolution mesurée en dpi –  
L'entraînement du papier
  - La facilité d'exploitation et d'entretien
  - Le coût d'achat et d'entretien



# Composants matériels

## Les Scanners de Production

- **Les principaux constructeurs de scanners de production :** CANON, FUJITSU, KODAK, BELL & HOWELL, etc.
- **Les principaux protocoles d'interface logicielle ou norme de communication entre les applications et les scanners :** TWAIN, WIA, ISIS, Interface vidéo propriétaire.
- **Les Copieurs Numériques Multifonctions :** connectés dans les entreprises pouvant être utilisés comme imprimante et/ou comme scanner en réseau. Des exemples de marque : TOSHIBA, CANON, XEROX, RICOH, KYOCERA, KONICA.

# Composants matériels

## Les supports de stockage

- Un(des)serveur(s) de données.
- Un serveur d'applications, où le système sera installé.
- Des stations de travail.
- une passerelle liant le système aux logiciels métier de l'entreprise.

# Composants logiciels

- Logiciel de Numérisation
- Logiciel de reconnaissance : OCR, LAD, RAD, ...
- Logiciel de Gestion Électronique de Documents
- Antivirus pour la protection des menaces extérieures ou intérieures (virus, mauvaises manipulations, ...)

# Composants logiciels

## Logiciel de Numérisation

- Permet de numériser les documents page par page, par lot, en choisissant la résolution (ex 300 dpi) et le format de fichier (ex PDF, TIFF, JPEG, ...).
- Il peut être fourni avec le scanner ou acheté séparément.

# Composants logiciels

## Logiciel de reconnaissance

- Les principaux moteurs OCR sur le marché :
  - FineReader (Abbyy)
  - OmniPage Pro (ScanSoft)
  - ReadIRIS Pro (IRIS)
- Certains logiciels de reconnaissance de caractères comme ReadIris Pro génèrent un texte éditable tout en conservant la mise en page. Il recrée par exemple les tableaux et colonnes du document original dans Word.
- **Préalable requis** : qualité de l'image;

# Les solutions de GED

- Pour les Administrations, Collectivités, Grands Groupes : Docubase, Gedox, DocuShare, CanonFile, ZeDOC, ZyLabs, etc.
- Pour les PME-PMI, TPE : Docubase, Gedox, Novaxel, ArchiDoc,  
AutoIndexIt, Documind Pro, Documentum,

# Les solutions de GED

