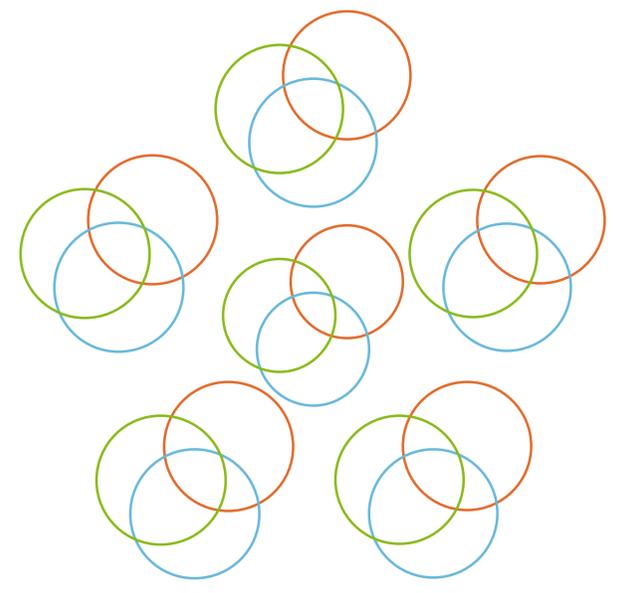




Nutzung automatisch extrahierter Metadaten

Redakteure oder Mediendienstleister übergeben Videomaterial mittels einer einfachen Datei-basierten Schnittstelle an unser **Analyse-System AMOPA**. Annotationsprozesse wie Szenenwechsellerkennung, Texterkennung und Spracherkennung extrahieren dabei automatisch inhaltliche und strukturelle Beschreibungen des Materials. Die gewonnenen Metadaten werden im einfach weiterzuverarbeitenden **XML-Format** anschließend über die Datei-Schnittstelle bereitgestellt.

Die erzeugten Metadaten bilden die Ausgangsbasis für **Webarchive**, **redaktionelle Recherchertools** oder **mobile, personalisierte Dienste**. Exemplarisch haben wir das Wiederauffinden audiovisueller Medien mit dem **Xtrivial Framework** realisiert. Das System konnte seine herausragende Qualität wiederholt in internationalen Vergleichen auf wissenschaftlicher Ebene sowohl in der klassischen Text- und Bildrecherche als auch bei der Klassifikation von Videos unter Beweis stellen.



```
<?xml version="1.0" encoding="UTF-8"?>
<text_extraction>
  <file name="TU_Chemnitz_Immatrikulation.mp4" videoLength="294684">
    <text_recognition>
      <config>
        <textRecognitionEngine>
          <textRecognitionEngine>
            <gradientThreshold>500</gradientThreshold>
            <gradientThreshold>170</gradientThreshold>
          </textRecognitionEngine>
        </config>
      </text_recognition>
      <textResult frame="969">
        <boundingBox id="1" x="9" y="42" width="105" height="105">
          <textLine>Prof. Dr. Klaus-Jürgen Matthes</textLine>
        </boundingBox>
        <boundingBox id="5" x="207" y="81" width="41" height="41">
          <textLine>Rektor der TU Chemnitz</textLine>
        </boundingBox>
      </textResult>
    </file>
  </text_extraction>
  <speech_recognition>
    <file name="TU_Chemnitz_Immatrikulation.mp4" videoLength="294684">
      <speaker_recognition>
        <config>
          <processingWindowLength>3.0</processingWindowLength>
          <processingWindowStep>0.5</processingWindowStep>
          <subsegmentLength>0.025</subsegmentLength>
          <predictingThresholdDistance>40.0</predictingThresholdDistance>
          <predictingThresholdDistance>1.8</predictingThresholdDistance>
        </config>
      </speaker_recognition>
      <segment id="0" start="0" end="2371">
        <recognitionEvent id="0" start="0" end="1182" confidence="0.6633883" duration="1182" start="0">
          <word start="0" duration="1182" confidence="0.934082839">liebe</word>
          <word start="1201" duration="237" confidence="0.887228">Studierenden</word>
          <word start="1431" duration="570" confidence="0.9094018">der</word>
          <word start="1942" duration="250" confidence="0.6274768">TU</word>
          <word start="2122" duration="249" confidence="0.7754549">Chemnitz</word>
        </recognitionEvent>
      </segment>
    </file>
  </speech_recognition>
```



Redaktionelle Weiterverarbeitung

z.B. Apple Final Cut Server

Anfrage

Xtrivial: Videosuche

Web Frontend

Themen: arbeit, barbara, besuch, chemnitz, cottbus, dachau, etc.

NIF Immatrikulation TU
Klassifikation: Student, Erziehung, Lehren und Lernen, Ausbildung

NIF tu future truck
Klassifikation: Student, Erziehung, Lehren und Lernen, Schule, Ausbildung

Resultat

Mobiler Zugang