

Workshop: Automatic Speech Recognition (ASR) in de Media Suite

Digitale Televisiegeschiedenis 2023-2024

In deze tutorial leren studenten:

- Hoe ze kunnen zoeken in de ASR transcripten
- Hoe ze een ASR zoekopdracht kunnen combineren met Boolean operators
- Hoe ze een ASR zoekopdracht kunnen combineren met Linked Data
- Hoe ze een ASR zoekopdracht kunnen combineren met Search facets
- Hoe ze in de ASR transcripten kunnen navigeren

→ Reflecteren op de ASR's beperkingen en potentieel in het digitale televisie archief

Stap 1: Voorbereiding

Studenten loggen in, en maken een 'user project' aan.

Stap 2: Selecteer de collectie en de ASR enrichment

Studenten selecteren de radio collectie, en hierna zowel de ASR-laag als het 'date field'.

Stap 3: Het formuleren en verfijnen van een zoekopdracht

3.1 Formuleren van een zoekopdracht

Reflectievraag: Vanwege de verschillende taalvarianties zijn '**Eurovisie**' en '**Eurovision**' de meest logische zoekopdrachten om mee te beginnen. Neem een moment om deze verschillende varianties uit te proberen. Wat voor resultaten krijg je? Welke term geeft de meest relevante resultaten?

Antwoord:

- 'Eurovisie' geeft 4565 resultaten, met pieken tussen 2003-2018. De resultaten zijn voornamelijk programma's over het Eurovisie songfestival zoals interviews met artiesten, nieuws over het festival of speciale thema programma's zoals '50 jaar eurovisie songfestival'
- 'Eurovision' geeft maar 24 resultaten, een groot verschil dus. Deze zijn ook zeer sporadisch tussen 1985 en 2022. Een hoop resultaten zijn hier vreemde zinnen, en zijn hoogstwaarschijnlijk foutjes van de spraakherkenner. We zien echter wel dat de ASR de

Engelse zoekterm wel kan herkennen, en dat enkele relevante resultaten dus ook wel naar voren komen.

- 'Eurovisi*' geeft 5007 resultaten. Ook met de meeste pieken tussen 2003-2018. Hier zitten dus nog extra variaties bij, zoals 'Eurovisiesongfestival' aan elkaar.

3.2. De zoekopdracht verfijnen: Boolean operators

Inhoudelijke vraag: Studenten oefenen met de AND / NOT boolean operators en noteren hun resultaten. Eventueel klassikaal bespreken bij welke case studies ze dit hebben geprobeerd, en met welke resultaten.

3.3 De zoekopdracht verfijnen: Search facets

Studenten proberen hun eigen opties uit, en slaan interessante items op in hun user project. Eventueel vragen welke facets ze hebben gebruikt, en welke resultaten ze hierbij kregen.

Details voorbeeld:

- Muziekprogramma, uitvoering en uitzending geven relatief vergelijkbare resultaten. Dit zijn grotendeels radioprogramma's rondom het thema muziek die verslag doen van het ESF, of hierover spreken. 'Muziekuitzending' geeft de meeste resultaten, en lijkt de tag te zijn die het meeste is gebruikt door het archief.
- 'Nieuws' lijkt voornamelijk integrale opname uitzendingen te geven, en is qua resultaten wat beperkt. 'Actualiteiten' daarentegen geeft meer dan 1500 resultaten die zeer verschillend zijn. Dit lijkt ook veel nieuws in andere programma's te zijn, waarin het ESF is besproken. Waarschijnlijk vanuit een andere hoek dan bij een muziekprogramma, dat weer meer inhoudelijk op de uitzending in zal gaan.

3.4 De zoekopdracht verfijnen: Linked Data

Studenten zoeken naar eigen voorbeelden, of het gegeven voorbeeld, via Linked Data en slaan interessante items op in hun user project. Eventueel vragen naar wie ze hebben gezocht, en met welke resultaten.

Is het ook bij iedereen gelukt de Linked Data, ASR en Boolean operator samen te gebruiken?

Stap 4: Close reading van het programma

4.1 (1) Reflectievraag: Wat hoor je tijdens de lichte en de donkere blokken?

Antwoord: De donkere blokken zijn de momenten waar de ASR spraak heeft herkend, en de lichte blokken zijn stilte, of in ieder geval momenten waarop de ASR geen spraak heeft kunnen oppakken.

4.1 (2) Reflectievraag: Hoe relateren de woorden in de woordwolk tot wat je hoort? (Hoe) kan dit overzicht gebruikt worden om door het segment te navigeren?

Antwoord: De woordwolk geeft de belangrijkste woorden weer die in het volledige spraak transcript aanwezig zijn. Het geeft je een idee over het onderwerp van dat segment. Het kan dus gebruikt worden om een idee te krijgen van het onderwerp van gesprek en zo te navigeren naar het moment in de uitzending die past bij je zoekopdracht, of om in het algemeen een idee te krijgen waar de uitzending over gaat.

Bijvoorbeeld: In de 'Met het oog op morgen' uitzending zien we de woorden 'Bulgaren, rust, slag, robin, persie, nul kort, strafschoep, extra, tijd, winnende, treffer etc. het is dus makkelijk te herleiden dat we hier een verslag van een voetbal wedstrijd kunnen vinden.

4.2 Reflectievraag: Vergelijk de woordwolk met het volledige transcript. Wat valt je op? Vind je de woordwolk zinvol voor analyse?

Antwoord: Persoonlijk antwoord, afhankelijk van de studenten. Maar over het algemeen voegt de woordwolk niet heel veel toe als je ook het speech transcript ernaast hebt. Het zijn dezelfde woorden, maar dan minder compleet. Hoewel de woordwolk handig kan zijn voor navigatie, kan je ook door de audio navigeren d.m.v. het volledige spraak transcript. Dit maakt de woordwolk wat overbodig.

Reflectievraag: Scan door het transcript van de gekozen uitzending:

- Klopt het transcript helemaal? Kan je ook fouten vinden? (Hoe vinden ze de kwaliteit, konden ze veel fouten vinden?)
- Hoe betrouwbaar lijkt het transcript? Wanneer is het minder betrouwbaar?
(**Antwoord:** Bijvoorbeeld bij snelle spraak, interrupties, andere taal of muziek is het transcript minder betrouwbaar.)
- Wat voor implicaties heeft dit voor de conclusies die je kan trekken uit je ASR zoekopdrachten?

(**Antwoord:** Dit betekent dat, wanneer we de ASR gebruiken voor het zoeken van items, de resultaten nooit 100% volledig zullen zijn (zoals bij elke zoekopdracht). We zullen

resultaten hebben die niet relevant zijn, en waarschijnlijk ook resultaten missen waarbij de zoekopdracht niet is herkend door de ASR.)

Plenaire reflectievragen:

- Voor wat voor soort zoekopdrachten zou je de ASR gebruiken?

Antwoord: De ASR doorzoekt de spraak, en zal dus ook items vinden waarbij je zoekopdracht niet centraal genoeg is om in de metadata beschrijving te staan. Neem bijvoorbeeld een talkshow waar het in de zijlijn wordt genoemd. Het is dus echt een 'verrijking' van de metadata en kan meer resultaten opleveren.

Ook is het handig als je op zoek bent naar resultaten met een transcript. Die kan een hoop werk schelen, als je dit niet handmatig hoeft te transcriberen.

- Voor wat voor zoekopdrachten vind je de ASR minder bruikbaar?

Antwoord: De ASR is minder bruikbaar voor woorden in andere talen, of wanneer je bepaalde spraak zoekt in een andere taal of met een sterk accent. Daarbij heeft maar een klein deel van de collectie al ASR gekregen, je moet dus altijd goed kijken waar je precies in zoekt.

Als je graag een zo compleet mogelijke corpus wilt, dan is het verstandig ook buiten de ASR te zoeken om je corpus zo aan te vullen. Door de fouten en gaten in de ASR willen er weleens items missen. Op deze manier kan je de twee zoektechnieken combineren.