

# Choral +

**Spraaktechnologie: ingezet  
voor de ontsluiting van audio-  
visuele archieven**

# De benadering



**Doel van het NWO CATCH project CHoral (2006-2011):  
onderzoek en ontwikkel geautomatiseerde annotatie-  
en zoek technologie voor audio archieven**

**Doel van het NWO Catch+ project CHoral (2009-2012):  
Ontwikkel een generieke tool op basis van het  
voorgaande project, die breed inzetbaar is voor het  
werkveld**

**Samenwerking tussen:**

- 1. Spraaktechnologie onderzoekers, University of Twente**
- 2. Archivarissen, Gemeentearchief Rotterdam**
- 3. X-MI**



## 'Radio Rijnmond' (RR) archieven:

- De regionale omroep van Rotterdam Rijnmond
- Eerste uitzending in 1983
- Opnames van uitzendingen omvatten meer dan 60.000 uur
- Deels gedigitaliseerd, maar grotendeels analoog
- Deels ontsloten, maar voor het grootste deel wachtend op ontsluiting
- Typische collectie voor audio archieven in de sector cultureel erfgoed



# Doorzoeken van de RR archieven

<p><b>Collectie:</b> Bewegend beeld en geluid <b>Titel:</b> Radio Rijnmond <b>Beschrijving:</b> 16 december 1986. CD 2166. RR Getapt. Muziek. Kinderdans. Kerstcircus. Amnesty International. Geen plaats voor circussen in Vlaardingen. <b>Vervaardiger:</b> <b>Datering:</b> 16/12/1986 t/m 16/12/1986 <b>Techniek:</b> : cd, audio <b>Auteursrechthouder:</b> RTV Rijnmond ; <b>Beschreven persoon:</b> <b>Trefwoorden:</b> <b>Catalogusnr.:</b> RR 2166 <b>Reproductie:</b> luisterkopie RR 2166; audio-cd / bewaarkopie RR 2166; <b>Annotatie:</b></p>
<p><b>Collectie:</b> Bewegend beeld en geluid <b>Titel:</b> Radio Rijnmond <b>Beschrijving:</b> 2 januari 1987. CD 2221. Bontebal Excelsior start op. 10 Jaar Vrouwenhuis. Column Laurens. Wethouder Henderson over sociaal onderzoekers. Filmproject Dick Rijnke. Banketbakker C.Veth, van winkeliersvereniging Winkelcentrum De Korf, Krimpen a/d IJssel. <b>Vervaardiger:</b> Beranova, Jana / redactrice/ontwerper ; <b>Datering:</b> 02/01/1987 t/m 02/01/1987 <b>Techniek:</b> : cd, audio <b>Auteursrechthouder:</b> RTV Rijnmond ; <b>Beschreven persoon:</b> Henderson, Johan ( Henderson, J.H.A., ); Rijnke, Dick ( ); <b>Trefwoorden:</b> Sociale recherche (nieuw trefwoord) ; Vrouwen <b>Catalogusnr.:</b> RR 2221 <b>Reproductie:</b> luisterkopie RR 2221; audio-cd / bewaarkopie RR 2221; <b>Annotatie:</b></p>

Minimale metadata set per uur data



# Problemen

**De belangrijkste problemen waar de ontsluiting van deze collectie op stuit:**

- 1. een grote hoeveelheid niet ontsloten materiaal  
→ data niet toegankelijk voor derden**
- 2. Tamelijk algemene metadata, als die al aanwezig is  
→ slechts beperkt mogelijk om aan de vraag naar informatie te voldoen**
- 3. de audio bevindt zich op analoge dragers en CD's  
→ interactieve en online raadpleging is vrijwel niet mogelijk**



# Ontsluiting met behulp van spraaktechnologie

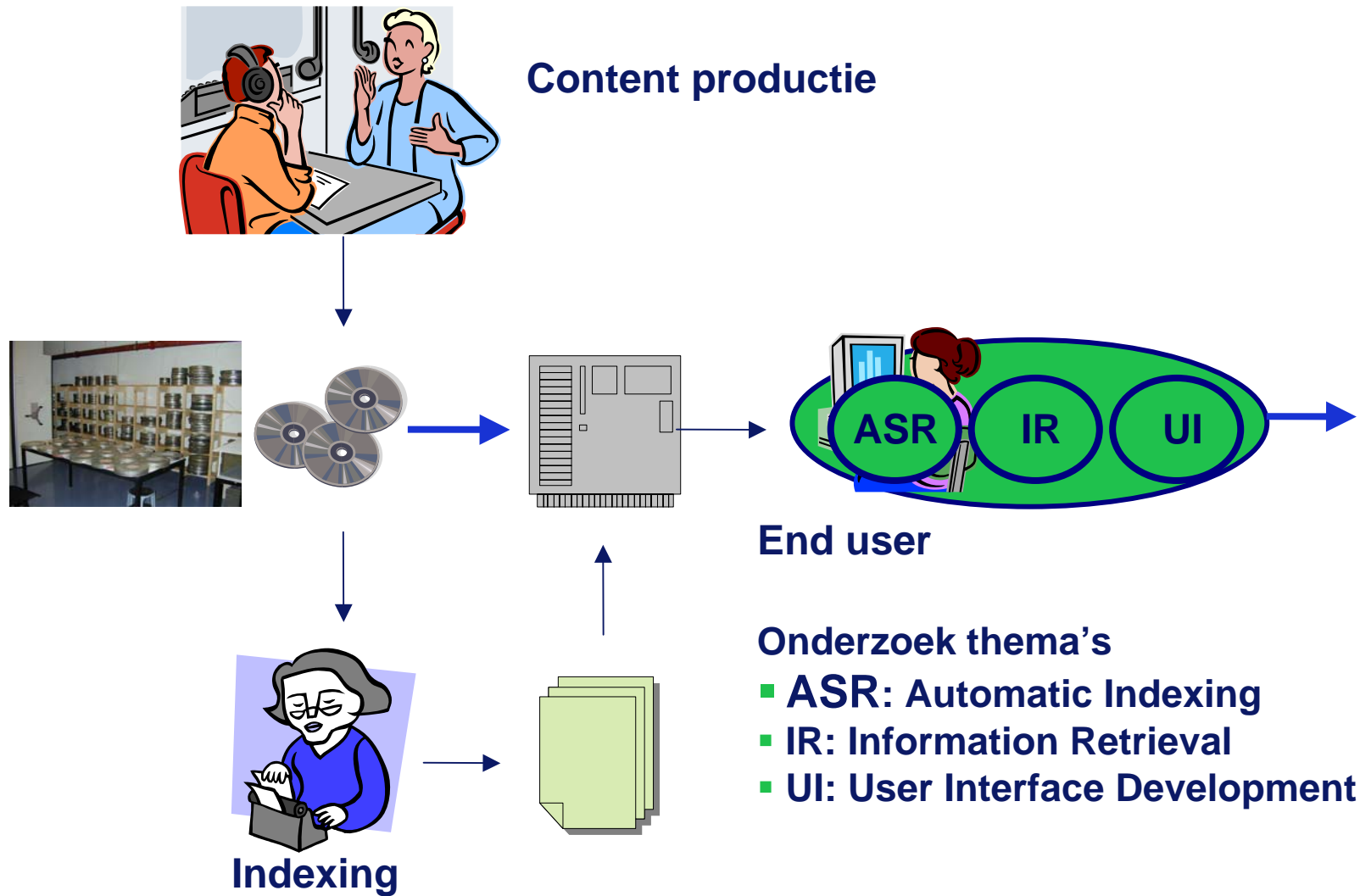
**Ontsluiting: genereer automatisch een tijdgemarkeerde inhoudsbeschrijving**

**→ Mogelijkheid om online audio fragmenten op te vragen met een bepaalde inhoud**

**Ontsluitingsmethode is afhankelijk van:**

- **Beschikbare metadata**
- **Beschikbaarheid van contextuele documenten (o.a. transcripties)**
- **Wanneer de Word Error Rates minder dan 40% bedragen, kan automatisch gegenereerde content beschrijving worden gebruikt als zoekindex**

# AV archiving workflow



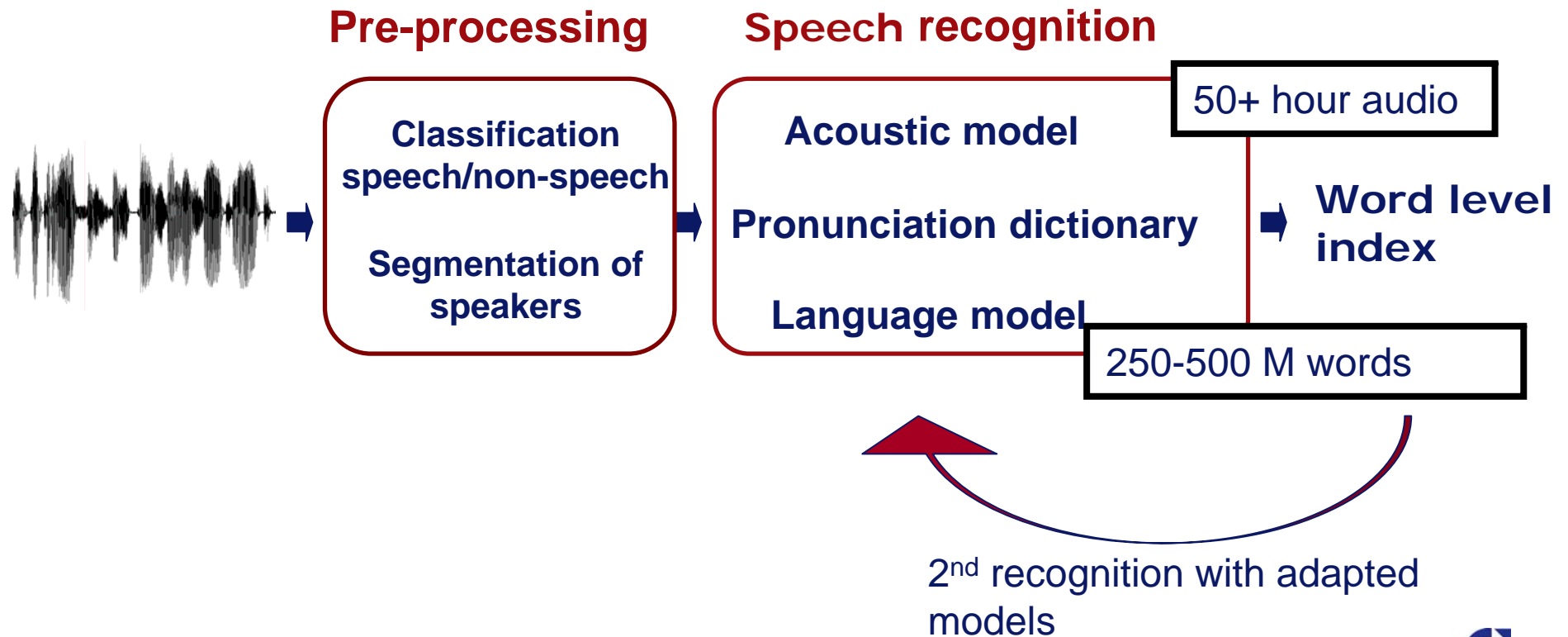
## Research en ontwikkeling in het huidige project Catch +

- **Automatische indexering d.m.v. spraaktechnologie:**
  - Ontwikkeling van robuuste ASR en classificatie tools
- **Information Retrieval:**
  - Ontsluiting van gesproken documenten gebaseerd op ASR output
  - Overbrugging van de semantische kloof tussen zoekvragen van de gebruiker en de gesproken content
- **User Interface ontwikkeling:**
  - Ondersteuning van zoeken en browsen in audio document
  - (Re)presentatie van audio content





# Automatic speech recognition



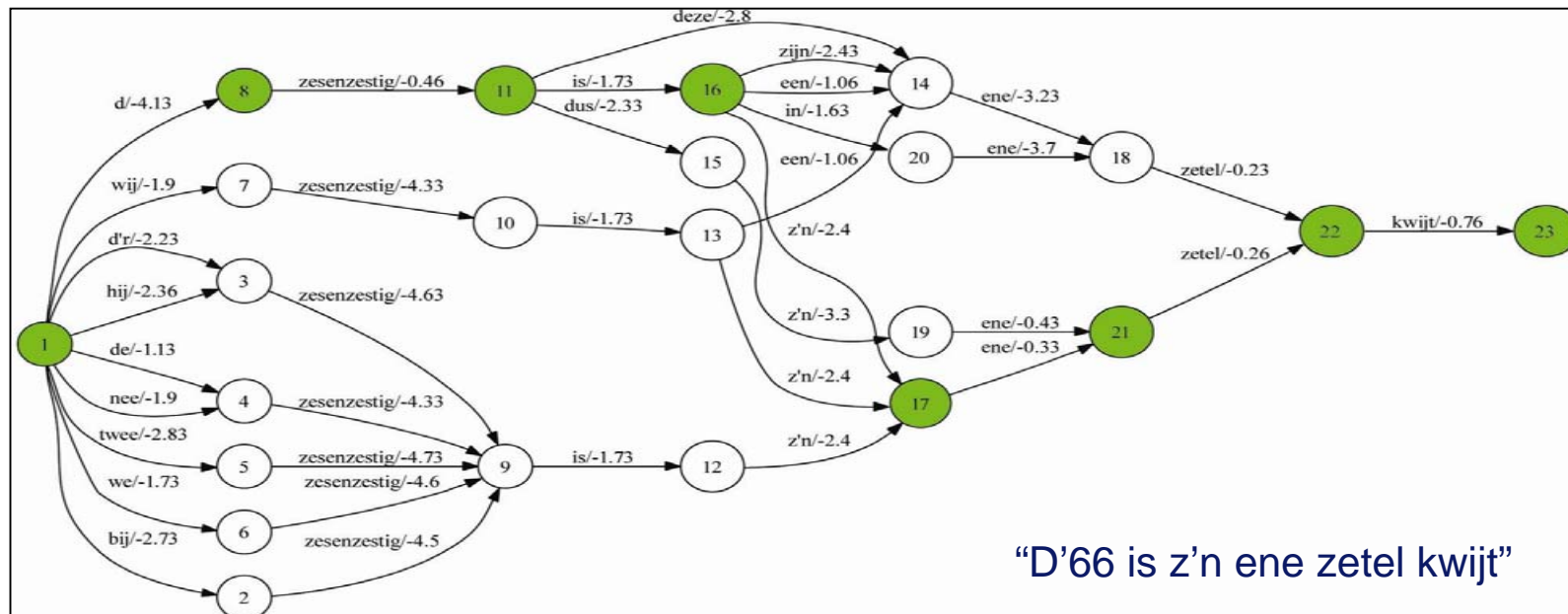
# Types of word level indexes

## Meest waarschijnlijke tekst:

ASR: Er is een bekend beeld voor veel ouders de grote show in onveilige situatie voor de school

TXT: 't is een bekend beeld voor veel ouders. De chaotische en onveilige situatie voor de school

## Lattice structures:



## Uitgangspunten en dilemma's

### **Om automatische annotatie succesvol te laten zijn:**

- Moet de Audio digitaal beschikbaar zijn, bij voorkeur op een centrale server

### **Om ASR modellen te optimaliseren voor hoge kwaliteit output:**

- Moet een deel van de gesproken tekst getranscribeerd worden
- Of moeten er aan de audio gerelateerde documenten beschikbaar zijn.

### **?Hoe valideren wij de geautomatiseerde indexen?**



# User interface ontwikkeling

- **Begrip voor de wensen en informatiebehoeften van de eindgebruiker**
- **Ondersteuning van selectie en browsen van gesproken content**
  - Bijvoorbeeld door representatie van gesproken tekst via andere media, zoals tekstueel, visueel of beide
- **Representatie in context met gerelateerde content binnen dezelfde of andere collecties**
- **Auteursrechtelijke perikelen**



# Uitdagingen

- De workflow en de dagelijkse praktijk bij audiovisuele archieven enerzijds en de state-of-the-art in technologie anderzijds vragen om een zorgvuldige afstemming om een goed resultaat neer te zetten
- zie: <http://hmi.ewi.utwente.nl/project/CHoral>

