

# Yes, I'm able to index audio files within Alfresco

2013

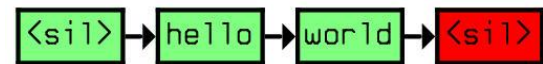
Fernando González

[fernando.gonzalez@ricoh.es](mailto:fernando.gonzalez@ricoh.es)

@fegorama

# Why?

- A lot of audio/video files in many companies
- The need to seek words in audio files
- Transcription of important conversations
- Efficiency in DAM



[fernando.gonzalez@ricoh.es](mailto:fernando.gonzalez@ricoh.es)  
@fegorama

#SummitNow

# What is it?



## AAT (Alfresco Audio Transcriber)

Alfresco Action (Java) for audio transcription  
with Sphinx-4 from Carnegie Mellon  
University

[fernando.gonzalez@ricoh.es](mailto:fernando.gonzalez@ricoh.es)  
@fegorama

*#SummitNow*

# What is Sphinx-4?



A group of speech recognition systems developed at Carnegie Mellon University.

These include a series of speech recognizers (Sphinx 2 - 4) and an acoustic model trainer (SphinxTrain).

[fernando.gonzalez@ricoh.es](mailto:fernando.gonzalez@ricoh.es)  
@fegorama

#SummitNow

# Elements of Sphinx-4



Language model:  
Grammars  
Dictionaries

Acoustic models:  
Hidden Markov Model (HMM)

[fernando.gonzalez@ricoh.es](mailto:fernando.gonzalez@ricoh.es)  
@fegorama

#SummitNow

# How does the action work?

The action...

- Transcribes by direct execution
- Transcribes using content rules
- Transcribes using UI-Actions
- Transcribes with Alfresco Scheduler



[fernando.gonzalez@ricoh.es](mailto:fernando.gonzalez@ricoh.es)  
@fegorama

#SummitNow

# Features



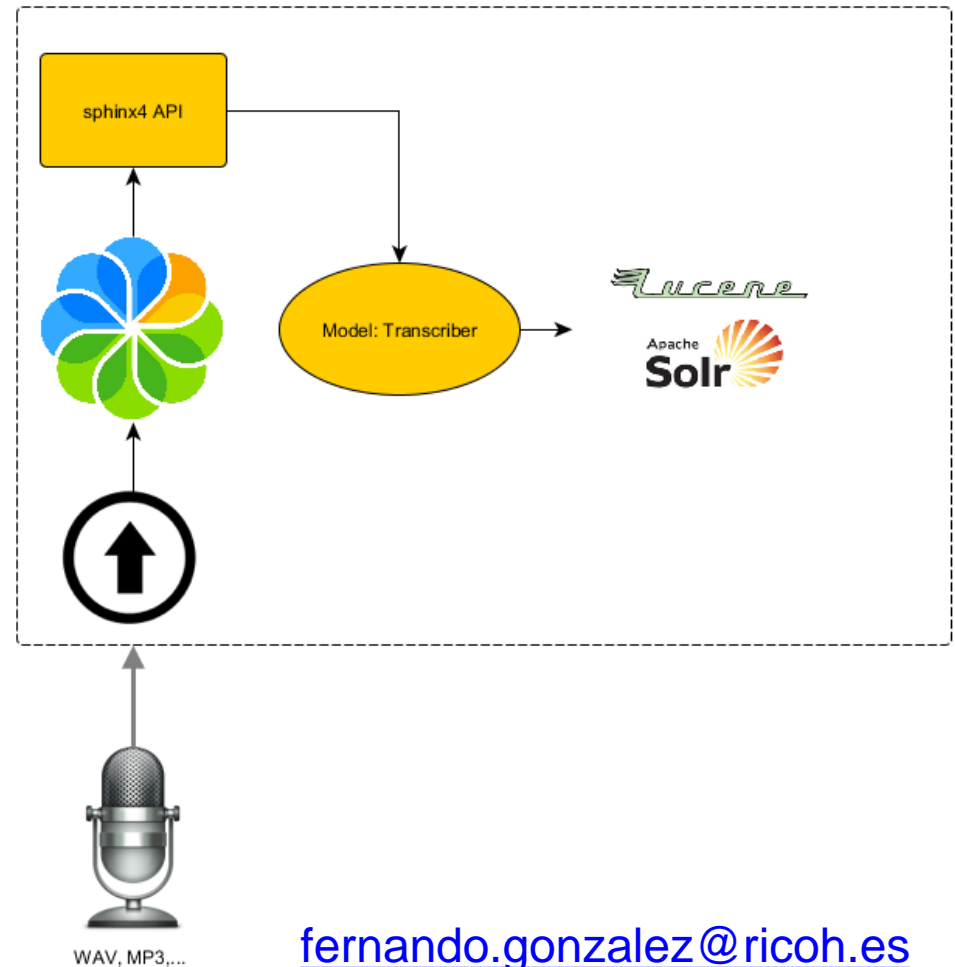
- Use of Sphinx-4 and JSAPI2 for recognition
- Use of "policies" to transcribe uploaded content
- Use of "scheduler" to transcribe spaces programmatically
- Use of action "Audio Transcriber" in user interfaces (Alfresco Explorer and Share)
- List of available Audio Files
- Assignment of "aspects" to control transcriptions

[fernando.gonzalez@ricoh.es](mailto:fernando.gonzalez@ricoh.es)  
@fegorama

#SummitNow

# Architecture

- Alfresco API (Actions)
- Share API (UI-Actions)
- JSAPI2
- Sphinx-4 API



[fernando.gonzalez@ricoh.es](mailto:fernando.gonzalez@ricoh.es)  
@fegorama

#SummitNow



# Transcriber Action

- Upload the file (WAV,...)
- **Run the Action**
  - Call to transcriber and recognizer
  - Capture words and other properties
  - Indexing...

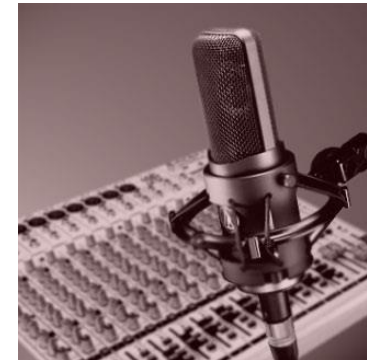
☒ Si se cumplen todos los criterios:

Tipo MIME en cm:content  es

☐ A menos que se cumplan todos los criterios:

**Realizar acción**

aat.transcriber.action



#SummitNow

# Model for audio-indexing

Aspect: Transcriber

Property: Words  
Index: Atomic and  
Tokenized

Property: Frames  
Index: No

Words and Frames are multiple

Navegador de nodos: cm:10001-90210-01803.wav

## Acerca de

**Referencia:** workspace://SpacesStore/2ee34b11-7a38-4674-96bc-d71dfd61b2a  
**Ruta primaria:** /app:company\_home/sts/sites/cm:aat-test/cm:documentLibrary/cm:Transcriber/cm\_x0031\_0001-90210-01803.wav  
**Tipo:** cm:content  
**Padre:** workspace://SpacesStore/f42b33ac-178d-41e8-8972-a64208e354ff

## Propiedades

| Nombre               | Tipo       | Valor  | Residual |
|----------------------|------------|--|----------|
| cm:name              | d:text     | 10001-90210-01803.wav  | false    |
| sys:node-dbid        | d:long     | 11957  | false    |
| sys:store-identifier | d:text     | SpacesStore  | false    |
| sys:locale           | d:locale   | es_ES  | false    |
| cm:content           | d:content  | contentUrl=store://2013/10/11/21/37/357c9561-0560-43f7-87d1-fe1b4e90295c.bin  mimetype=audio/x-wav  size=265914  encoding=UTF-8  locale=es_ES_  id=445 | false    |
| audio:sampleRate     | d:int      | 16000  | false    |
| sys:node-uuid        | d:text     | 2ee34b11-7a38-4674-96bc-d71dfd61b2a  | false    |
| cm:modified          | d:datetime | 11 oct 2013 21:37:54 GMT+0200 (CEST)   | false    |
| audio:sampleType     | d:text     | 16int  | false    |
| cm:created           | d:datetime | 11 oct 2013 21:37:53 GMT+0200 (CEST)   | false    |
| sys:store-protocol   | d:text     | workspace  | false    |
| cm:creator           | d:text     | admin  | false    |
| aat:Words            | d:text     | Colección<br>one zero zero zero one<br>nine oh two one oh<br>zero one eight zero three   | false    |
| cm:modifier          | d:text     | admin  | false    |

#SummitNow

# Ways to transcribe

- **Automatic transcription**
  - Upload/Create and Load documents
  - Actions/Rules
- **Programming transcription**
  - Scheduled Actions
- **Interactive transcription**
  - Repository action running
  - UI Action running



#SummitNow

# Fields of application

**DAM (Digital Asset Management)**

**Trials recording**

**Movies and Songs**

**Radio and TV**

**Education**



*#SummitNow*

# To Do...

New formats of audio files for transcriptions

Internationalization (Grammars and Acoustic models)

Specialized Dictionaries

Refactoring, refactoring and refactoring...

*#SummitNow*

PUT YOUR  
CONTENT  
TO WORK

