Video Search: Are Algorithms All We Need?

Jeff Ubois
Archival.tv
Berkeley, CA 94708
jeff@ubois.com

Jamie Davidson
YouTube
San Bruno, CA 94066
Davidson@google.com

Marko Grabolnik
Jožef Stefan Institute
Ljubljana, Slovenia
Marko.Grobelnik@ijs.si

Paul Over
NIST
Gaithersburg, MD 20899
Over@nist.gov

Hans Westerhof
Beeld en Geluid
Hilversum, The Netherlands
hwesterhof@beeldengeluid.nl

ABSTRACT
This panel will debate various approaches to improving video search, and explore how professional cataloguing, crowd sourced metadata, and improvements in search algorithms will evolve over the next ten years. Panelists will explore the needs of large scale video archives, and compare these against the current capabilities of video search.

Categories and Subject Descriptors
H.3 INFORMATION STORAGE AND RETRIEVAL
H.3.7 Digital Libraries

General Terms
Algorithms, Design, Experimentation.

Keywords
Video, search, algorithms, design, archives

1. INTRODUCTION
During the 1990s, search technology improved sufficiently to handle large volumes of textual material without the need for manual abstracting, indexing, and cataloging. Taking professional cataloguers out of the process of text indexing created enormous value; an analogous set of advances is underway with video in the 2010s. But will it be possible to take professionals out of the process of cataloging video manually? How might different approaches to video search to evolve, and which ones will prove most useful? What kind of societal value can we reap by making all of our broadcast history as readily accessible as books, journal articles, and newspapers are now? How can the W3C help with the development of appropriate standards for video description and search? Answers to and debate about these questions will be the focus of this panel, which has representatives from some of the largest, most viewed, and carefully indexed collections of digital video available online.

WWW 2010, April 26–30, 2010, Raleigh, North Carolina, USA.
ACM 978-1-60558-799-8/10/04.

2. PARTICIPANT BIOGRAPHIES
Jamie Davidson is Product Manager, on Search and Algorithmic Discovery, YouTube. Video discovery at YouTube includes relevance based search ranking, for which Google is well known, as well large scale video recommendation engines. Prior to YouTube, Jamie worked as an engineer on the Ads Quality team for Google's search ad system for auctions, ranking and pricing.

Marko Grobelnik is Technical Associate in the Department of Knowledge Technologies, Jožef Stefan Institute in Ljubljana, Slovenia, and Program Manager for VideoLectures.net, which provides free and open access of a high quality video lectures presented by distinguished scholars and scientists at the most important and prominent events like conferences, summer schools, workshops and science promotional events from many fields of Science, and operates within the FP5, FP6, and FP7 European Framework Programs.

Paul Over is a computer scientist at the US National Institute of Standards and Technology and founding project leader for the TREC Video Retrieval Evaluations (TRECVID) - an international effort since 2001 to promote open, metrics-based research in video retrieval and analysis systems, both interactive (human-in-the-loop) and fully automatic. Paul has also been responsible at NIST for evaluation of interactive text retrieval systems within the TExt REtrieval Evaluations (TREC). Before joining NIST he applied his interests in linguistics and computer science in advanced software product development for IBM.

Jeff Ubois is a consultant on digital preservation to the Beeld en Geluid, and to Fujitsu Labs of America. Prior to this, Jeff was part of the Preserving Digital Public Television Project at Thirteen/WNET, and a staff research associate at the iSchool at UC Berkeley. For the Internet Archive, Jeff has developed policies to manage images of war, and limit the collection and retention of digital library usage data.

Hans Westerhof is program director Images for the Future at the Netherlands Institute for Sound and Vision (Beeld en Geluid), which looks after, and releases, 70 per cent of the Dutch audio-visual heritage. In total, Sound and Vision takes care of 700,000 hours of television, radio, music and film, making it a supreme collection full of national and international memorable events. In the project Images for the Future, Sound and Vision will digitise 280.000 hours of material.