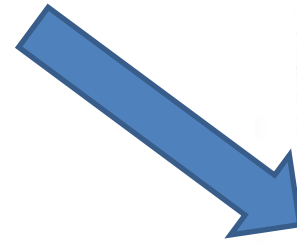


# Language Technologies – A Must for Innovation in Broadcasting

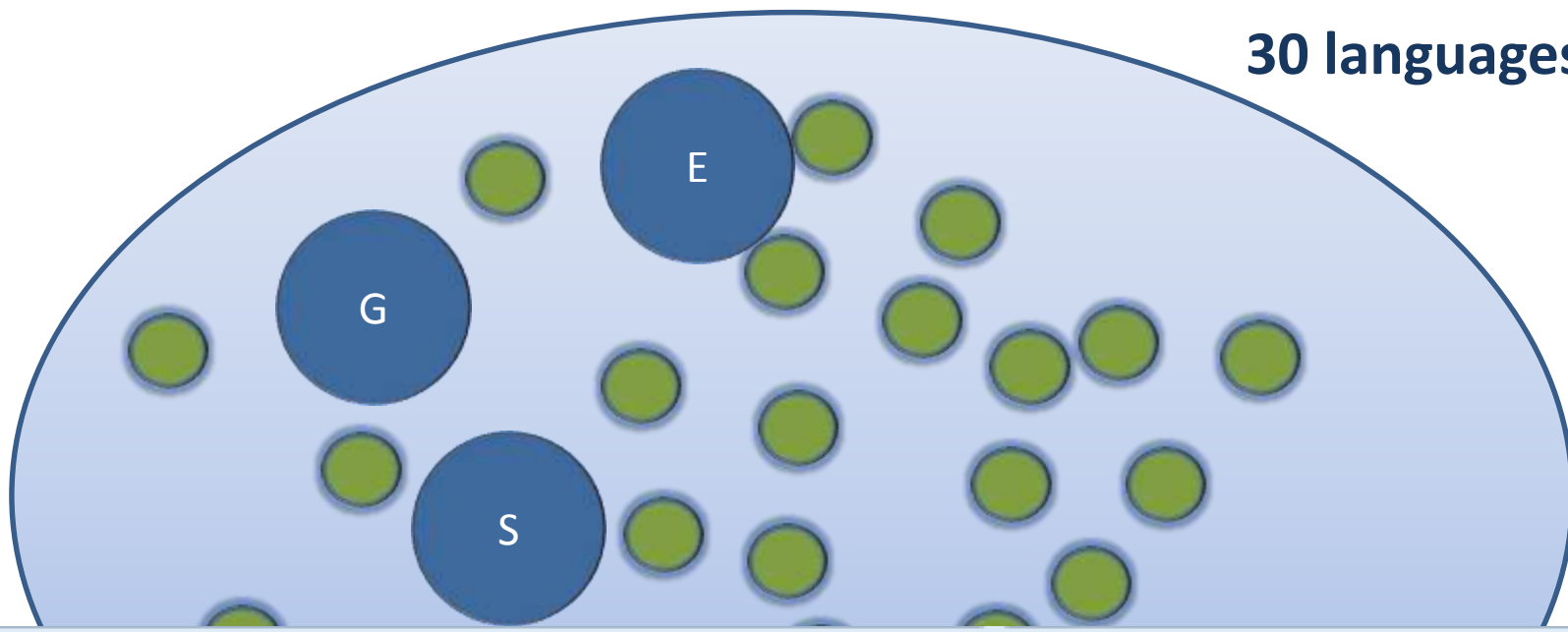
Peggy van der Kreeft  
Deutsche Welle/Innovation Department  
Riga Summit 2015 on the Multilingual Digital Single Market  
27-29 April 2015

# Broadcasting Objectives



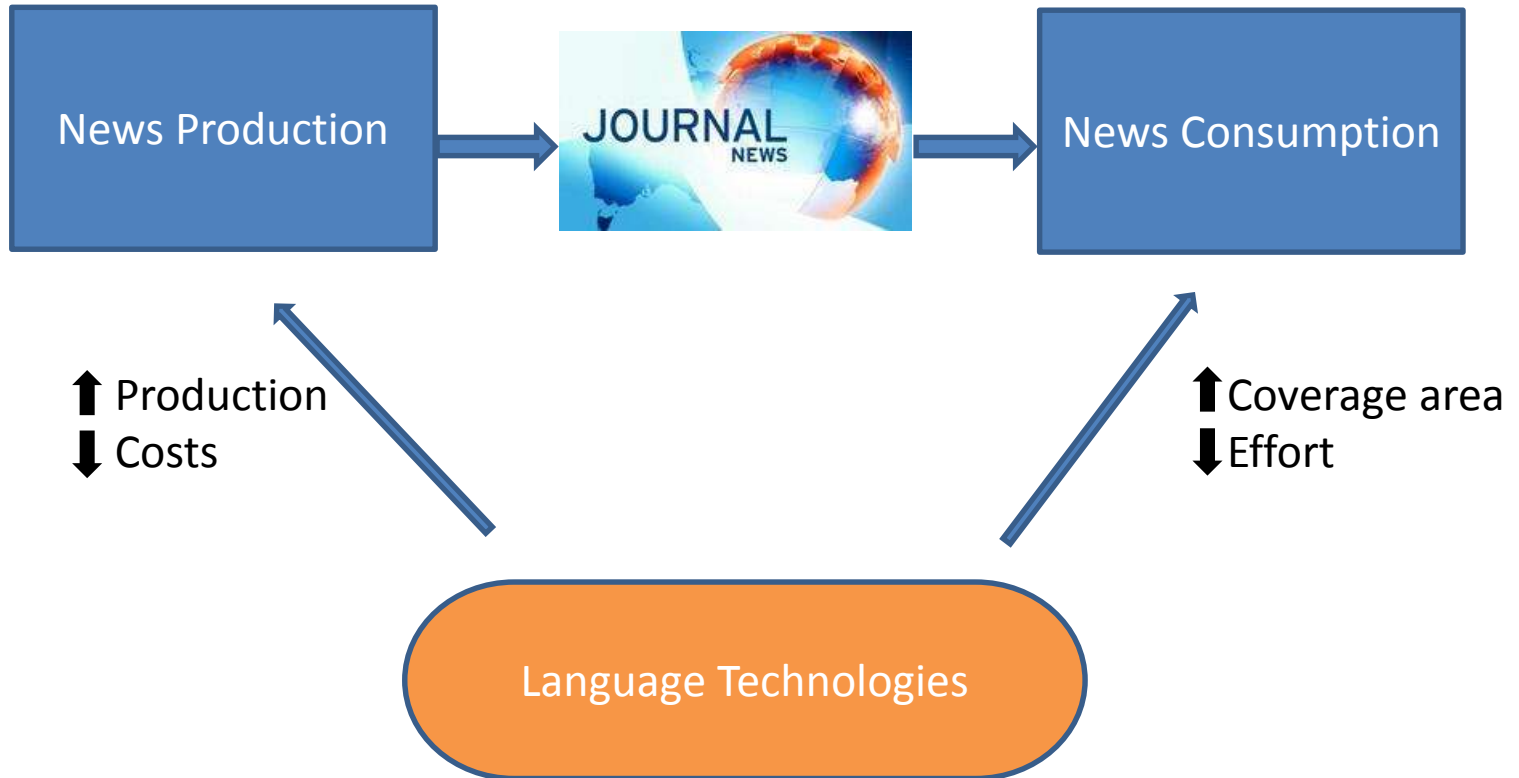
# Deutsche Welle World Broadcaster

30 languages

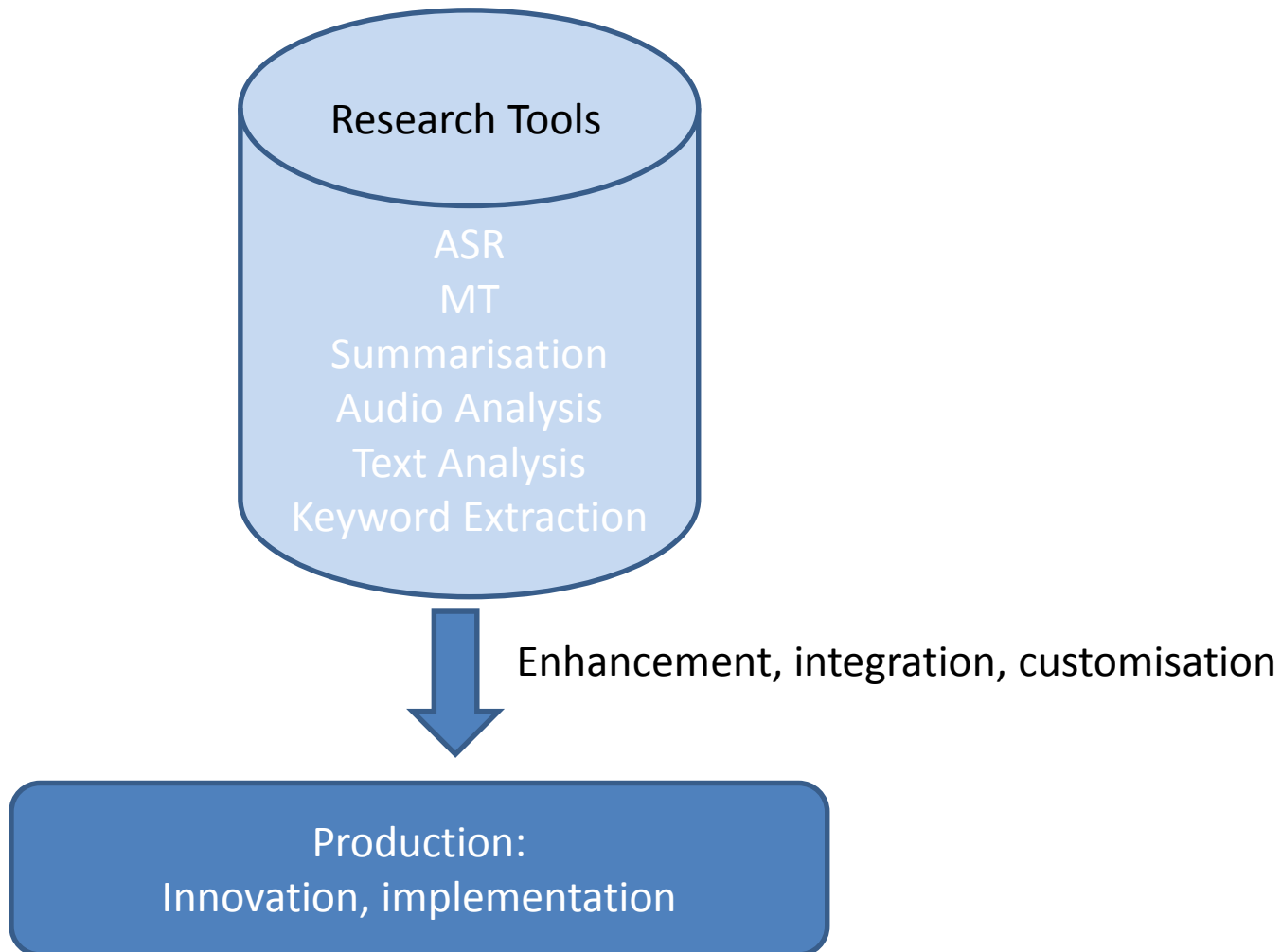


- |                     |                             |                                  |                       |
|---------------------|-----------------------------|----------------------------------|-----------------------|
| Albanian Shqip      | Dari دری                    | Kiswahili Kiswahili              | Russian на русском    |
| Amharic አማርኛ        | ✓ English English           | Macedonian Македонски            | Serbian Српски/Srpski |
| Arabic العربية      | French Français             | Pashto پښتو                      | Spanish Español       |
| Bengali বাংলা       | German Deutsch              | Persian فارسی                    | Turkish Türkçe        |
| Bosnian B/H/S       | Greek Ελληνικά              | Polish Polski                    | Ukrainian Українська  |
| Bulgarian Български | Hausa Hausa                 | Portuguese Português do Brasil   | Urdu اردو             |
| Chinese 中文          | Hindi हिन्दी                | Portuguese Português para África |                       |
| Croatian Hrvatski   | Indonesian Bahasa Indonesia | Romanian Română                  |                       |

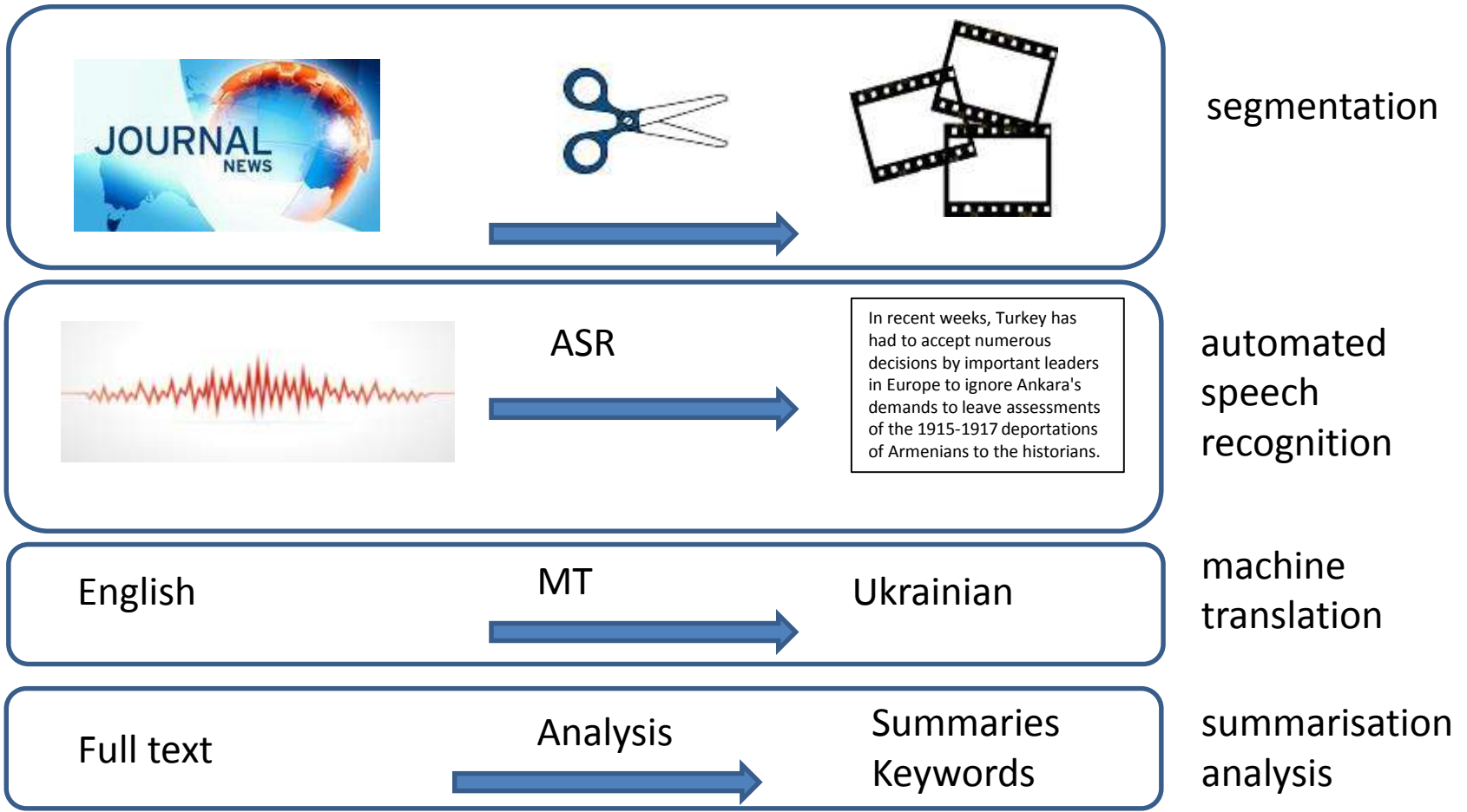
# Language Technologies



# Innovation



# Combination of Technologies



# Text Processing

Manual

English

German

French

Spanish

Ukraine

Automated

English

MT  
↓  
+ post-editing

French

Monitoring  
Reuse of content  
Breaking news

# AV Content Processing



ASR



*Post-editing*

English Transcript

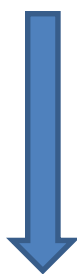
MT



*Post-editing*

Portuguese text + subtitles

TTS



Voiceover in Portuguese



Compilation



Subtitles in English  
Subtitles in Portuguese

Reuse of content  
Video-on-demand  
Visually impaired  
Breaking news



# Some solutions from research projects

## CoSyne Project

Multilingual Content Synchronisation with Wikis

Automated translation

Translation     Origin  
 Concepts     Align     Overlap

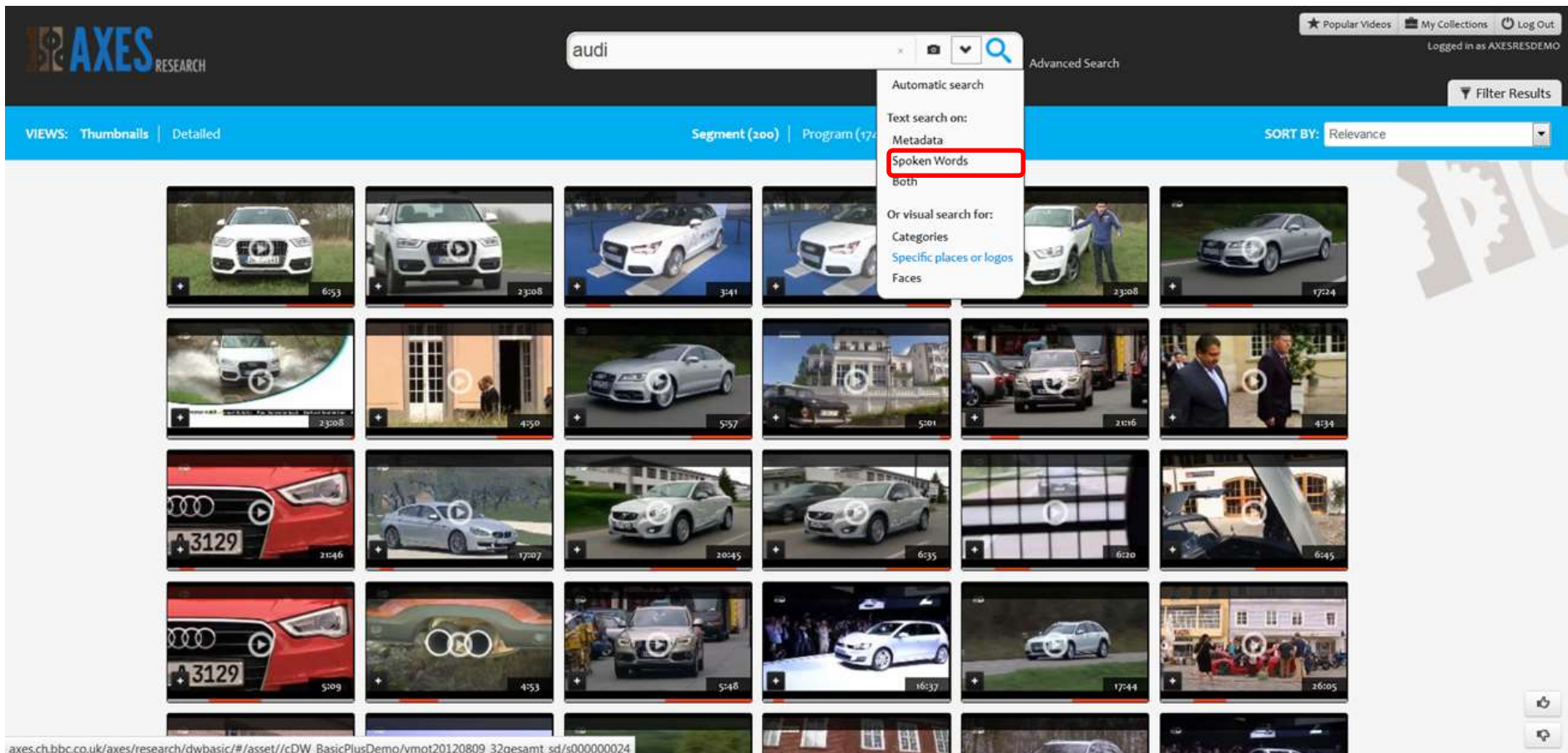
<p>The port city of</p> <p>Amsterdam is the capital of the Netherlands and its most populated city. Although the seat of government of the country, as well as the royal residence, are 60 kilometres away in the Hague, Amsterdam since 1983, is under the Dutch constitution, the capital of the Netherlands.</p> <p>Amsterdam is located in the province of North Holland, at the mouth of the Amstel and the IJ into the IJsselmeer.</p> <p>Amsterdam is connected with the North Sea by the Noordzeekanaal, and is world famous for its large number of Grachten.</p>	<p>Amsterdam is the largest city and the capital of the Netherlands. Its principal language is Dutch. The city's status as the capital of the nation is governed by the constitution.</p> <p>Amsterdam is located in the province of North Holland.</p> <p>It comprises the northern part of the Randstad, one of the larger conurbations in Europe, with a population of approximately 7 million.</p> <p>Amsterdam's name is derived from Amstelredamme, indicative of the city's</p>

Alignment at sentence and document level

Identify overlapping and unique information across languages

# AXES Tool

## Access to Audiovisual Archives



The screenshot displays the AXES Research interface. At the top left is the AXES RESEARCH logo. A search bar contains the text 'audi'. To the right of the search bar is an 'Advanced Search' button. Below the search bar, there are navigation options: 'Views: Thumbnails | Detailed', 'Segment (200) | Program (174)', and 'Sort By: Relevance'. A dropdown menu is open from the search bar, listing search options: 'Automatic search', 'Text search on: Metadata', 'Spoken Words' (highlighted with a red box), 'Both', 'Or visual search for: Categories', 'Specific places or logos', and 'Faces'. The main area shows a grid of video thumbnails, many of which feature a play button icon, indicating they are video results. The thumbnails show various Audi vehicles and scenes. At the bottom left, there is a URL: [axes.ch.bbc.co.uk/axes/research/dwbasic/#/asset//cDW\\_BasicPlusDemo/vmot20120809\\_32gesamt\\_sd/s000000024](https://axes.ch.bbc.co.uk/axes/research/dwbasic/#/asset//cDW_BasicPlusDemo/vmot20120809_32gesamt_sd/s000000024)

Audio Search + Language-independent visual search using Überclassifiers


# ASR Tool

nm-audioportal.iis.fraunhofer.de/Videoportal\_fwakes/portal?search Zoeken

< Aktueller Clip (1 von 1): 14\_mgwide-5011-made-in-g01ep\_tunesien\_sd\_dwdownload >

**Clips zum Thema produktion**

14\_mgwide-5011-made-in-g01ep\_tunesien\_sd\_dwdownload



**ABDALLAH ADUAM**  
Produktionsleiter Standort Tunesien

**MADE IN GERMANY**

Vide

ohne wirke ehrlich wir waren total blockiert vor allem die Jungen die wussten gar nicht was sie machen sofen die können ja gar nicht sagen man konnte nur über banale Dinge sprechen zum Beispiel was im Fernsehen läuft auf MTV sie kann nicht über ihre Probleme über ihre Arbeitslosigkeit Leben das ist jetzt anders nochmal

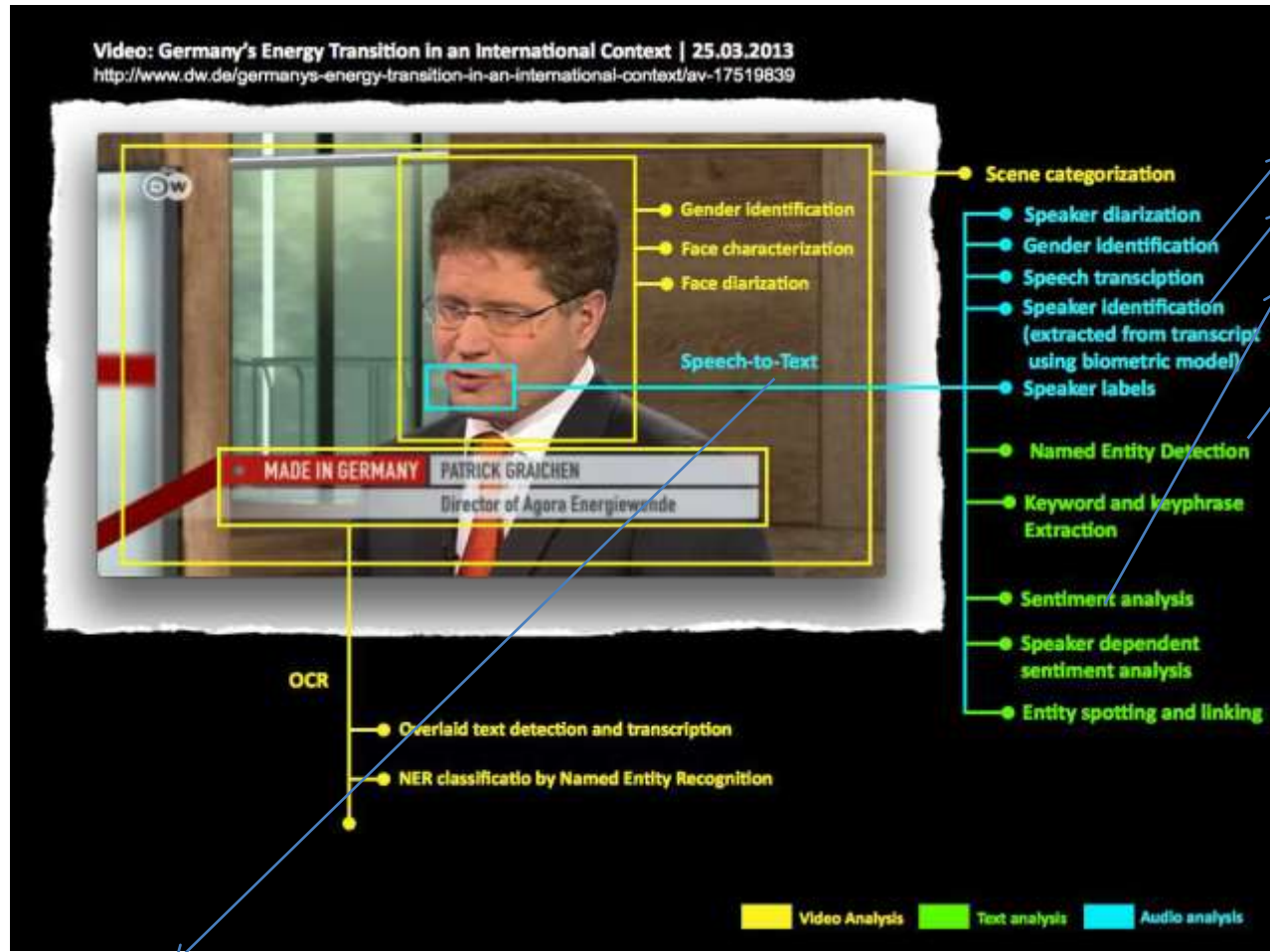
mit

der Heine Get lebt von der deutschen Firmen



# EUMSSI Tool

Event Understanding through Multimodal Social Stream Interpretation



gender identification  
speaker identification  
sentiment analysis  
named entity detection

STT: different language modules

Storytelling using analysis across languages

# Resulting in...

Entity Extraction

Keywords

Multilingual Transcripts

Multilingual Subtitles

Summaries

# Potential Applications

Broadcast content enhancement  
Text, audio, video  
Transcripts  
Subtitles

Social media  
User-generated content

Metadata  
enrichment

Data mining  
Data-driven journalism  
Multilingual media monitoring  
Business intelligence



Thank you !  
Questions?

