



**HAL**  
open science

## TexMix - Navigation dans des archives de journaux télévisés

Morgan Brehinier

► **To cite this version:**

Morgan Brehinier. TexMix - Navigation dans des archives de journaux télévisés. RFIA - Reconnaissance des Formes et Intelligence Artificielle, Jan 2012, Lyon, France. pp.978-2-9539515-2-3. hal-00660955

**HAL Id: hal-00660955**

**<https://hal.science/hal-00660955>**

Submitted on 19 Jan 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# TexMix – Navigation dans des archives de journaux télévisés

Morgan BREHINIER

INRIA Rennes – Equipe TexMex

INRIA – 263 avenue du Général Leclerc – 35042 RENNES  
morgan.brehinier@inria.fr

## Description

Cette démonstration a pour objectifs de naviguer dans une collection de journaux télévisés de manière innovante. Pour cela, nous partons d'une transcription automatiquement générée de la partie audio du journal télévisé afin de détecter les ruptures lexicales, et redécouper ainsi le JT en une série de reportages. Cette transcription nous permet aussi d'extraire des mots-clés afin de caractériser chaque reportage, et de proposer, en rapport à chaque reportage, des liens Internet pertinents et des reportages traitant de sujet similaire. A l'aide de détection d'entités nommées, chaque journal télévisé est enrichi de manière dynamique par des informations sur les personnes ou les locations citées. D'autre part, nous utilisons aussi une technologie de recherche d'images similaires afin de chercher les images similaires parmi toute notre collection de JT.

## Mots clés

HTML5, hypervideo, chapitrage automatique, entités nommées, images similaires, résumé dynamique.

## Bibliographie

<http://texmix.irisa.fr>

