



PROFESSUR MEDIENINFORMATIK



Nutzung automatisch extrahierter Metadaten

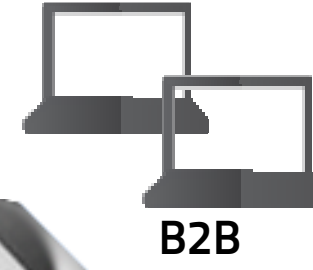
Redakteure oder Mediendienstleister übergeben Videomaterial mittels einer einfachen Datei-basierten Schnittstelle an unser **Analyse-System AMOPA**. Annotationsprozesse wie Szenenwechselerkennung, Texterkennung und Spracherkennung extrahieren dabei automatisch inhaltliche und strukturelle Beschreibungen 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 **Xtrieval 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.

Video



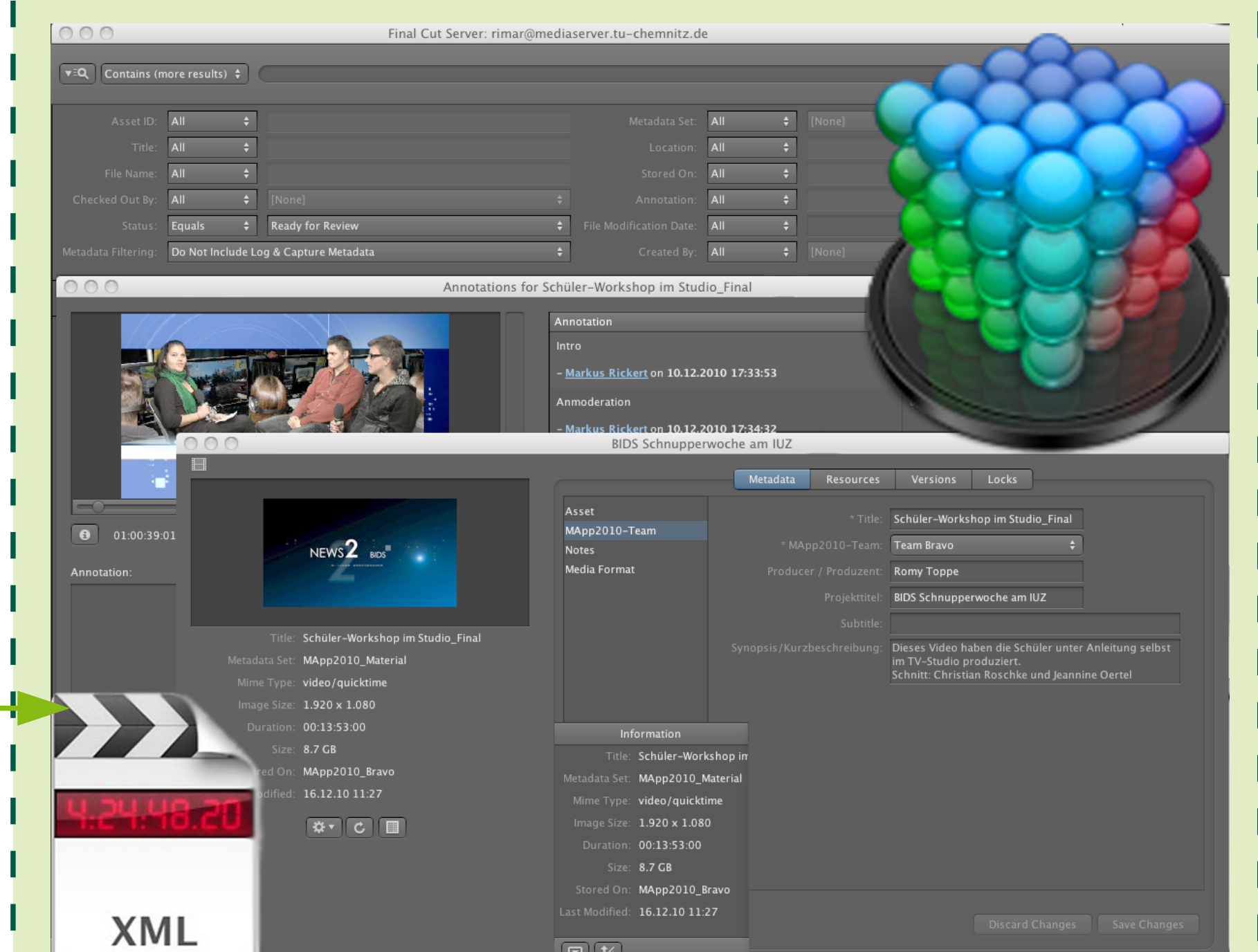
```
<?xml version="1.0" encoding="UTF-8"?>
<text_extraction>
  <file name="TU_Chemnitz_Immatrikulation.mp4" videoLength="294684">
    <text_recognition>
      <config>
        <TextRecognitionEngine>tesseract</TextRecognitionEngine>
        <PixelAmountThreshold>500</PixelAmountThreshold>
        <GradientThreshold>170</GradientThreshold>
      </config>
    </text_recognition>
    <textResult frame="960">
      <boundingBox id="1" x="9" y="43" width="105" height="
      <textLine>Prof. Dr. Klaus-Jürgen Matthes</textLine>
    </boundingBox>
    <boundingBox id="5" x="297" y="81" width="41" height="
    <textLine>Rektor der TU Chemnitz</textLine>
    </boundingBox>
    </textResult>
  </text_extraction>
  <speech_recognition>
    <file name="TU_Chemnitz_Immatrikulation.mp4" videoLength=
    <speaker_recogniton>
      <config>
        <ProcessingWindowLength>3.0</ProcessingWindowLength>
        <ProcessingWindowStepSize>0.5</ProcessingWindowStep>
        <SubsegmentLength>0.025</SubsegmentLength>
        <PredictingThresholdDistance>40.0</PredictingThreshold>
        <PredictingThresholdAlpha>1.8</PredictingThresholdAlpha>
      </config>
    </speaker_recogniton>
    <segment id="0" speaker_id="0" startTime="0" endTime="2371">
      <recognitionEvent id="0" confidence="0.662383" duration="2371" startTime="0">
        <word startTime="0" duration="1181" confidence="0.334062839">Liebe</word>
        <word startTime="1201" duration="250" confidence="0.6877283">Studierenden</word>
        <word startTime="1451" duration="570" confidence="0.9094018">der</word>
        <word startTime="1941" duration="190" confidence="0.624768">TU</word>
        <word startTime="2232" duration="180" confidence="0.7754349">Chemnitz</word>
      </recognitionEvent>
    </segment>
  </speech_recognition>
```



B2B

XML

Redaktionelle Weiterverarbeitung



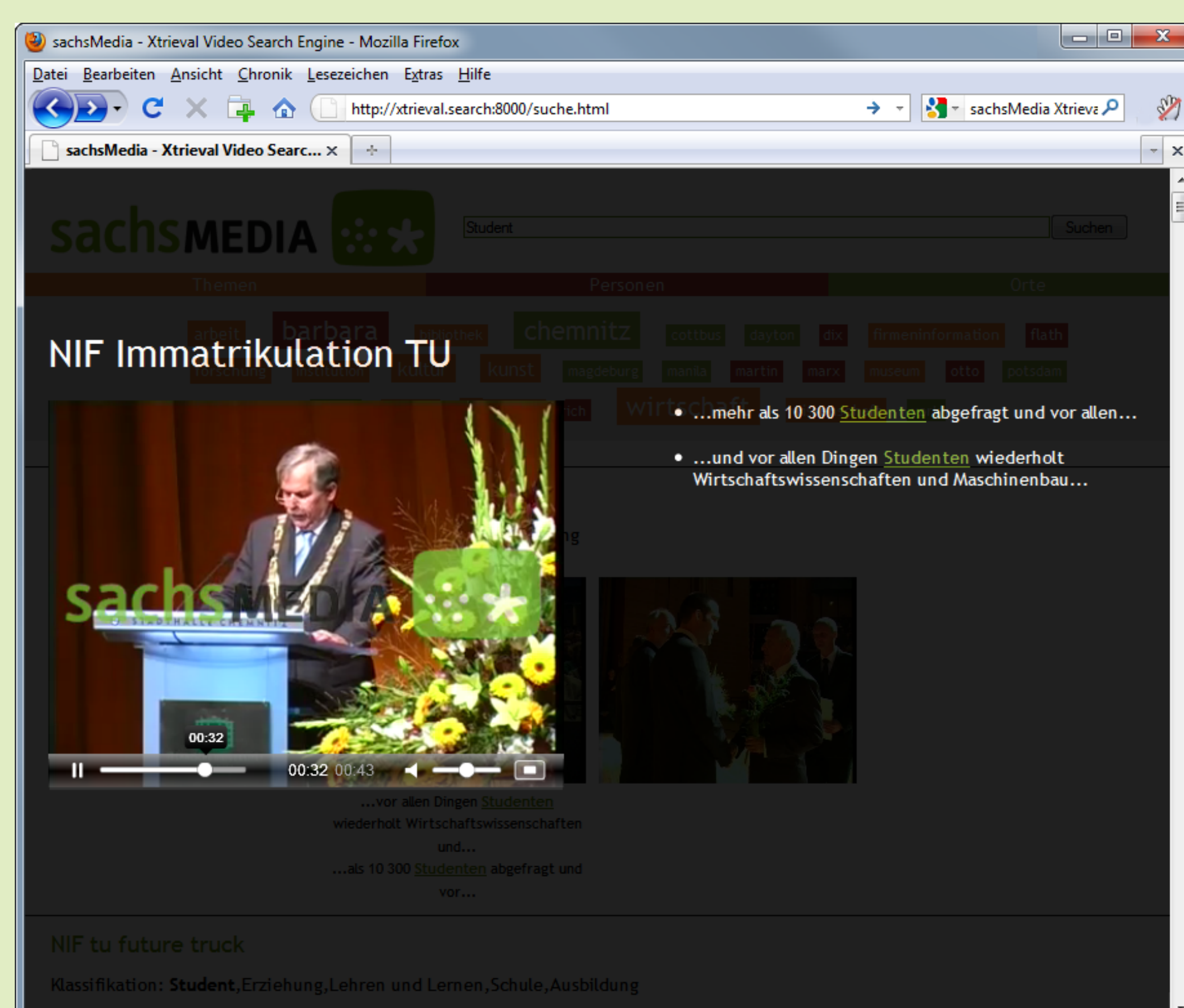
z.B. Apple Final Cut Server

Xtrieval: Videosuche

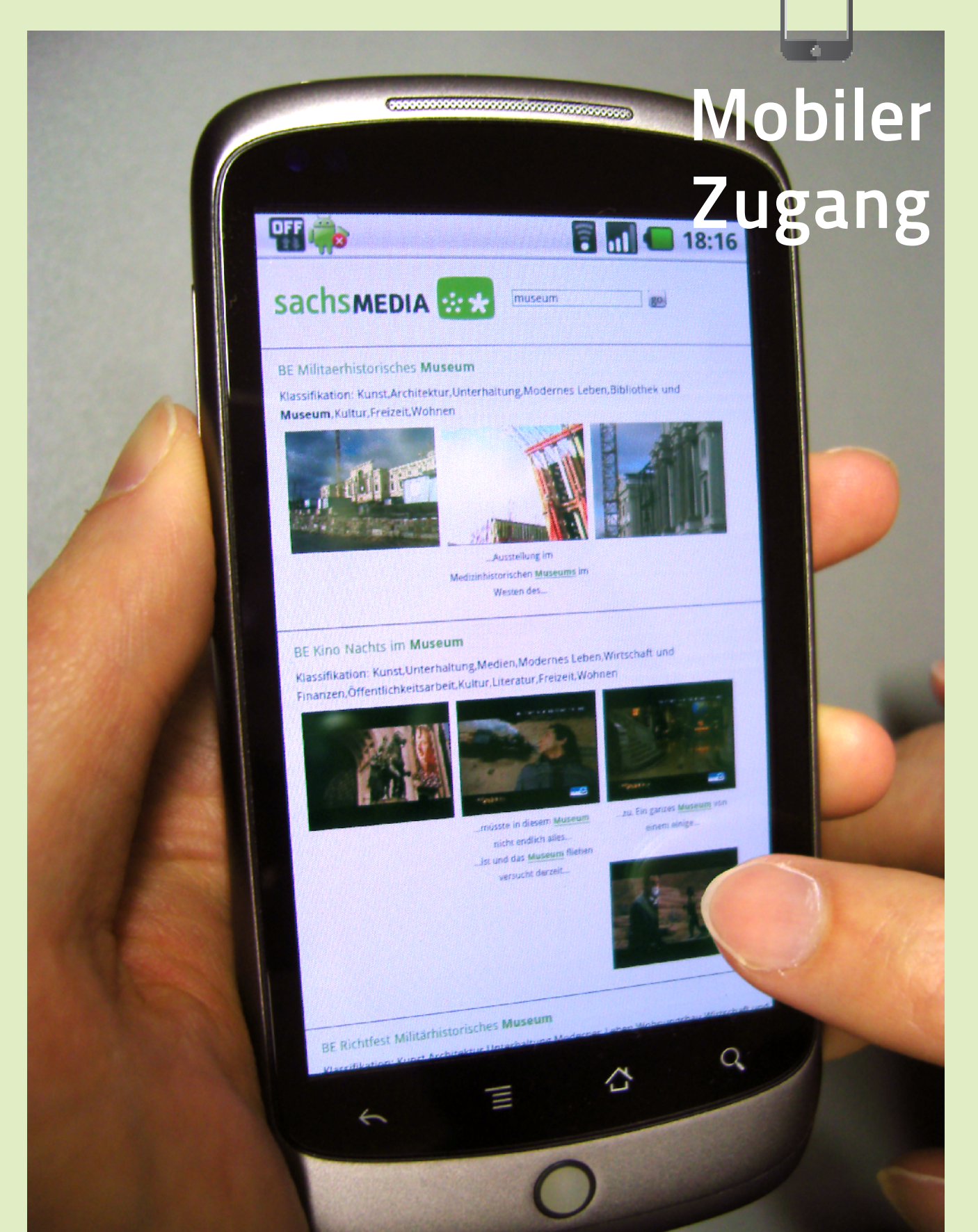
Anfrage



Resultat



Mobiler
Zugang



Kontakt:

Dipl.-Inf. Jens Kürsten · jens.kuersten@cs.tu-chemnitz.de
Technische Universität Chemnitz · Professur Medieninformatik
Straße der Nationen 62 · 09111 Chemnitz