

# Automatisation des processus dans IJCan

par

François HARVEY

(Programmeur-analyste)

# Automatisation du processus de publication des décisions sur IJCan

- Problématique
- Processus de publication
- Conclusion

# Problématique

- Publication des décisions de 65 tribunaux
- Traitement de plus de 1500 décisions par semaine
- Bilinguisme
- Équipe éditoriale de 3 personnes

# Processus de publication

- Cinq étapes
  - Acquisition des documents
  - Extraction et validation des méta-données
  - Insertion dans le système et conversion
  - Génération des pages Web
  - Vérification post-publication

# Acquisition des documents

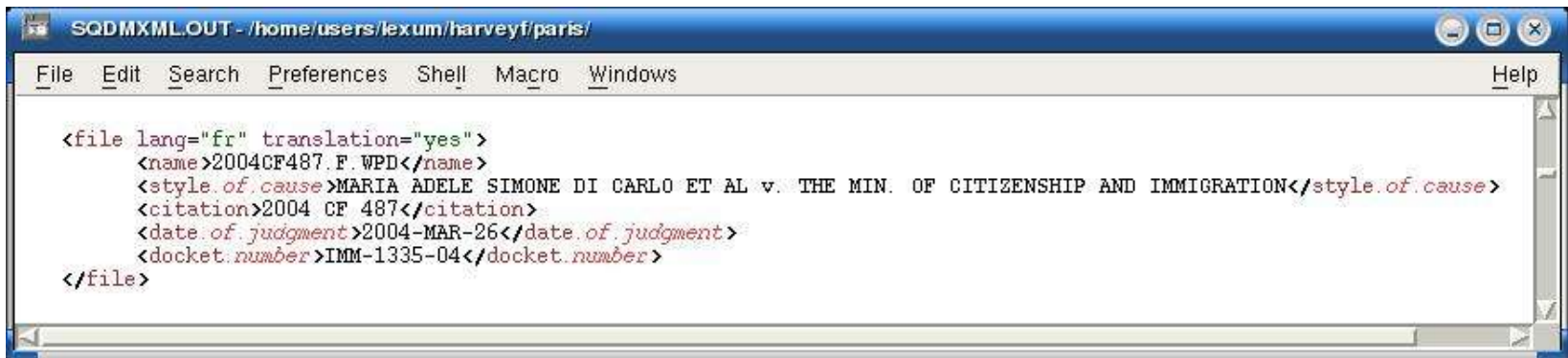
- Dépend de la source des documents
  - Courriel
  - Téléchargement (HTTP ou FTP)
  - Autres supports (CD, disquettes, ...)

# Extraction et validation des méta-données

- Automatiquement
- Manuellement

# Extraction automatique

- Informations fournies sous un format normalisé (XML, SGML, ...)



The screenshot shows a text editor window titled "SQDMXML.OUT - /home/users/lexum/harveyf/parts/". The window contains the following XML code:

```
<file lang="fr" translation="yes">
  <name>2004CF487.F.WPD</name>
  <style.of.cause>MARIA ADELE SIMONE DI CARLO ET AL v. THE MIN. OF CITIZENSHIP AND IMMIGRATION</style.of.cause>
  <citation>2004 CF 487</citation>
  <date.of.judgment>2004-MAR-26</date.of.judgment>
  <docket.number>IMM-1335-04</docket.number>
</file>
```

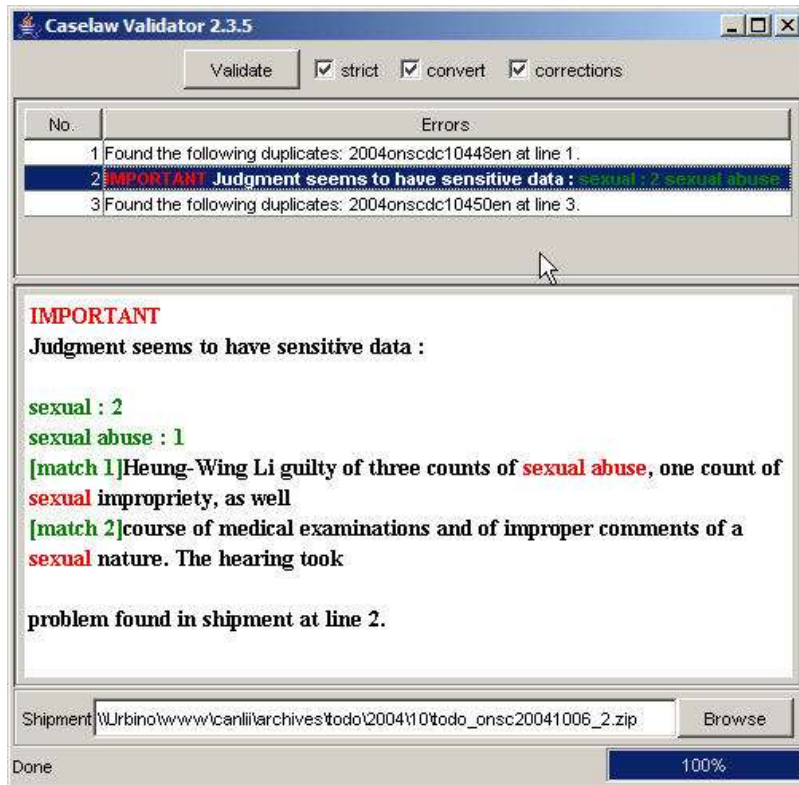
- Uniformité des documents permettant l'utilisation d'expressions régulières

# Extraction manuelle

- Opération éditoriale d'extraction des méta-données du document impliquant :
  - Ouverture du document
  - Création d'un fichier regroupant les méta-données
- Solution non optimale
  - Processus plus long
  - Génère une possibilité d'erreurs



# Validation



- Méta-données validées à l'aide d'un programme
- Tests
  - Vie privée
  - Doublons
  - Et autres

# Insertion dans le système et conversion

- Validation de l'intégrité des informations pour permettre leur insertion
- Insertion dans la base de données
  - Documents
  - Méta-données
- Conversion des documents au format HTML
- Insertion d'hyperliens dans le texte des décisions

# Génération des pages Web

- Utilisation de Coldfusion pour la création du site statique :
  - Application de la signature visuelle d'IIJCan aux documents HTML déjà convertis
    - Ajout des entêtes et pieds de page
  - Génération des pages de navigation d'IIJCan
  - Sauvegarde de la page dynamique générée pour en faire une page statique

# Vérification post-publication

- Vérification manuelle de chacun des documents publiés, assistée de 23 tests automatisés
  - Apparence
  - Contenu
  - Méta-données
  - Informations sensibles

# Vérification post-publication

The screenshot displays the Mozilla Firefox browser window at the URL `http://localhost:8081/canlii-editor/select.do`. The interface is divided into several sections:

- Sidebar:** A list of HTML files, including `6 - 1998abca141.html` through `14 - 1998abca173.html`.
- Main Table:** A table with columns: Info, Editer, Historique, Documents, Tests, Flags, and Afficher le contenu. The 'Tests' column contains a 'Rapport du Canlii-verificator' with a status of '2 / 2'. Other rows in the table show various document checks with statuses like 'Absent' or 'à Vérifier'.
- Test Details:** A section below the table showing details for a specific test:
  - Nom: 9 - Vie Privée
  - Date de l'identification: 2004-10-08 19:11:17.439123
  - Statut: Échoué
  - Date du dernier test: 23 septembre 2004
  - Description: Tente d'identifier les décisions qui pourraient nécessiter un caviardage à l'aide de mots clefs.
  - État du test: Ce test est à jour
- Test Results:** A text area showing the results of the test, including matches for 'sexual assault' and 'young offender' in a legal document snippet.

# Conclusion

- Notre procédure est complète, mais
  - plusieurs étapes manuelles demeurent
  - certaines de nos collections ne passent pas par l'ensemble de ce processus
- Nous voulons donc uniformiser le traitement plus tôt dans le processus :
  - Étendre notre processus à l'ensemble de nos collections
  - Devancer l'étape de l'insertion des documents afin de minimiser les manipulations
  - Promouvoir des normes d'uniformisation (référence neutre, préparation des jugements)