

Audiovisual collections – turning sound and vision into data

Kungliga biblioteket, hörsalen, 2 februari 2018

The potential indicated in the expression “a picture says more than thousand words” is today gradually being fulfilled in its literal sense. The development of digital applications for information processing means that this saying now comprises not only still-images (as in Optical Character Recognition) but also extends to information extraction in relation to moving images and sound (i.e. speech to text, object and feature detection). How can this technology be used to enhance the management of audio-visual collections within the heritage sector?

Heritage institutions have traditionally relied on manual methods for classification and indexing of the holdings. Audio-visual material has always presented challenges in this respect due to the dependency on technical equipment for access. This aspect makes the description of audio-visual content a resource intensive and time-consuming task, and potentially an insurmountable obstacle in relation to large collections.

The possibility of information extraction offered by the digital format is increasingly utilised in the management of text-based documents. The transfer to machine readable text enables automated methods of indexing and genre identification as well as research approaches, based on data driven analysis, commonly referred to as Text and Data Mining. Similar applications are found in relation to audio-visual information carriers although here the development tends to be driven by commercial interests rather than the need among heritage institutions to improve the management of their collections. The digital format consequently offers a great, and in many respects unexploited, potential for the dissemination of audio-visual content across the boundaries posed by the traditional institutional types within the heritage sector.

This conference is intended to present research projects and running operations where digital methods are used in order to extract text and metadata from audio-visual collections. We will examine some challenges and potential promises in this area and at the same time hopefully inspire to further development and applications among the institutions and organisations that manage audio-visual collections.

Programme

9:00-9:45	<i>Registration & coffee</i>	
9:45-10:00	Welcome	Eva-Lis Green, head of Digital collections department, National Library of Sweden
10:00-10:45	Embracing automation, enhancing discoverability; a state of the art from an end-user perspective	Jacqui Gupta, Vice President FIAT/IFTA and Lead Technologist/Information Scientist BBC
10:45-11:15	Smart selections – methods for choosing and processing speech and video for indexing	David House, professor & Jens Edlund, associate professor, Royal School of Technology, Stockholm
11:15-11:45	Automated metadata extraction at Yle – Combinations of people and technology for a broadcaster's various needs	Lauri Saarikoski, development manager, Yle, the Finnish Broadcasting Company
11:45-12:15	Extracting text and metadata from sound and video collections at the British Library	Richard Ranft, Head of Sound and Vision, the British Library
12:15-13:45	<i>Lunch</i>	
13:45-14:15	Harvesting data from video using commonly available tools	Peter Krantz, CIO, National Library of Sweden
14:15-14:45	The I-Media-Cities project – extracting metadata from video	Uwe Kühhirt, Head of Metadata Department, Fraunhofer Institute for Digital Media Technology
14:45-15:15	Illusive Proximity/Concrete Distance: On Mining & Mapping Local Audiovisual Data via Distant Access	Mats Jönsson, professor, Department of Cultural Sciences, University of Gothenburg
15:15-15:45	<i>Coffee</i>	
15:45-16:30	Summary and observations	Johan Oomen, head of Research and Development dpt., Beeld en Geluid, the Netherlands Institute for Sound and Vision

Moderator: Pelle Snickars, professor, Department of Culture and Media studies, Umeå university.
The conference is held in English