AIDAR:
Adressage et Indexation de Documents Multimédias Assistés par des techniques de Reconnaissance Vocale

Acteurs :

- VOICE-INSIGHT SA Frédéric Beaugendre : frederic.beaugendre@voice-insight.com

- ULB Gianluca Bontempi : gbonte@ulb.ac.be

- Asbl Titan: Guy Maréchal : gmarechal@brutele.be

12 mars 2025

Objectifs du projet

- Indexation thématique de flux sonore « brut » pour des données radiophoniques
- Traitement sans préparation préalable des données
  - Implique l'utilisation de techniques statistiques, travaillant sur de grandes quantités de données
  - Utilisation de données étiquetées en grande quantité pour la phase d'apprentissage
- Exploitations possibles
  - Indexation de potentiellement tout type de documents multimédias audio, extension possible à d'autres langues
  - Radio/Television : indexation (semi-)automatique des archives, aide à la reprise d'antériorité

Avantage de la technologie pour aider à la reprise d'antériorité

- Minimisation des coûts **d'indexation :**

  une indexation manuelle suppose en effet l'écoute intégrale des documents afin d'y attacher une série d'index prédéfinis

- **Méthodologie d'indexation** cohérente**,** unifiée **et** reproductible

… pour la masse sans cesse croissante de documents disponibles

Planing et Participants

Planing : 24 mois (Janvier 2005 à Janvier 2007)

Partenaires

- Voice-Insight SA : coordinateur du projet, responsable des tâches de reconnaissance de parole, interface utilisateur
- ULB (Machine Learning Group) : classification et segmentation textuelle
- TITAN : architecture, dissémination

Collaborations

- LIA (laboratoire d'informatique d'Avignon, France) : *Speeral*, moteur de reconnaissance vocale en français
- RTBF : définitions des besoins utilisateurs, données multimédias
- Skema : développement de l'AXIS manager

Problématique générale de l'indexation automatique de documents audionumériques

- Segmentation du corpus en segments « parole / musique / bruit »
- Identification de la langue
- Délimitation des frontières de parole pour un locuteur
- Reconnaissance du locuteur
- Etiquetage de l'ensemble du texte
- Reconnaissance de mots-clefs au sein du texte transcrit
- Reconnaissance thématique :
  - Segmentation par thème : segmentation du corpus en sous corpus traitant chacun d'un thème donné
  - Détection de thème : détection automatique d'un nouveau thème
  - Etiquetage de thème : association d'un thème donné avec un (une série d') index identifié(s)

Outils de traitement audio/parole utilisés : Speeral (LIA)

- Système de reconnaissance grand vocabulaire
  - Segmentation -> indications de ruptures
    - Du flux acoustique (parole, musique, bruit, reconnaissance de jingles)
    - En locuteur
  - Reconnaissance et suivi du locuteur
  - Reconnaissance du type de parole (téléphonie, office, …)

    ➔sélection des modèles acoustiques
  - Transcription textuelle des segments de parole
  - Alignement automatique de partie retranscrites préalablement
  - Adaptation au locuteur
  - Activation de grammaires dynamiques relatives au contexte

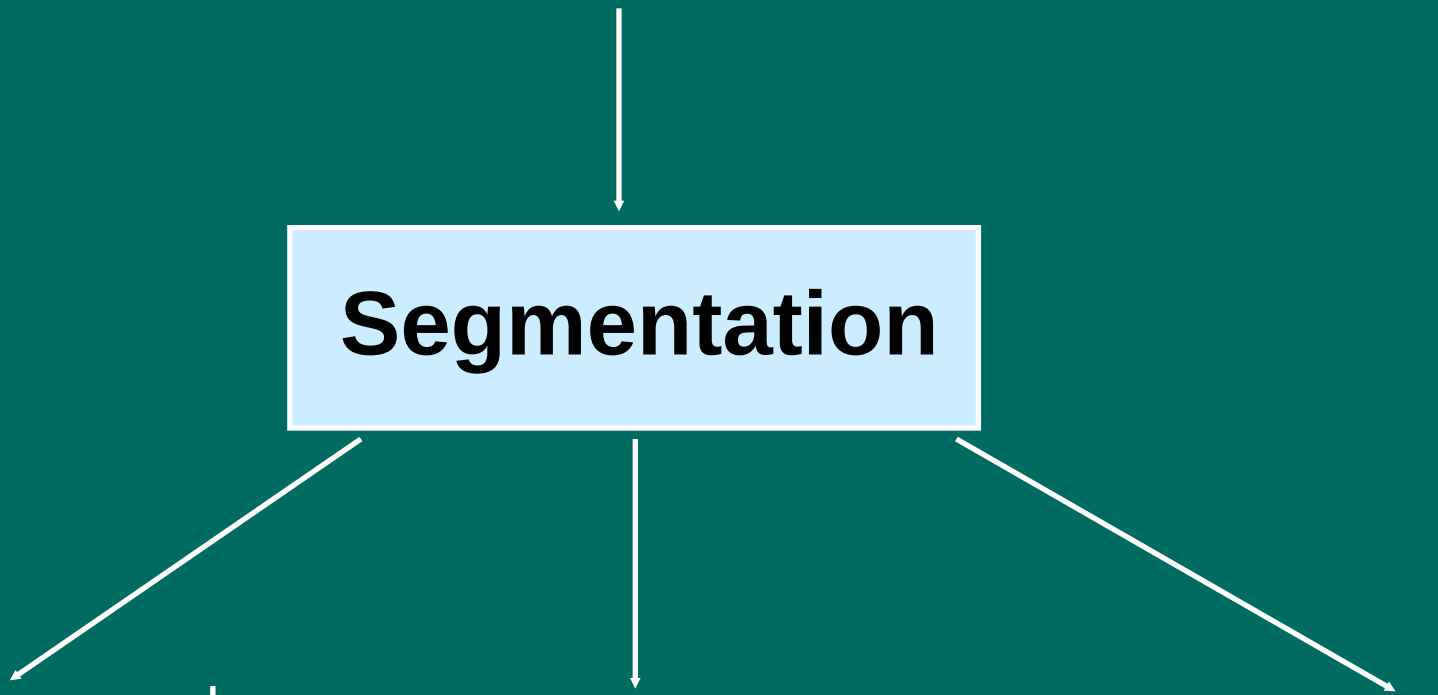
Technique d'analyse textuelle

- Entrée: transcription textuelle faite par le moteur de reconnaissance vocale
- Processus d'analyse textuelle pour extraire l'information pour la créations des metadonnées
- Utilisation de deux méthodes:
  - Segmentation linéaire (1ère passe) : Permet de segmenter un texte en parties sémantiquement disjointes
  - Classification supervisée (2ème passe) : Classer un texte ou une portion de texte selon un ensemble de thèmes prédéfinis

Processus de segmentation

(Transcription textuelle)

Phrase 1 Phrase 2 Phrase 3 Phrase 4 Phrase 5 … Phrase i … Phrase N



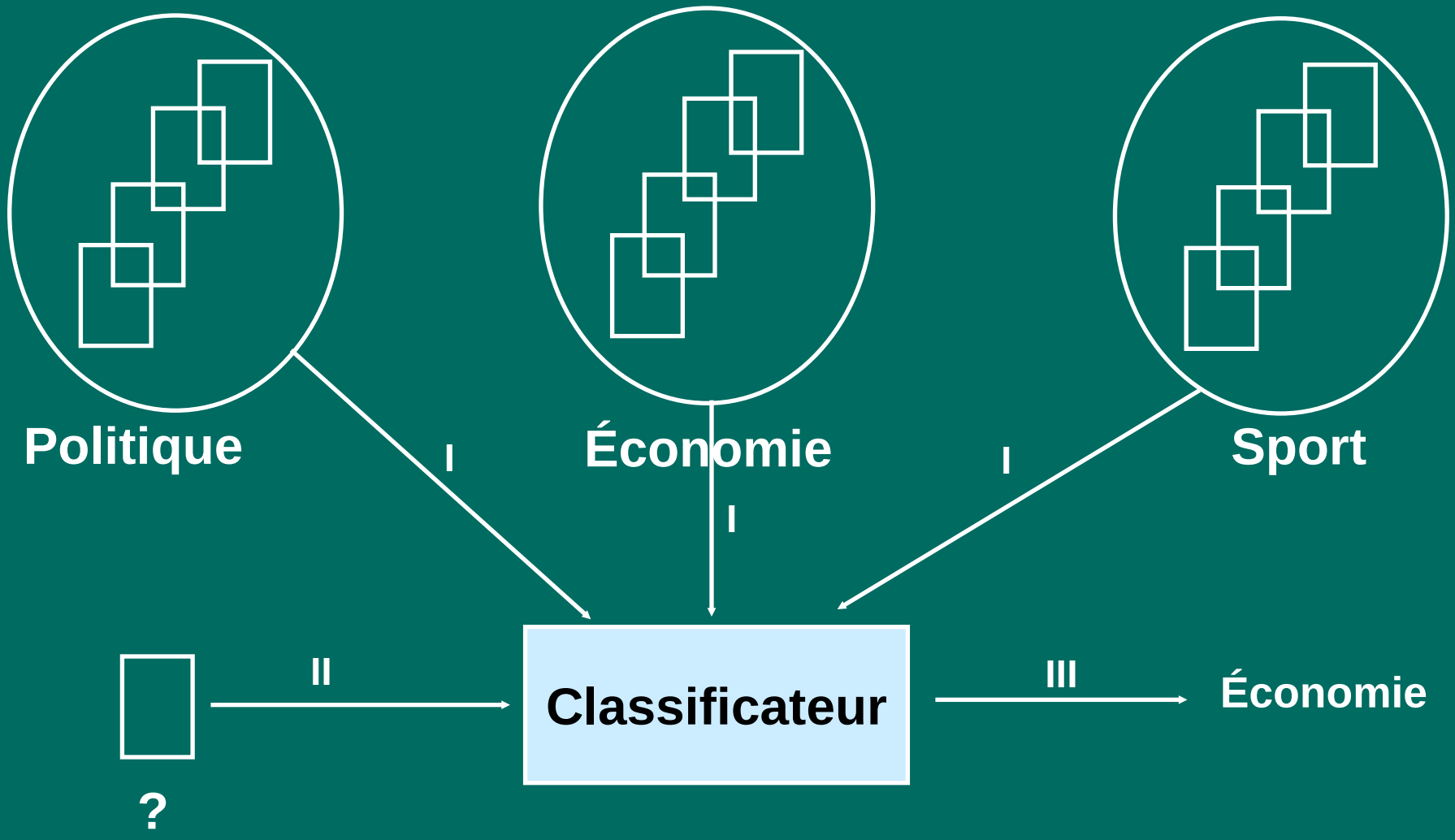
Phrase 1 Phrase 2 | Phrase 3 Phrase 4 Phrase 5 | … Phrase i … Phrase N

(Document 1) (Document 2) (Document 3)

Classification textuelle

- Classer un texte ou une portion de texte selon un ensemble de thèmes prédéfinis (Politique, Sport, Économie, etc.)
- Création d'un algorithme basé sur un apprentissage supervisé
- Besoin d'un corpus de documents déjà étiquetté pour réaliser l'apprentissage
- Utilisation d'un thesaurus thématique

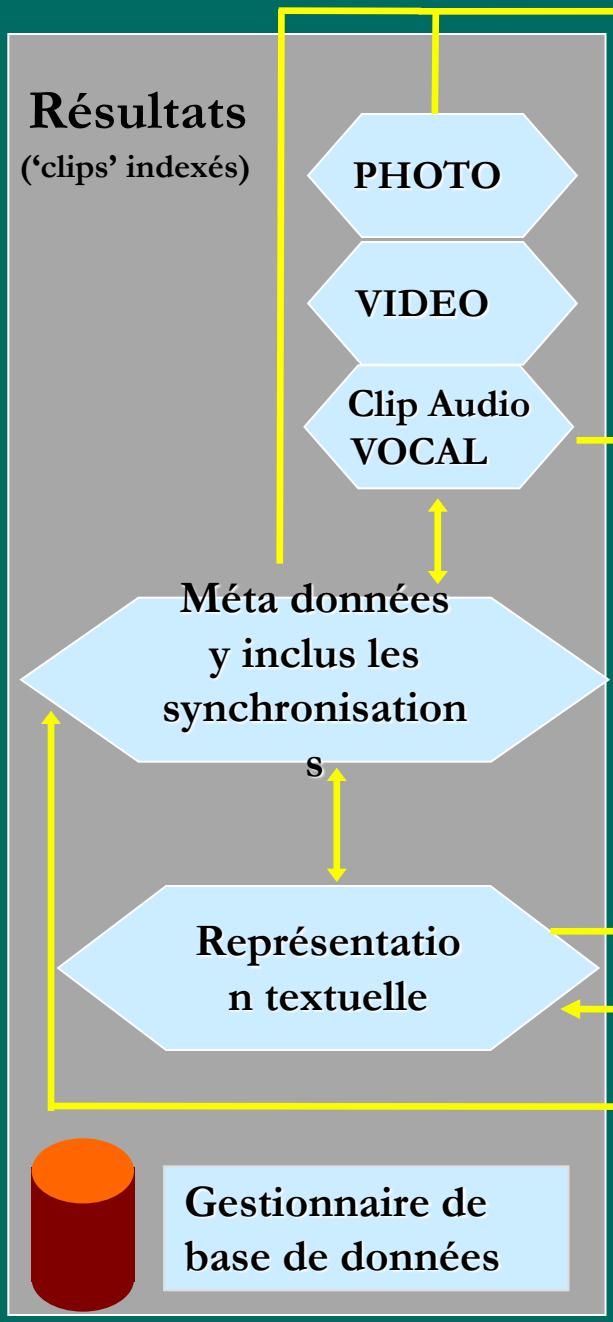
Processus de classification

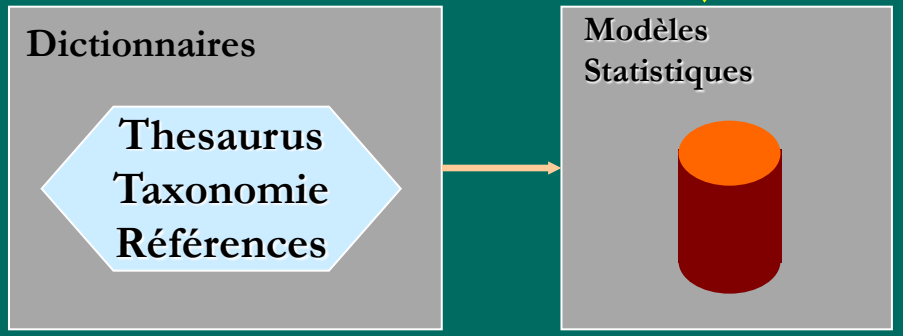


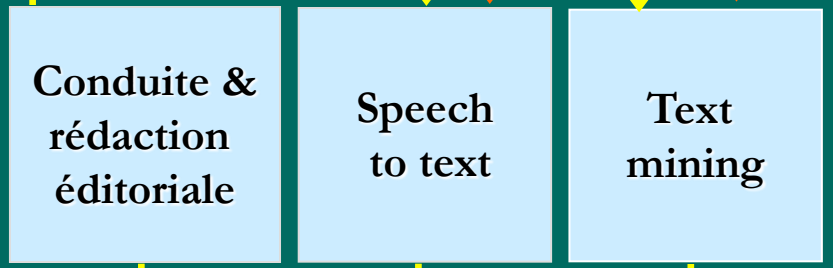
Modèles de langages

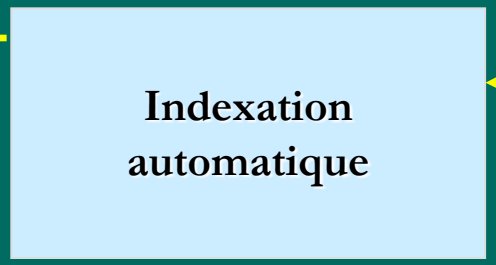
- Création de modèles de langages suivant une structure hiérarchique:
  - Modèles généraux, spécifiques à un domaine particulier
  - Dépendants des modèles et thesauraus utilisés pour la segmentation et classification
- Objectif
  - Permet une meilleure reconnaissance et donc une meilleure transcription textuelle
- Méthodologie : reconnaissance multi-passes
  - Première passe avec un modèle générique
  - Deuxième passe avec des modèles spécifiques au domaine ou focalisés sur une période particulière

Projet AIDAR (La diversité des flux d'information)









Démonstration au stand Voice-Insight