

## TailoredMedia

Tailored and Agile enrichment and Linking FOR sEmantic Description of multiMedia

|                                 |   |                        |               |
|---------------------------------|---|------------------------|---------------|
| <b>Programm / Ausschreibung</b> | IKT der Zukunft, IKT der Zukunft, IKT der Zukunft - 8. Ausschreibung (2019) | <b>Status</b>          | abgeschlossen |
| <b>Projektstart</b>             | 01.11.2020  | <b>Projektende</b>     | 31.12.2022    |
| <b>Zeitraum</b>                 | 2020 - 2022   | <b>Projektlaufzeit</b> | 26 Monate     |
| <b>Keywords</b>                 | KI, Semantische Technologien, Multimedia, Medienproduktion                  |                        |               |

### Projektbeschreibung

Audiovisuelle Medien sind zu einem allgemein verwendeten Kommunikationsmittel geworden, sowohl in den traditionellen als auch in den sozialen Medien. Die etablierten Methoden zur Dokumentation und Indizierung von Multimedia-Inhalten, z.B. bei der Medienbeobachtung und Archivierung, ermöglichen (i) keine feingranulare Beschreibung (z.B. auf Szenen- und Objektebene) und (ii) skalieren nicht mit der ständig wachsenden Menge an Inhalten. Zwar liegen die Inhalte aller Modalitäten (Text aus Newsfeeds, Web, Produktionsinformationen, aber auch Audio, Bilder, Video) inzwischen in digitaler Form vor, doch werden die potenziellen Vorteile der Digitalisierung noch nicht voll ausgeschöpft: Unterschiedliche Inhaltsquellen werden oft unabhängig voneinander verarbeitet, ohne sie mit anderen Quellen oder externen Informationen in Beziehung zu setzen, und Kontextinformationen werden weder zur Steuerung des Annotationsprozesses noch zur Verbesserung seiner Robustheit genutzt.

TailoredMedia zielt darauf ab, die jüngsten Fortschritte bei der automatischen Analyse von visuellen Inhalten mit Hilfe von KI-basierten Methoden zu nutzen, um die Extraktion von Metadaten und die semantische Anreicherung für Anwendungen im Journalismus und in der Medienarchivierung zu unterstützen. Aufbauend auf zeitgemäßen Methoden für die visuelle Analyse (z.B. Objekterkennung, Szenenklassifikation, Gesichtserkennung) wird das Projekt zu KI-basierten Methoden für die multimodale Fusion unter Einbeziehung des Kontexts forschen. Der Einsatz von Active und Online Learning-Ansätzen wird auch "few shot learning" ermöglichen, d.h. effizientes Lernen neuer Klassen aus sehr geringen Mengen von annotierten Daten (z.B. 5-10 Beispielen). Die Analysewerkzeuge werden durch eine multimediale Wissensrepräsentation unterstützt, die semantische Informationen aus verschiedenen Quellen (einschließlich Linked Open Data) und Modalitäten integriert und eine interoperable Darstellung bietet. Dies ermöglicht es, Text- und Multimediainhalte mit semantischen Metadaten anzureichern. TailoredMedia verfolgt einen Human-Centred Design Ansatz, der es den Benutzern ermöglicht, die Kontrolle über komplexe Analyse- und Suchprozesse zu behalten. In der Designphase werden Anwender der Projektpartner und externer Organisationen einbezogen, wobei der Schwerpunkt auf Systemexperten und Redakteuren liegt. Im Rahmen des Projekts werden ein Interaktionsprototyp, ein funktionaler Prototyp und schließlich ein Demonstrator entwickelt und mit den Nutzergruppen validiert.

Um die Interoperabilität zu ermöglichen und die Bindung an einen bestimmten Anbieter zu vermeiden, werden die Ergebnisse von TailoredMedia als ein Ökosystem von Mirco-Services und Anwendungen realisiert, das offenen Standards wie

EBU MCMA folgt. Dies wird die Nutzung der Services auf eigener Infrastruktur und in Private oder Public Clouds ermöglichen. Die Services werden für den Informationsaustausch Metadaten-Standards, wie z.B. MPEG-7, EBU Core, W3C Web Annotation und W3C PROV-DM verwenden.

## **Abstract**

Audiovisual media have become the dominant means of communication, in both traditional and social media. The established practices to document and index multimedia content, for example in media monitoring and archiving, (i) do not allow for fine-grained description (e.g., on scene and object level) and (ii) do not scale to the ever-increasing amounts of content. While content of all modalities (text from newsfeeds, web, production information, but also audio, images, video) is now available in digital form, the potential benefits of digitisation are not yet fully leveraged: Different content sources are often processed independently, without interrelating them with other sources or external information, and contextual information is not exploited for guiding the annotation process or improve its robustness.

TailoredMedia aims to leverage the recent advances in automatic analysis of visual content using AI-based methods in order to support metadata extraction and semantic enrichment for use cases in journalism and archiving. Building on state-of-the-art methods for visual analysis tasks (e.g., object detection, scene classification, face recognition), the project will research AI-based methods for multimodal information fusion, and context-aware AI methods. The use of active and online learning approaches will also enable few shot learning, i.e. efficiently learning new classes from very limited amounts of labelled data (e.g., 5-10 samples). The analysis tools will be backed by a cross-media knowledge representation, integrating semantic information from different sources (including linked open data) and modalities, and providing an interoperable representation of contextual knowledge. This will enable enriching textual and media content descriptions with semantic metadata, and enable discovery and reasoning using the knowledge representation.

TailoredMedia will follow a human-centered design approach, allowing human users to stay in control of analysis and search processes using complex tools. In the design phase, stakeholders from the partners and beyond will be involved, focusing on system professionals and editors. In the project, an interaction prototype, a functional prototype and finally a demonstrator will be developed, and validated with the target user groups.

In order to enable interoperability and avoid vendor lock-in, the TailoredMedia results will be deployed as an ecosystem of microservices and applications, following open standards such as EBU MCMA. This will enable using the services on premise and in private or public cloud infrastructures. The services will follow metadata standards, for example MPEG-7, EBU Core, W3C Web Annotation and W3C PROV-DM, for the exchange of information.

## **Projektkoordinator**

- JOANNEUM RESEARCH Forschungsgesellschaft mbH

## **Projektpartner**

- Technisches Museum Wien mit Österreichischer Mediathek
- Österreichischer Rundfunk
- Hochschule für Angewandte Wissenschaften St. Pölten Forschungs GmbH
- RedLink GmbH