CHANGING SCENERIES
CHANGING ROLES PART VI
METADATA AS THE CORNERSTONE OF DIGITAL ARCHIVING

SELECTED PAPERS FROM THE FIAT/IFTA MEDIA MANAGEMENT SEMINAR, HILVERSUM 2013
CHANGING SCENERIES
CHANGING ROLES PART VI
METADATA AS THE CORNERSTONE OF DIGITAL ARCHIVING

SELECTED PAPERS FROM THE FIAT/IFTA MEDIA MANAGEMENT SEMINAR
CHANGING SCENERIES, CHANGING ROLES PART VI
HILVERSUM 16TH - 17TH MAY 2013

ORGANISED BY
THE FIAT/IFTA MEDIA MANAGEMENT COMMISSION
AND BEELD EN GELUID

DESIGNED BY AXEL GREEN FOR FIAT/IFTA
COVER AND SECTION PHOTOS BY AAD VAN DER VALK
PRINTED VERSION BY BEELD EN GELUID AT MULLERVISUAL COMMUNICATION 2013
# CONTENTS

**PREFACE**  
by Annemieke de Jong & Ingrid Veenstra  

**ACKNOWLEDGEMENTS**  
by Eva-Lis Green  

## 1. THE NETWORKED SOCIETY

1:1 THE NETWORKED SOCIETY  
by Nick Ceton  

## 2. LINKED METADATA

2:1 ON APPLES AND COCONUTS  
by Seth van Hooland  

2:2 GEMS: SEMANTICS FOR AUDIOVISUAL DUMMIES  
by Xavier Jacques-Jourion  

## 3. PRESERVATION METADATA

3:1 PRESERVATION METADATA  
by Karin Bredenberg  

3:2 AV PRESERVATION WORKFLOWS  
by Annemieke de Jong, Beth Delaney & Daniel Steinmeier  

3:3 HELP!: THE LEGAL DEPOSIT FOR ELECTRONIC DOCUMENTS IS HERE  
by Eva-Lis Green & Kaisa Unander
4. AUTOMATICALLY GENERATED METADATA

4:1 A ROSETTA STONE FOR IMAGE UNDERSTANDING
by Cees Snoek 97

4:2 MOOD BASED TAGGING
by Sam Davies 108

4:3 SPEECH TO TEXT AND SEMANTIC ANALYSIS
by Sarah-Haye Aziz, Lorenzo Vassalo & Francesco Veri 126

5. USER GENERATED METADATA

5:1 THE BBC WORLD SERVICE
by Yves Raimond 136

5:2 RTÉ TWITTER PROJECT
by Liam Wylie 146

6. EPILOGUE

6:1 CONCLUSIONS
by Brecht Declercq 156

OBSERVATIONS
by Ian Matzen and Jacqui Gupta 162

INDEX OF STILLS 175
IN THE DOMAIN of digital archives and digital archiving the concept of ‘metadata’ has become crucial. It's only with the help of metadata that archives can make their treasures accessible to their users, be it manually or increasingly – automatically. Metadata is capable of linking collections, forming a huge worldwide network of online sound and images. Ingesting, managing and preserving digital files, would be impossible without the controlling power of standardised metadata. More than ever, more than before, metadata is key.

It is for this reason that the Media Management Commission of FIAT-IFTA dedicated its 6th Seminar in the serial “Changing Sceneries, Changing Roles” entirely to the phenomenon of metadata and its new significance to access, collection management and preservation. More than 175 archivists, documentalists, preservation staff, educators and archive managers from all over the world, came to Hilversum on the 16th and the 17th of May 2013, where they were welcomed by Jan Müller, CEO of the
Netherlands Institute for Sound and Vision and president of FIAT/IFTA. The Seminar was divided in 4 sessions: Linked Metadata, Preservation Metadata, Automatically Generated Metadata and User Generated Metadata. For each session, moderator Thomas Prehn, introduced a keynote speaker and two case studies, presented by a broadcaster or a national archive. As in all previous MMC Seminars, every session was then evaluated by a variety of AV-practionairs in a discussion lead by Elena Brodie Kusa.

From the very first edition in 1998 on, it is a tradition to document the Seminar paper in a book. As did the Seminar, the 2013 Edition carries the title “Metadata as the Cornerstone of Digital Archiving”. It's opened by an article by Nick Ceton of primary seminar sponsor Ericsson, on the networked society and the role of AV-archives. The book then follows the seminar structure, having each of the 4 sessions form a chapter of its own. By way of an epilogue, a synthesis of the presented metadata concepts was developed by Brecht Declercq, member of the MMC. The book ends with a series of observations of the seminar presentations and the panel discussions, written by Ian Matzen and Jacqui Gupta.

With this book the Media Management Commission hopes to have delivered another important contribution to the current insights in digital audiovisual archiving, in this case metadata developments. The Institute for Sound and Vision is proud to have been given the opportunity to facilitate the 6th MMC Seminar, that functioned as the book's foundation.

Annemieke de Jong
Sr. Policy Advisor Digital Preservation

Ingrid Veenstra
Manager Selection & Registration

Netherlands Institute for Sound and Vision
Hilversum, July 2013
PRESENTATIONS FROM THE seminar are collated in this book, and I thank all the writers for their contributions, as well as contributors to the added papers. I also want to mention the seminar host Beeld en Geluid, Wendy Bekkers and Erica van den Bent for their administrative support, the Sponsors, and the moderators Elena Brodie and Thomas Prehn. Thanks to the Photographer Aad van der Valk and SVTBild for letting us use their photos in this book, and Axel Green for the book design and stills research.

And last but not least the MMC organisation Committee, Annemicke de Jong, Ingrid Veenstra, Jacqui Gupta, Xavier Jacques-Jourion, Elena Brodie, Vicky Plaine and Brecht Declercq.

Eva-Lis Green

Chairman, MMC Committee

Stockholm, July 2013
1. THE NETWORKED SOCIETY
Nick Ceton has been working in the field of media digitisation and access services since 1998. He has been involved in one of the first video on demand platforms in the Netherlands and worked as a consultant and project manager for the Nederlands Omroepproduktie Bedrijf to realise the Digital Archive of the Netherlands Institute of Sound and Vision. Nick was also involved in the development of the services for the ‘Images of the Future’ project.

Throughout the past few years Nick has been active as a business and product development manager within Technicolor mainly in the field of digital storage, access services and media distribution. One of his focus areas are the challenges around the longevity of digital assets. He holds an MSc degree in Electro-technical Engineering. In 2012 Technicolor Broadcast Services was acquired by Ericsson. Ericsson is the world leader in communication technology and services and with this acquisition also one of the leading independent play out service providers.
‘THE NETWORKED SOCIETY’ is Ericsson's vision of the Future. In this new world, everyone and everything will be connected everywhere, in real time. This will fundamentally change the way we innovate and collaborate, the way people will consume and interact with content and the way businesses are built and society is shaped. This will not only have impact on things such as education, government, telecommunications, broadcasting and archiving, but will also affect our daily lives.

Ericsson expects 50 billion devices around the world to be connected by 2020. That's more than 6 devices per person. It is not about the connections however, but rather about the impact these are having on our world.

In 2007 I visited the FIAT-IFTA conference in Lisbon. It seems not that long ago but 2007 was the year Google introduced ‘Street View’ and Apple its first iPhone which now both seem to have been around for ages and we can hardly imagine life without the features these, and other
products, offer nowadays. Back in Lisbon 2007, I had quite a hard time just trying to send an email from my phone over a GSM network. In April 2013 several Dutch mobile operators reported that almost 50 percent of the traffic on their mobile networks consisted of video streams or video files. With the upcoming LTE (4G) networks we expect that in 2015 over 90 percent of the traffic will be video in some form: user generated video, YouTube or streaming video from content owners, broadcasters or archives.

It is difficult to predict the future. But one thing will be sure: it will be different from today and probably different to what we expect. Who imagined that in 2013 it is already common to receive a message from your car on your phone, reminding you, you left the car unlocked and offering you the possibility to remotely lock it. Or you can switch on any device in your house from wherever you are in the world. And all that with a free app.

**CONTENT ANYPLACE, ANYWHERE, ANYTIME**

Connectivity offers new possibilities and new ways of consuming content as well. Linear television will not disappear, especially not for news, sports, live events, game shows et cetera, but the consumer will find more ways of watching content, wherever they are, at any time, on any device.

If you have missed the eight o' clock news why not catch up later that evening on your tablet or laptop. Why not watch an episode of your favorite sitcom on your smartphone, while commuting in the train. Hook up into an alternative live stream on your second screen app on your tablet, to view a different angle during a sports game or perhaps experience some additional footage that is not shown on the main screen. Both current devices and current mobile networks offer this capability.
At the same time the offering of video services grows exponentially. Not only the number of international available channels increases, but also the number of video portals and the amount of user generated content expands on a daily basis. Each minute, according to Google, more than 70 hours of video content is uploaded to YouTube. With this in mind, it will become harder to catch the viewers’ attention. But this is not the only game changer.

**CHANGING BUSINESS MODELS**

In the traditional value chain, a content producer, a rights holder, a broadcaster and a distributor all work together to deliver content to the consumer. The business model is clear and hasn’t changed very much over the last decade. You pay your monthly subscription fee to the distributor and maybe the broadcaster. The content rights holder is paid by the broadcaster and is ultimately responsible for paying the content producer. And so everybody gets their share in this traditional, simple, business model.

Nowadays in the increasingly connected world this business model is under pressure. Today you don’t need a cable subscription anymore to watch television or video content. Any broadband internet is sufficient to receive hundreds of channels from all over the world. The broadcaster can easily bypass the traditional distribution parties. However for the same reason the content producers and rights holders can easily bypass the broadcasters and deliver their content right in to the homes of the viewers. Services like Netflix and Google TV are already capable of offering this and are really shaking up the content industry.

We don’t have to look very far to find another industry that has already faced similar challenges to their traditional business model. In the music industry for example: in 2007 10 dollars would buy you a CD from your favorite singer or band. Today 10 dollars will give
you one month’s subscription to Spotify with access to more than 20 million songs. It is just an example of changing business landscapes our industry needs to adapt to.

**NEW OPPORTUNITIES**

These changing business models and ways of content consumption can be both a threat and a new opportunity for archives. Even though the volume of audio and video in the world increases on a daily basis, in the end it is still the quality of the content that counts. Fortunately archives all over the world are safeguarding and preserving the most valuable audio and video assets there are. This content ‘only’ needs to be found by the viewers.

Of course this is much more complex than it seems at first. What will be important is to integrate the content seamlessly into the daily lives of the consumers. Consumers need to be able to access content at the time, place and device they want it, plus it has to be just a click or swipe away and not hidden in an unfriendly user interface.

**THE ROLE OF ARCHIVES**

What is the possible role of archives in the Networked Society? Imagine a research student or scholar accessing a video database anywhere in the world or a distributor who wants to offer content to one of their niche customer groups, or a consumer that wants some in-depth information about a subject that’s stored in many archives around the world.

Today this may seem a virtual impossibility, how can content stored “somewhere” in the world be found at the click of a button? Well, by connecting these archives to the rest of the world; by partnering
with companies or organisations that will build new ways of content consumption, in education, research, entertainment or any other business. Archives need to find new business models as well, by opening up their assets to new types of usage. This will not be possible overnight, issues such as rights or search limits archives across the world.

**THE INFRASTRUCTURE CHALLENGES**

The ability to offer video or audio content in these kinds of future applications requires an intelligent and very flexible infrastructure from source to end-user. Video formats and bitrates need to adapt to changing devices and dynamic network circumstances. At the same time this infrastructure will have to be very cost effective and provide seamless integration with your digital archive.

As a first challenge, the content in your archive is likely to be stored in high resolution formats, such as MXF D10-50 or MPEG2 HD in an MXF wrapper. First, these file formats are large, varying from 30-50 Gigabyte per hour of content so are not particularly suited for distribution purposes. These growing archives with large files offer a challenge of their own. How to take care of the digital preservation of these assets and who, in the distribution chain will pay for the yearly costs of the digital storage that involves datacenters, connectivity and a lot of digital storage? Digital longevity has its challenges not only to preserve the files on the right technical carriers (and what is this right carrier anyway – disk, tape, optical?), but even on the level of the right file (codec and wrapper). Who guarantees the files of today can be read in 50 years, even if they are preserved consistently? The ancient Egyptians, for example, took care of the first by carving the hieroglyphs in stone that survived the centuries, however omitted the latter by not preserving the codec that explained what the hieroglyphs actually meant.
In addition to this format challenge, these formats are not formats used on end-user devices. So offering content to end users for any purpose, being streamed over the web, broadcast via a thematic broadcast channel or sent to any mobile device will need transcoding to a popular format. These formats change rapidly, the newest format for end-usage is MPEG-4 HVEC or H265 and will reduce the necessary bandwidth by 50 percent while retaining image quality. In addition, the new MPEG-Dash standard will enable changing bitrates to compensate for variations in the network. These new standards will enable the end-user to consume the content at any time, any place on any device but puts the archive to the challenge of “transcoding” the assets to the suitable format in a cost effective but efficient way.

The second big challenge is to enable the content to be found. Without proper metadata the essence doesn’t qualify as valuable content, so the usage of rich metadata, being skillfully added by AV-documentalists and archivists, (semi)automatically generated or even user generated may contribute to the traceability of content and in the end its usage. Joining forces as archives to collaborate in a universal virtual collection and offering a federated search, for example by using a controlled vocabulary, will mean a network of linked archives that in total offers more than the sum of the individual contributions.

**FOCUS ON YOUR CORE COMPETENCES**

In times of changing technology, changing business models and financial crisis it is very important to realise what value your organisation can add in the changing society. Find ways of offering your valuable assets in new ways and new business models to new ways of usage, together with new partners in new consortiums. This may not always be a financial gain but a business model can also add cultural value. Your assets, your brand, your
skills and resources can add unique parts to the concept of a linked archive. Free up your time, financial resources and attention by looking at other business models for your internal activities as well. In the synergy of connectivity, mobile networks and video, Ericsson adapts in this changing landscape by offering unique managed services to broadcasters, archives and content producers with services like digitising, storage, publishing and broadcasting. This enables our customers to focus on their new role in the Networked Society.

In the Ericsson Networked Society we say: When one person connects, their world changes. With everything connected our world changes. I'd like to add: When audio visual archives connect, society's view on the use of audiovisual content changes. If people can access audiovisual content that is stored all over the world in the various archives seamlessly only a click or swipe away, they will appreciate the valuable assets that are carefully preserved and with this create a new business model that will help archives to enter the networked society.
2. LINKED METADATA
Seth van Hooland holds the chair in Digital Information at the Information and Communication Science department of the Université Libre de Bruxelles (ULB), Belgium.

Within the current move towards a more structured and semantic environment for electronic resources, van Hooland wants to bridge the gap between domains (Document and Records Management, Open and Linked Data, Digital Humanities), practices (manual and automated metadata creation, data cleaning and enrichment) and communities (academia, industry, activists and practitioners) in order to make resources as accessible as possible in a sustainable manner.
AS A FOUR year old child, I developed an obsessive fascination for coconuts. I must have seen one on television in some context, and when I spotted them in our local supermarket I persistently begged my parents to have one. As a result of weeks of nagging, I finally managed to drag one home. It took me hours to figure out how I would open it. During the process I managed to lose the coconut water. Once fully opened, I was quite disappointed with the coconut flesh, which was probably all dried up after months of being shelved at the supermarket.

Over the last years, I perceived the same type of disappointment among collection holders and content providers who have been trying to get to grips with semantic web and linked data principles. The prospect of connecting resources from heterogeneous collections over the web in an automated manner obviously holds a tremendous potential. However, when confronted with the concrete output of large-scale research projects, con-
tent providers tend to be quite disappointed when gazing at the billions of RDF statements contained in triple stores, which can be exploited through an obscure query language.

Time has come to take a step back from a merely hype-driven vision of linked data. Linked data can offer big advantages for the dissemination of audio-visual archives, but let us spot the low-hanging fruits (the apples) which are currently within reach, and identify long term goals (coconuts) which might require substantial efforts and whose business value is rather uncertain. Before we go there, it is essential to understand why RDF, the data model underlying the linked data vision, has been developed and how you should see this particular way of modelling data in regards to other models which have been developed, and will continue to coexist over the next decades. The technological hype cycle leads us to believe that every new technology supplants the previous one, but in the field we must acknowledge their coexistence.

**WORLD’S SHORTEST INTRODUCTION TO DATA MODELLING**

Suppose someone handed you a stack of about 50 DVDs, and asked you to describe the collection. What would be the most intuitive and natural thing to do? Chances are high that you would take a sheet of paper or create a spreadsheet and identify columns in which you cluster the most important metadata of the DVDs (title, director, date, actors, etc). Figure 1 illustrates how we might develop this type of overview of the DVDs you need to catalogue.

During centuries, this practice underlay the creation of indexes and catalogues, and in a lot of contexts it is still the first step undertaken. It offers the big advantage that it is almost self-explanatory: all metadata are contained within columns next to one another. However, how long is this data model sustainable? A quick look at the metadata in Figure 1 gives an
overview of the limitations of the flat file approach. For example, the family name of the director is expressed in two different manners (“Jarmusch” and “Jarmush”). Imagine we don’t have four records in our flat file but four million. If we want to rectify errors in the misspelling of the name, we have to go through all of the four million records. On a computational level, search and retrieval is very inefficient with this approach, as again all four million records have to be checked to see whether they contain a specific value. Flat files also don’t offer the right tools to impose rules on how we encode values, resulting in inconsistencies in the way we encode dates, for instance.

For all of these reasons related to inconsistency and performance, databases were introduced. The relational model, developed at the end of the 1970s, has by far been the most successful for representing data, and

<table>
<thead>
<tr>
<th>DIRECTOR</th>
<th>TITLE</th>
<th>ACTORS</th>
<th>DATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jim Jarmusch</td>
<td>Dead Man</td>
<td>Johnny Depp, Gary Farmer</td>
<td>1995</td>
</tr>
<tr>
<td>Jim Jarmusch</td>
<td>Mystery Train</td>
<td>Steve Buscemi, Joe Strummer</td>
<td>End of the eighties</td>
</tr>
<tr>
<td>Jim Jarmusch</td>
<td>Stranger Than Paradise</td>
<td>John Lurie, Eszter Balint</td>
<td>'84</td>
</tr>
<tr>
<td>Et cetera</td>
<td>Et cetera</td>
<td>Et cetera</td>
<td>Et cetera</td>
</tr>
</tbody>
</table>

*Fig 1: Example of an overview of DVD catalogueing*
will continue to be used in the decades to come. The essential difference with the flat file approach is the organisation of our world in independent entities (“Director”, “Films” and “Actors”), which are characterised through their attributes (“ID”, “Name” and “Date of birth” in the case of the entity “Director”). Entities are connected through relations with one another. This data modelling approach offers the tremendous advantage of storing the data only once in regards to, for example, a director. Every time he or she makes a new movie, we don’t have to re-encode metadata in relation to the director (as we need to do with the flat file approach); we simply refer to the ID of the director when we add a new film. This approach ensures a lot more consistency obviously. Software built on top of this model, referred to as Relational Database Management Software (RDMS), also holds numerous possibilities to impose constraints on how data are encoded. This model has become the standard approach in how organisations create and manage their metadata.

Early collection management systems for the cultural heritage sector were also built on top of RDMS. From the very beginning, institutions thought about interoperability and how records could be exchanged between collection holders. There was (and there sometimes still is) a strong belief that acquiring the same software will provide the needed basis for interoperability. However, the customisation of these software tools, and their database back-ends, inherently resulted in different approaches regarding metadata encoding. This made the exchange of records between institutions, which might have been using exactly the same software, problematic. Interoperability and sharing data is really the point where people noticed drawbacks of the database approach. If I want another institution to have my database, it’s not just a matter of copy/pasting over some files and installing software on a server. Collection management databases that are used within an archives/curatorial environment tend to be quite complex, as they can be built up by hundreds of interconnected tables. First of all, you need to have the same version of the RDMS software, which might require a heavy
investment in a licence in the case of proprietary software, such as Oracle or Microsoft’s SQL Server. Secondly, understanding the relations between hundreds, or sometimes thousands, of tables can be challenging, especially if there is no explicit documentation on how the database was built and its evolution through time. Hopefully, this makes it clear that swapping out databases is not the way to go to share information between institutions.

Fast forward to the year 2000 – The Web became popular and accentuated the need to publish and share metadata in a transparent and cost-efficient manner. In this context, XML (eXtensible Markup Language) became a format of choice to import and export data both on and offline. What makes XML so special? Unlike the binary files of a database application, XML files are just marked up text which can be opened in any text processing application, such as Notepad.
Figure 3 gives a simple example of how metadata can be encoded in XML. The brackets identify the markup, in between which data can be stored. Here it is important to realise that the user has the possibility to define the markup elements, hence the adjective “extensible” in the name of XML. In this sense, XML is a meta markup language, as it gives you the tools to create your own markup language, specifically suited for an application domain. These elements are documented in what we call an XML schema. The big advantage of XML is its platform and application independence.

With its open and standardised format, XML offers us tremendous interoperability advantages in comparison to applications built on top of databases. But even if native XML based applications exist, they are not as efficient and easy to maintain as databases for the day-to-day management of data. More importantly for us, if you (or a computer) want to understand what
the values contained within the XML elements actually mean and how they interact with one another, you still need to access and process the XML schema in order to make sense out of the data.

We have come all the way with our example to come to this specific point. Both databases and XML offer wonderful possibilities to create and store structured metadata, but they both have the big drawback that you need to understand the schema describing the structure and interaction between the data. This is exactly where RDF (Resource Description Framework) comes in.

By adopting an extremely simple data model consisting of triples, data represented in RDF become in a sense “schemaless”. An RDF triple consists of a subject, a predicate and an object. This model allows one to express simple statements, such as Johnny Depp (subject) performed in (predicate) Dead Man (object). Figure 4 represents the domain of movies
which we have been documenting throughout our examples, but the meta-
data are now expressed in RDF triples. By simplifying the data model
to a maximum, all of the semantics are made explicit by the triple itself.
By doing so, there is no longer a need for a schema to interpret the data.
Within the world of databases and XML, only the data conform to the
rules defined in the schema may exist and be encoded in the database or
XML file. With RDF, you just make statements about facts you know, but
these statements might interact with statements made outside your inform-
ation system. This data model allows heterogeneous data to connect and
interact.

This connection of heterogeneous data relies on the use of unique iden-
tifiers for concepts. Therefore, linked data use URLs to express both sub-
jects (e.g. “Jim Jarmusch” can be identified through the URL describing
his entry in IMDB, predicates (e.g. “has made” can be expressed by using
the Dublin Core field creator, referred to as http://purl.org/dc/terms/
creator) and objects (e.g. the film “Dead Man” by the Wikipedia URL
URLs on the Web allows for other people to include the same URLs in
other RDF statements. For example, if I publish an RDF statement say-
ing that I live in New York, and I use the same URL to identify the object
“New York” and the predicate which expresses that I am based some-
where, we can automatically derive a new statement based on the triples
represented in Figure 4, describing the fact that Jim Jarmusch and I live in
the same city. Deriving this type of new information is the exciting aspect
of linked data. However, the fact that “anyone can say anything about
anything” can obviously also have serious drawbacks in regards to quality
assurance of data on the Web.
STILL SOUNDS COMPLEX? DON’T PANIC!

If this still sounds overly complex to you, don’t panic. Until recently, metadata practitioners lacked accessible methods and tools to experiment with linked data. On the website http://freeyourmetadata.org, you can start experimenting with freely available tools and services to evaluate the use of linked data on your own. The quickly evolving landscape of standards and technologies certainly continues to present challenges to non-technical domain experts, but it is important to point out the low-hanging fruit which are currently hanging in front of our noses.

Before you think about interconnecting your metadata with external data sources, you should first streamline and clean your metadata in order to enhance the linking process between heterogeneous data sources. This used to be a task that had to be performed by people with both domain and technological skills. Often, many thousands of records require similar operations. This is either a tedious manual task, or something that needs to be automated on a per-case basis. Luckily, the advent of Interactive Data Transformation tools (IDTs) allows for rapid and inexpensive operations on large amounts of data, even by domain experts that do not have in-depth technical skills. OpenRefine (http://code.google.com/p/google-refine) has recently gained a lot of popularity and is rapidly becoming the tool of choice to efficiently process and clean large amounts of data in a browser based interface. A screencast on the Free your Metadata website details how to make use of filters, facets and clustering algorithms to identify and act on metadata quality issues.

Once your metadata set has been cleaned, you can start experimenting with the reconciliation of your metadata with external knowledge bases, such as Freebase of DBpedia. The DERI research group has developed an RDF extension for OpenRefine, which allows users to add SPARQL endpoints of their own choice to the reconciliation process. This means that you can automatically check whether the content of a keyword type of
field from your collection can be matched to concepts from the Library of Congress Subject Headings (LCSH) and the Arts and Architecture Thesaurus (AAT), which are both available from a SPARQL endpoint.

If your metadata does not contain structured keywords, you might make use of the Named Entity Recognition (NER) extension that Ruben Verborgh of the Free Your Metadata project developed. In the context of audio-visual archives, you could apply NER techniques on the unstructured text which has been produced on the basis of Speech-to-Text techniques. The NER extension automatically identifies concepts, which are available in knowledge bases, inside unstructured text. By offering a URL to the identified concepts, the extracted terms are also disambiguated.

**CONCLUSION: SHORT, MID-TERM AND LONG TERM STRATEGIES**

The practical overview of data modelling should allow collection holders to have a clearer understanding of the possibilities and limitations of the four data models (flat files, databases, XML and RDF) which we have been using over the last decades to manage metadata, and which will continue to coexist in the years to come. As a collection holder, archivist or curator it is fundamental to understand on a conceptual level when to opt for what model, especially in the current linked data hype. Don’t let an academic or a consultant convince you to migrate your relational database for a RDF triple store; unless you are absolutely sure of its added-value (the academic nor the consultant won’t be there for the day to day management of your millions of records over the next decades). Over the years to come, we’ll probably see a gradual enrichment of “classic” metadata with concepts linked to knowledge bases and controlled vocabularies which are ingested in extra tables in your relational database.

This does not mean you cannot already start picking some low hanging fruits. Despite the intrinsic complexities of linked data techniques, it
should hopefully be clear by now that every institution interested in the linked data vision can start experimenting on its own. On a short-term, every institution can gain immediate benefit from the use of tools such as OpenRefine to clean its metadata. Freely available extensions exist to apply reconciliation and NER techniques on your existing metadata and to evaluate in a hands-on manner the potential benefit of creating links to outside knowledge on the Web.

The most important mid-term decision to make really relates to identifying your metadata core-business and outsourcing all fields which are already managed by external parties in a sustainable manner. If the Virtual Authority Files (VIAF) project maintains persistent and updated authority files for artist, corporate and geographic names, why should your institution spend valuable time encoding the same information? Therefore, sum up all the metadata fields that your institution is currently encoding and try to identify the authoritative sources to which you could point instead of developing the metadata yourself. The big unknown factor here is obviously how sustainable projects such as VIAF are.

On a long-term level, an institution should reflect on the potential added value of becoming an authoritative source for the metadata which reflect its core business, by publishing these metadata as linked data. By doing so, your institution offers others the opportunity to reuse your metadata in a persistent manner. Taking this step requires an important institutional engagement (issuing and managing persistent identifiers and the attached metadata is a costly affair), and should only be taken if the upper management is convinced of the long-term value (overall reputation, increased presence on the Web, etc.) of this investment.
Xavier is Head of Archives at the Radio-télévision belge de la Communauté française (RTBF), the Belgian French-speaking public broadcaster. He started his career as a sound engineer in Radio in 1997, then turned to project management in 2007 before joining the Archives in 2009.

Over the years, he has led various projects at the interface between broadcasting technology, information technology and content production. Aside from the daily tasks of archiving and retrieval for both TV and Radio, the RTBF Archives department is in charge of charting the future in terms of metadata for the whole company. Xavier is a member of the FIAT-IFTA Media Management Commission.
AUDIOVISUAL ARCHIVES THE world over are at the crossroad of antagonistic trends: on the one hand the ineluctable laws of demography make sure that our longest serving archivists are leaving the workforce for a well deserved retirement, while on the other hand current market trends tend to show that we are all going to archive an ever growing amount of material with the same resources. Tomorrow is all about “Big Data” and the archiving of more complex audiovisual products dispatched over the air or over the Internet.

Our deep archives are by nature imperfect, mostly because the standards of indexing and documentation have evolved over some 60 years of television production and about 90 years of radio, along with the cataloguing tools used over time, from the card file cabinets to modern databases. Audiovisual archivists who have worked their entire career within television know every nook and cranny within the archive and are able
to find the unique and exclusive footage that our end-users are looking for. Footage that will enable the creation of richer, more entertaining and overall more interesting content for our spectators, no matter what the actual delivery medium turns out to be. Every time such an expert retires, we lose irreplaceable knowledge of the implicit structure of our archives and swathes of our collections disappear in darkness, hidden from view.

At the opposite end of the scale, we live in a world of “Big Data”, where every bit of information produced is stored, documented and expected to be retrievable instantaneously. Archive material is used more and more in new productions, re-used in order to create a wealth of context at a lower total cost of production. Audiovisual production is becoming ubiquitous as it is increasingly published on the internet. With the cost of storing digital information going down, more and more content will enter the archive and it is at this point in time difficult to predict if we will have armies of archivists at our disposal to properly index all this content.

Based on our observation of these two trends, at RTBF we have identified the field of semantics as a strategic direction, which could help us resolve the issue at hand. This paper aims to provide a simple introduction to the field of semantics and explain how the technology was used in building a working proof of concept application in the first half of 2012.

SEMANTICS

Much has already been said about semantics as a research field, and interesting applications keep popping up every day in various aspects of information technology (IT). Semantic technology is part of the various layers of tools that together form what is commonly known as the semantic web.

Interestingly enough, the vision that led to the current and latest evolution of the World Wide Web and the underlying technology isn’t new and dates back to the initial memo that Tim Berners-Lee sent to his boss at
CERN in March 1989. This important proposal, laying the foundations of the web, was commented as “Vague but exciting” at the time.

But what is semantic technology, and more importantly what does it do? In a few words, it enables computers to make more sense of the data they store and manage than they currently do and, based on this, provide the users with better, more accurate search results for instance.

**HOW DOES IT WORK?**

Semantic technology introduces a paradigm change in the way information is stored and managed. Classic databases as we know them are nothing more than turbocharged spreadsheet tables, where the data is stored
in rows and columns, and only the human user can make sense of the column title, thus barring the computer from extracting any meaning from the underlying data.

Instead of rows and columns, semantic engines store the data as triples: subject, predicate and object. Information is thereby qualified and links are built between the various data points. As an example, please refer to Fig. 2 for a triple representation of the sentence “The sky is blue”.

The different predicates used within a data collection form a system, which can be used as a way to structure knowledge. This system is called an ontology. Consider for a moment the example of genealogy as a system, and try to represent typical family links as triples. This fairly simple example easily demonstrates the power of the semantic treatment of the information.

Based on those elements, computers that were up to now used to manage
mostly raw data, can be used to manage information (in the form of qualified data points), and up to a point also enable us to manage knowledge (in the form of an ontology).

The data triples, which are also known as graphs (or graphical representations of information) also facilitate the interoperability between different information servers. Based on W3C standards, data owners are able to publish both the information contained in their systems and the way the data is organised, enabling what is usually known as Linked Data. Many content owners have chosen to completely open their databases to the outside world, allowing other databases to link data to theirs, thereby accumulating knowledge on a much larger scale. This is known as Linked Open Data (LOD), in which instead of copying data from another database, data points are linked together across multiple databases spanning the entire breadth of human knowledge. Practical examples of LOD are dbPedia (the semantic version of Wikipedia), MusicBrainz (the musical database), GeoNames (a reference for places with geographical coordinates).

Linked Open Data is a treasure trove of information waiting to be used, allowing the archive to enrich its metadata collection by putting its resources in perspective against the backdrop of multiple and rich sources of information.

THE GEMS PROJECT: FROM SEARCH TO FIND

Starting from the business issues described in the opening of this article, and based upon learnings from its participation in the MediaMap project, RTBF identified semantic technology as a strategic field of research as early as 2009.

Towards the end of 2011, the GEMS project was defined in order to extract emerging technologies from R&D and work on applying those on real-world data and problems. We decided to focus on technologies that were
available at the time, although at varying stages of development and maturity, and work with independent partners towards the creation of an innovation ecosystem where the best technicians would assemble a brand new solution.

Together with our partners Perfect Memory and Memnon, we built a working proof of concept for a semantic-based multimedia browser interface, using raw extracts from our media databases and the following technologies:

- A semantic engine was used to assemble the knowledge previously spread across multiple databases,
- Interoperability wickets were designed to import data from our legacy databases,
- Linked Open Data sources were imported into the system to enrich our data,
- Speech-to-text was used to automatically generate metadata, although the results of this was only used as an indexing field, hidden from the users,
- Natural Language Processing was used to detect named entities in a natural, free-flowing text and match those against our reference lists of names, places and concepts,
- And finally an ergonomically designed human user interface based on HTML5 standards.

The user interface, built using the Transmedia Player from Perfect Memory, was articulated in three parts: search results, media player and graph navigation. Search results are presented in a very rich way, with thumbnails, element
title and a summary of the item, along with the series titles, broadcast channel and date, and finally the item duration. Clicking on an item would take the user to the second tab for media playout.

The Media player screen presents the user with a wealth of metadata, displayed alongside the video or audio content as it plays out:

- Human-encoded keywords, resulting from our archiving processes are displayed prominently in the upper-right corner of the screen next to the video player,

- The bottom half of the screen is used to display machine-detected keywords along the media timeline, allowing for a very efficient navigation inside the media items,
The left part of the screen displays summary metadata for the sequence currently playing and the overall show that the sequence is a part of,

- Thumbnails are extracted from Linked Open Data, for instance showing a portrait for each name in our reference list,

- Clicking on a tag or keyword can display the relevant Linked Open Data page,

- Alternatively, clicking on a tag or keyword can also take the user to the Graph Navigation screen.

Fig 4: GEMS media playout
The Graph navigation screen displays the actual data triples relevant to the item chosen by the user, displaying media items for which the tag is the main subject, items in which the tag is mentioned, or concepts/tag that are related to the chosen tag, which appears in the center of the graph. Clicking on another tag or media generates a new graph, centered on the new item or tag.

Graph navigation is very intuitive and creates a space in which serendipity can happen, allowing a researcher to navigate across concepts and discover links between existing data points that he or she might not have been aware of.

Early user testing seemed to indicate a very short adoption curve and certainly generated a lot of interest for the software. Interestingly enough many users quickly forgot that they started off looking for illustrative images that
could be reused in a new production as they focused on the links between different (and sometimes seemingly unrelated) points of information.

CONCLUSION

Our conclusion from the proof of concept experiment is very positive: indeed we have proved that semantic technology combined with automatic metadata generation and linked open data can bring incontestable added value to the way we access and navigate through rich media collections. Far from actually replacing the old expert archivist leaving on retirement, the semantic multimedia browser offers an entirely new experience that could enable serendipitous results otherwise impossible to obtain.

Furthermore, the market of semantics-based software development and solutions is beginning to reach some maturity as more and more players switch from pure R&D to having systems actually used or put to their paces in production.

Numerous questions remain open for debate and still need an adequate answer before using such a system in the context of an audiovisual archive, but we are confident – now more than ever – that semantics are the way to go and that this path will deliver results in the coming months.
3. PRESERVATION METADATA
Karin Bredenberg holds a position as IT architect at the National Archives in Sweden (SNA). She graduated in Computer Engineering at the Royal Institute of Technology in Stockholm in 2006. Bredenberg mainly works with Swedish adaptations of international archival metadata standards. She worked with archivists from around the world on the development of EAC-CPF, a standard for encoding archival authority records.

Currently she serves as member of the PREMIS Editorial Committee (Preservation Metadata: Implementation Strategies) and the Society of American Archivists; she is part of the Schema Development Team and of the Technical Subcommittee on EAC. Since 2011 Bredenberg has been project manager for part of the E-archiving specifications for the Swedish project E-archiving and registration eARD.
SINCE WINNING THE 2005 Digital Preservation Award, the PREMIS Data Dictionary for Preservation Metadata has become the international standard for preservation metadata for digital materials. Developed by an international team of experts, PREMIS is implemented in digital preservation projects around the world, and support for PREMIS is incorporated into a number of commercial and open-source digital preservation tools and systems. It is a comprehensive guide to core metadata to support the digital preservation process.

PREMIS is based on an underlying data that defines five key entities – Objects, Events, Rights, Agents and Intellectual Entities – associated with digital preservation and is grounded in the concepts of the OAIS Information Model. To promote applicability in as wide a range of contexts as possible, the Data Dictionary is neutral in terms of the preservation strategies or encoding used for implementation, although it does provide the PRE-
MIS XML schema as one alternative.

This paper addresses the importance of preservation metadata and how it supports many of the functions of preservation repositories. It explores the PREMIS Data Model, key features of the PREMIS Data Dictionary and how it fits into OAIS. As PREMIS has matured, both open-source and commercial tools and services have been built to support it, and the paper discusses some of the options.

Since PREMIS is intended to address all format types, it doesn’t specifically address audiovisual materials. This paper will cover some of the implementation issues that audiovisual archives face given the complexity of these materials. It will also summarise work being conducted in two working groups of the PREMIS Editorial Committee: the PREMIS OWL ontology, which is compatible with Linked Data and the Semantic Web, and explorations of PREMIS compliance in regards to common usage and the ability to share digital objects and their metadata across repositories.

**THE CHALLENGE OF DIGITAL PRESERVATION**

Saving digital objects of all kinds and formats has posed major challenges for all institutions regardless of geographic location or user community. If these digital objects are someday going to replace the analog, they need to remain accessible next week, a decade from now, and so on.

How the digital world will change is largely unknown as is our knowledge and experience regarding validating the longevity of digital objects. The Open Archival Information System (OAIS)\(^1\) reference model provides us with a basis for the digital archive to help us to secure access in the

---

\(^1\) [http://public.ccsds.org/publications/archive/650x0m2.pdf](http://public.ccsds.org/publications/archive/650x0m2.pdf)
future to digital objects and also gives us guidance on how to store “Digital Preservation Metadata”. The OAIS model can be seen as a description of the functions and entities in a digital archive, allowing us to store the digital objects in a uniform way. The OAIS model also states that “Preservation Description Information” is needed to ensure that the archive has sufficient information about the digital object to allow for long-term preservation. The PREMIS Data Dictionary was developed to build on OAIS, providing the specific information that is needed.

**BASIC PREMIS CONCEPTS**

The PREMIS Data Dictionary is a listing, with detailed specifications of the information you need to know for preserving digital objects, using the OAIS Information model as a starting point. The acronym PREMIS is commonly used and it stands for PREservation Metadata: Implementation Strategies.

The PREMIS website is hosted by the Library of Congress, and the standard itself is maintained by the PREMIS Editorial Committee (EC). The EC consist of members from different types of institutions from all over the world. The EC members are responsible for the directions and the priorities of PREMIS, handling change requests, revising both the DD and the accompanying XML-schema, and promoting the standard in their respective countries. The PREMIS EC also conducts events for users to better understand the Data Dictionary and provides a forum for users to share issues and experiences. These are the PREMIS Tutorials and PREMIS Implementation Fairs. All these events are announced through the PREMIS website.

---

2. Metadata that is essential to ensure long-term accessibility of digital resources. Where metadata means data about data
PREMIS defines units of preservation metadata as representation, context, provenance and fixity information in OAIS terms. This information is expressed as “semantic units”, a term referring to discrete pieces of information that an archiving system needs to know for long-term preservation.

**DATA MODEL OF PREMIS**

The original PREMIS Working Group designed a data model consisting of 5 interacting entities to organise the Data Dictionary and to understand the entities involved in the digital preservation process. The data model provides a useful framework for distinguishing applicability of se-
mantic units across different types of Entities and different types of Objects. However, it is not a formal entity-relationship model, i.e. it alone is not sufficient to design databases.

The data model includes:

- Entities: “things” relevant to digital preservation that are described by preservation metadata (Intellectual Entities; Objects; Events; Rights; Agents)

- Properties of Entities (semantic units)

- Relationships between Entities

For each entity (except for the Intellectual Entity) the Data Dictionary provides semantic units that describe the entity’s preservation metadata. Intellectual Entities are out of scope in version 2.2 of the DD (except for an identifier), but this will change in version 3.0. As a result of user requests, the EC decided to make Intellectual Entities another level of Objects.

**THE SCOPE OF PREMIS**

The PREMIS DD serves as a common data model for organising and thinking about preservation metadata. It is also considered to be a standard that is implementable and technically neutral, providing core metadata that can be used for exchanging information regarding preservation metadata between repositories. Thus the PREMIS DD is not an out-of-the-box solution that satisfies all preservation needs, since it does not provide all needed metadata for rights management and lifecycle management of ob-
jects outside the digital repository, nor does it provide for format-specific technical metadata (although it provides an extension mechanism for it).

**INFORMATION PACKAGES**

When a digital object has been placed in archival storage an archive often uses a standard for creating a package that stores information about the structure of the digital object, who is the owner of the package, which files are in the package, markers of authenticity, relationships and dependencies among the files, and technical information like digital preservation information. One common standard for this is the “Metadata Encoding and Transmission Standard” known by its acronym METS. The METS standard is also based on the OAIS model and therefore is well suited for the purpose of creating all different kinds of “Information Packages” that exist in the model. METS is a standard that can accommodate different ways of creating a package; when you use the standard it is beneficial to create a profile describing the structure of the package, the metadata standards used, and examples of applying the profile. These profiles can be registered at the METS Board, allowing for access to that profile for users all over the world. METS specifies PREMIS as its endorsed standard for providing preservation metadata. Guides for using the two standards together and for making decisions concerning elements that are redundant between the two standards have been developed and are available through the PREMIS website. There are other container formats such as MXF or MPEG21, but these tend to be very complex and not as accessible.

---

ADDITIONAL METADATA IN PREMIS

In PREMIS additional metadata that is not covered by PREMIS can be inserted in extension elements. These can be local elements or elements from another standardised metadata scheme. For example format-specific technical metadata regarding a still image describing how the image was created can be included with the help of the standard Metadata for Digital Still Images in XML (MIX)\(^4\). This includes metadata elements such as bit depth, colour space, etc. For audio and video there are some schemas available to describe for example format-specific technical information such as physical format, speed, encoding, etc. but they are fairly new. One reason it has taken so long for these standards to be developed is there hasn’t been as much of a need until after the beginning of the 21st century, when audio and video digital formats became increasingly used. A lot of them are created by converting analogue material to digital objects. Previously, for the archival world the most common digital object until now was a database often delivered as text files, but now archives also have to deliver audio and video objects. Regardless of how the audio or video digital object is created, born-digital or converted, it still needs to have preservation metadata so that these types of digital objects can be useable in the future.

AUDIo AND VIDEO METADATA

Audio and video metadata standards have been slow to develop. For the most part the lack of standards can be explained by the complexities of

\(^4\) http://www.loc.gov/standards/mix/
the audio and video file formats themselves, which require expertise to develop and implement a metadata standard and require complex relationships between multiple file format types. Often programs capable of creating audio and/or video can export some metadata in XML-format. We have to ask ourselves in our preservation environment if this is sufficient or do we need more information? The answer found to the question (which perhaps is a hard one to answer) needs to be evaluated and needs to be the basis for a preservation plan regarding audio and video.

AES

The “Audio Engineering Society” (AES) in 2011 released two standards for audio metadata, one with core audio metadata and the other for audio object structures for preservation and restoration. Both of these audio standards are accompanied with their own XML-schema. So far there are a small amount of use cases to be found and they are mainly presentations about the standard with an educational orientation.

AUDIOMD

One simple standard for audio metadata is audioMD that is hosted at the Library of Congress. The standard is a result of work that started in 2000 and the version now published was an improvement made during 2009 and 2010. The schema is developed with simplicity in mind with

information needed by cultural heritage institutions (libraries, museums, archives) for their designated communities. The use of one of the earlier versions of the schema can be found in one registered METS profile.

**VIDEOMD**

For video metadata the simple standard is videoMD⁹. The standard has the same host and similar features to audioMD. Both audioMD and videoMD are attractive to the archives that are interested in using simpler schemas, although the AES schemas may provide more detailed metadata.

**SMPTE**

The work of the “Society of Motion Picture and Television Engineers” (SMPTE¹⁰) has resulted in the standard “Material Exchange Format” (MXF¹¹) and an accompanying metadata dictionary. The metadata dictionary consists of 1700 elements and it is regularly updated, and deletions are not permitted. The standard is a container or wrapper format which supports a number of different streams of coded “essence”, encoded with any of a variety of codecs, together with a metadata wrapper which describes the material contained within the MXF file.

---

¹⁰. https://www.smpte.org/
¹¹. The standard document can be bought by ordering “SMPTE 377M: The MXF File Format Specification (the overall master document)” on the homepage of SMPTE.
EBUCORE AND PBCORE

Both The European Broadcasting Union (EBU\textsuperscript{12}) standard EBUCore and the Corporation for Public Broadcasting standard PBCore\textsuperscript{13} was originally created as an enhancement of Dublin Core\textsuperscript{14}. EBUCore have now been adopted and published as AES core audio metadata. It provides a limited set of technical metadata and is mostly focused on descriptive information, but is worth mentioning because they cover some technical preservation related elements.

FAGDI

Federal Agencies Digitisation Guidelines Initiative (FAGDI\textsuperscript{15}) is a collaboration between the National Archives and Records Administration (NARA) and other federal agencies in the US. The work is based on developing standards for digitised video, working with AudioVisual Preservation Solutions\textsuperscript{16} on tools. Two tools can be found as open source, MediaInfo\textsuperscript{17} and AVI MetaEdit\textsuperscript{18}. MediaInfo is a convenient unified display of the most relevant technical and tag data for video and audio files. AVI MetaEdit, inserts essential metadata into selected video files such as information about how the file was made. An XML-schema for video technical metadata called reVTMD\textsuperscript{19} is also under development. The XML-schema is

\begin{footnotes}
\item http://tech.ebu.ch/lang/en/MetadataEbuCore
\item http://pbcore.org/index.php
\item http://dublincore.org/
\item http://www.digitizationguidelines.gov/
\item http://www.avpreserve.com
\item http://mediainfo.sourceforge.net/en
\item https://github.com/usnationalarchives/AVI-MetaEdit
\item http://www.archives.gov/preservation/products/reVTMD.xsd
\end{footnotes}
a structure to organise important technical information, such as how big the image is or how fast it should be playing, in a reusable XML format. reVTMD is now incorporated into MediaInfo, which is mentioned above. This new functionality allows the user to view and export metadata from a wide variety of video files into the reVTMD organisational structure.

MPEG-7 AND MPEG-21

The Moving Picture Experts Group (MPEG\textsuperscript{20}) is a working group of ISO/IEC with the mission to develop standards for coded representation of digital audio and video and related data. MPEG-7\textsuperscript{21} is a suite of standards for description and search of audio, visual and multimedia content and MPEG-21\textsuperscript{22} is a suite of standards that define a normative open framework for end-to-end multimedia creation, delivery and consumption. MPEG is to be seen as a digital film and TV format which means that the digital object is of the format MPEG.

TOOLS

To support the use of PREMIS some repositories have made tools available to the public. One of the tools are “PREMIS in METS toolbox” also known as PiM\textsuperscript{23} created by the Florida Center for Library Automation (FCLA\textsuperscript{24}). The tool actually consists of three different modules that are

\begin{itemize}
\item\textsuperscript{20} http://mpeg.chiariglione.org/
\item\textsuperscript{21} http://mpeg.chiariglione.org/standards/mpeg-7
\item\textsuperscript{22} http://mpeg.chiariglione.org/standards/mpeg-21
\item\textsuperscript{23} http://pim.fcla.edu/
\item\textsuperscript{24} http://fclaweb.fcla.edu/
\end{itemize}
used together as a set of open-source tools\textsuperscript{25}. The tool uses the PREMIS in METS guidelines and especially supports PREMIS implemented with the use of METS as a container format. It also extracts metadata from the objects. Other open source tools available are the DAITSS Digital Preservation Repository Software\textsuperscript{26} and Archivematica\textsuperscript{27}.

THE CONTROLLED VOCABULARIES DATABASE

The controlled vocabularies database\textsuperscript{28} can also be seen as a tool. This tool is provided by the Library of Congress to make available controlled vocabulary values for standards that it maintains. These controlled lists are represented using SKOS\textsuperscript{29} and MADS/RDF\textsuperscript{30} as well as alternative syntaxes, so are represented as Linked Data. For PREMIS there are vocabularies for the semantic units preservationLevelRole, cryptographicHashAlgorithm and eventType available. Additional PREMIS controlled lists will be made available when the PREMIS OWL\textsuperscript{31} ontology is released.

PREMIS OWL ONTOLOGY

LINKED DATA

A technique often used now is “Linked Data”\textsuperscript{32} which is a feature of the

\begin{itemize}
\item \textsuperscript{25} http://pimtoolbox.sourceforge.net
\item \textsuperscript{26} http://daitss.fcla.edu/
\item \textsuperscript{27} https://www.archivematica.org/wiki/Main_Page
\item \textsuperscript{28} http://id.loc.gov
\item \textsuperscript{29} http://www.w3.org/2004/02/skos/
\item \textsuperscript{30} http://www.loc.gov/standards/mads/rdf/
\item \textsuperscript{31} http://www.w3.org/TR/owl2-overview/
\end{itemize}
“Semantic Web”\(^{33}\) where links are made between resources. These links go beyond hypertext links (i.e. between web pages) and instead allow for links between any kind of object or concept. In Wikipedia Linked Data is defined as “a term used to describe a method of exposing, sharing, and connecting data via dereference able URI:s on the Web”. One of the benefits is that users can use links to find similar resources and aggregate results. The next step is to use “Linked Open Data” where instead of embedding data (in text, in tables, etc), you can simply share the data on-line. The resources are identified by a URI\(^{34}\) (e.g. a PREMIS description, controlled vocabulary terms). The data can be linked internally or externally since the resources are uniquely identified.

**THE ONTOLOGY**

The need for a PREMIS OWL ontology was raised by the Archipel project\(^{35}\) to be able to use PREMIS in an Open Archives Initiative Object Reuse and Exchange (OAI-ORE\(^{36}\)) solution. An OAI-ORE can simply be described as description and or exchange of Web resources. With the PREMIS ontology the data model and data dictionary may be used within the context of the Resource Description Framework (RDF), where the result is a serialisation of preservation metadata as a data management function in a preservation repository. At the same time in making PREMIS available in a different format it is important to remain as close as possible to the Data Dictionary’s clearly defined semantic units. The plan also is to interoperate with other preservation Linked Data efforts such as Unified Digital Formats Registry\(^{37}\).
(UDFR). as well as with the PREMIS controlled vocabularies at http://id.loc.gov.

**IMPLEMENTATION ISSUES**

When implementing a standard there are always issues about how it is used. In using PREMIS, as with other standards, conformance issues need to be considered to determine whether metadata describing digital objects actually conform to the specifications of the standard. The EC has defined rules for determining whether your implementation is conformant. The conformance can be seen on different levels, whether you use selected semantic units and are conformant to these individually or whether you implement the whole data dictionary, staying conformant to the whole. In being conformant you also abide by the rules of obligation, i.e. mandatory vs. optional. There are other technical issues, and implementation choices need to be made. For instance, will your system store XML-files with all information? Do you already have a system describing the agents associated with the repository and can you point to that system? You also need to decide how much information you want to keep regarding events, for instance, which event types should be stored and whether information about every event be recorded?

**CONCLUSION**

The complexity of audio and video objects demands the use of technical and structural metadata and to accomplish this both IT and the archives need to work together.

When digital objects, including audio and video objects are stored for the long-term many decisions are needed regarding preservation. Some common questions are:
• Why do we store it?
• How do we store it?
• For whom do we store it?

Answering these questions may involve multiple decisions. One decision will be whether we will store the audio/video object as it is with accompanying metadata or if the object will be stored with the help of a container format. Regardless of the choice regarding how we store the digital objects, preservation metadata is important, and the widely implemented standard PREMIS provides a solution. Creating and storing preservation metadata is essential, since digital objects change over time and it is necessary to employ digital preservation strategies, such as format migration. PREMIS is a key piece of the long-term digital preservation infrastructure.
Annemieke de Jong (Sound and Vision) is responsible for setting out the strategic policies for digital AV-collection management and AV-preservation. De Jong is also involved in the development and international dissemination of knowledge and expertise on digital AV-archiving and has several publications to her name.

Beth Delaney has been defining collection management system needs, implementing metadata standards and developing policy in audiovisual archives. As a consultant, she specialises in digital preservation processes, preservation related business requirements and preservation metadata standards.

Daniel Steinmeier is a technical specialist in the application management department at Sound and Vision. Metadata has always played an important role in his work and he has worked for years with metadata standards such as IEEE-LOM and protocols for metadata exchange such as OAI-PMH.
SINCE ITS ESTABLISHMENT in 1997, the Netherlands Institute for Sound and Vision has always served two functions. Firstly it is the central production archive for all public broadcasters (20) in the Netherlands. In this role the archive is responsible for storing and providing access to broadcasted television and radio programs. At the same time Sound and Vision functions as the Dutch national audio-visual archive. Thus, its collection also includes Dutch cultural heritage audio-visual material, including scientific material, documentary film collections, photos and objects, amateur film and AV collections from businesses and social organisations. In addition, Sound and Vision plays a central coordinating role in the Dutch broadcasting and AV cultural heritage landscape, gathering and disseminating knowledge in the digital storage and exhibition domains. Currently Sound and Vision manages a collection that includes more than 800,000 hours of radio, television and film programming. Yearly, thou-
sands of hours of digital born material, created in the broadcast domain, is added to the archive.

During the recent past, Sound and Vision has become best known for its analogue digitisation work. The huge project “Images for the Future” has seen the creation of digital files for more than half of its analogue film, video and audio collection. Another important development was the establishment of a direct IT interface between the archive and the digital broadcast production systems. Since 2007, all the broadcasted radio and TV programming as well as its related metadata has been ingested into the archive via an automated digital workflow system. In Sound and Vision’s depot, approximately 6 petabytes of digital born and digitised material is stored in its digital repository system. Everyday, sound and image fragments from the repository are delivered to hundreds of users in the professional broadcast domain as well as at home, in educational institutions, in businesses and as visitors to Sound and Vision’s museum, The Experience.

**PRESERVATION**

Now that Sound and Vision has digitised a great deal of its collections and thousands of new digital files are ingested daily, it faces an important new challenge: the management of a fast growing digital storage repository and the increasingly related complexity that brings with it. The chance of failure that the storage of digital material brings with it is great; it is necessary to maintain firm control over the lifecycle of digital files. In order to achieve long-term preservation, more processes, procedures and metadata need to be incorporated into the system. The role and responsibility of all parties, producers, archivists and users, needs clear definition.

The primary question is: how do we ensure that stored digital material stays accessible for the users? The different formats, size and location of digital files, delivered daily to a larger and larger client base is huge and
growing still. How do we make sure that we are able to deliver up-to-date formats that the Designated Communities can actually use? When the first IT integration with the broadcast production environment was designed, long-term preservation was not yet a major consideration. Thus, during the initial digital ingest and storage infrastructure design the processes needed to ensure long-term preservation were not explicitly incorporated. Yet, in both roles, as a broadcast production archive and as a national AV heritage coordinator, guarantees of trustworthiness and long-term preservation must be met.

Sound and Vision has taken this new challenge head on and has made it its most important goal for the coming period. Sound and Vision wants to become a “trustworthy digital repository” for Dutch audio-visual cultural heritage collections. Whether it is public radio or television programs, or other audio-visual material, it must be safely stored and made permanently available for whoever wants to use it. In 2012, this new strategic goal has led to the establishment of a project wherein the requirements for a trusted AV archive repository were to be defined. What was to be delivered was a set of normative policy documents that could guide the establishment of an OAIS compliant audio-visual archive environment. Its primary reference framework would be the “Quality Requirements Digital Archive Sound and Vision”. In addition to a set of policy documents outlining OAIS compliant organisational and administrative requirements, the central driver of the project was the development of Digital Object Management requirements. This consisted of the development of a workflow model for ingest, storage and dissemination of digital files and metadata. Data description would be enhanced through the development of a preservation metadata dictionary, composed of technical, administrative and provenance attributes. Together, the workflow and preservation metadata dictionary form the Information Model for Sound and Vision’s digital archive. To date, all the project documents have been completed.
OAIS

The Information Model’s workflow and metadata development was mostly inspired by the OAIS reference model. The ISO-standard OAIS\(^1\) defines the functions that ensure “trusted” long-term preservation and access to digital information objects. Establishing a common language to describe the objects and processes, it offers a conceptual approach upon which actual business processes can be defined, allowing the archive to fulfil the preservation goals a TDR promises. Its central theme is the guarantee, offered to those whose material is deposited in such a repository, to long-term accessibility. The model provides guidelines for defining and formalising processes and data throughout the archive chain: from ingest, to storage to dissemination. By tracking and registering certain defined events in the lifecycle of each individual digital object in “preservation metadata”, the authenticity of the ingested object can be demonstrated, and thus the basic requirement of ‘trustworthiness’ fulfilled. The Archive or repository is able to demonstrate its responsibility to its depositors as well as its users.

OAIS is a worldwide reference model that to date has seen wide take-up primarily in digital libraries and traditional archives. This can also be said for PREMIS\(^2\), the most important preservation metadata standard. The applicability and implementation of this standard in the media archive domain, where the emphasis has always been on access and re-use – is still scarce. How to manage and preserve fast growing volumes of digital material, in a rational and responsible way is really a question that increasingly confronts broadcast and other large audio-visual collection holders.

\(^{1}\)ISO 14721:2003. The Open Archival Information System Reference Model

\(^{2}\)http://www.loc.gov/standards/premis/
The Information Model Digital Archive Sound and Vision 1.0 formulates Sound and Vision’s first, normative answer to this question. In the model, OAIS and PREMIS based processes and metadata have been modified to fit the specific situations and needs of AV files managed in a dynamic production environment. Its development particularly reflects the domains in which Sound and Vision operates, as a production archive for Dutch public broadcasters (NPO), as a national AV cultural heritage repository and as an online service provider of AV content to various user groups.

**AUTHENTICITY AND INTEGRITY**

To date, Sound and Vision could not provide guarantees concerning two essential aspects found in a TDR: the integrity and authenticity of the objects in the repository. Upon what do such guarantees lie? In general an OAIS compliant repository must be able to demonstrate the integrity and authenticity of digital files/objects held in the repository through some sort of audit trail. Such a trail is created over time and documents preservation events that have occurred throughout a defined preservation workflow.

For example, demonstrating the integrity of a file requires documenting that data streams have not been corrupted during transport or while in storage. Validation and error checking mechanisms allow the repository to track and verify that no corruption has occurred. Fixity information demonstrates file/object integrity over time. The degree of authenticity is judged on the basis of evidence.\(^3\) Demonstrating the authenticity of a file/object over time requires documenting its provenance; that is the

---

creation, chain of custody and change history over time. Thus, two essential things are required: defined “preservation” business processes that ensure that “preservation” related events take place; and a mechanism by which an audit trail can be generated and maintained, allowing the archive to demonstrate the outcomes of these events. The TDR project set out to define these processes as well as a robust set of technical and preservation metadata, wherein essential lifecycle management information could be managed.

---

INFORMATION PACKAGES

Different types of digital objects exist in an OAIS compliant repository; the nature of these objects depends on where in the process the object exists. These “packages” are formed during defined workflow processes. The Producer/Donor submits files according to a Submission Agreement that consists of a set of terms between a producer and archive over the delivery and preservation of digital material. The files are packaged and sent to a staging area where a variety of auditing and metadata generation steps are carried out. All original files as well as all the additional files generated by the repository are then packaged and ingested into Archival Storage. Defined strategies, processes and required workflows lead to the creation of these manageable objects.

The OAIS defined digital archive is split up into different functional areas, each with associated processes. Three primary functional areas are ingest, storage and access. The digital objects and their associated metadata form an Information Package (IP). The IP is delivered by the producer and is called the Submission Information Package (SIP). The version of IP stored is the Archival Information Package (AIP). The package delivered to end-users is the Dissemination Information Package (DIP). The Information Model describes the distinction between ingest, storage and access workflows.

It defines all the actions an object undergoes, thereby describing the entire lifecycle of an object, from ingest to dissemination. The object’s lifecycle history is documented in metadata, generated during certain defined events the object undergoes during its journey from ingest to dissemination. Metadata that documents what has occurred with the object over time makes up the “digital provenance” part of the preservation metadata schema.

The Information Model lays out when these “events” are to take place in the workflow and where it is documented in the preservation metadata.
This digital provenance trail then provides the reference framework by which the actions an object has undergone can be verified for compliance with an archive’s lifecycle management policy. By comparing the resulting provenance metadata with the defined workflow events, an archive can be assured that an object has not undergone any unexpected processing. In the end this provides the end-user with the needed authenticity guarantees.

**WORKFLOW SCHEMAS**

The model describes not only the actions an object undergoes during various workflows, but the properties of the objects themselves as well, represented by two different schemas. The workflow schemas describe the object’s lifecycle from a linear perspective, while the object schema illustrates which metadata is present during different stages in the object’s lifecycle.

The different object types are defined as sets of essence, associated support files and metadata files. In the SIP phase, the accompanying metadata is limited to what the Producer has delivered with the file. It contains a minimal set of descriptive and rights metadata and possibly additional information about the source of the file or other technical metadata. The AIP contains the most metadata of all three object types. It includes not only descriptive metadata but the complete set of preservation metadata, made up of technical information, provenance information, source metadata and rights. The provenance metadata documents all the events that have played a role in the lifecycle of the object. Generally, the DIP contains only a subset of the AIP metadata. If requested in the Submission or Order Agreement, all the preservation metadata can be included in the DIP.

A basic set of events to ensure the long-term preservation of the digital object can be found in various requirement and best practice documents.
The first version of the Information Model focuses on workflows reflecting the type of audio and video material Sound and Vision provides to broadcast and cultural institutions.

The first step in the preservation process actually occurs before ingest. This consists of the negotiation and documentation of terms between the owners of the material and the archive and results in the creation of a Submission Agreement. Such contractual documents explicitly define all the actions that are to take place during the archiving process and includes requirements such as formats to be delivered, rights to preserve the material, when and what kinds of validation the objects will undergo and how error reporting is to be handled. Upon ingest a virus check takes place in order to prevent any possible damage to the entire digital repository. The “fixity” check follows, which involves the comparison of two checksums in
order to verify that the repository has received a file that has not been damaged during transport. This check provides the Producer and the Archive with guarantees that the ‘integrity’ of the file is still intact upon arrival in the repository system. If Producers are not able to deliver checksums with their files, the Submission Agreement should spell out the risks involved: the Archive can only “fix” the state of the file after it is received into the repository; thus, the “integrity” guarantee begins from that moment on. The repository cannot be held accountable for any corruption they may have taken place during transport to the system.

The next step is file characterisation, done by extracting technical metadata. Such characteristics include things such as aspect ratio, colour space, codec identification and bit rate. The extraction of technical metadata provides the repository with a detailed picture of the technical nature of
the files it stores. Only by exposing these characteristics can the repository identify potential risks associated with the files. Identifying a particular risk may lead for example, to the migration of a format. This characterisation step also allows the archive to verify that the files submitted actually comply with those agreed to in the Submission Agreement.

The last section of the workflow describes how DIPs are created. The DIP workflow always begins by authenticating the user. This is followed by a search for certain types of material and a determination of the use case. When the system has determined that the material can be made available for this purpose, the DIP is generated. If the DIP is equal to the AIP, a checksum is generated and a fixity check performed to ensure the integrity of the DIP. If a DIP is requested that requires transcoding to another format, or part of a file is requested (a partial-restore), a checksum
is generated for this new version and delivered along with the DIP. This workflow is relatively generic in that a ‘request’ could include a request simply to search in the catalogue as well as a ‘request’ for the delivery of an MXF file.

**DEVELOPING A PRESERVATION METADATA DICTIONARY**

Although some technical metadata is present in various applications in Sound and Vision’s current IT architecture, it is not structurally defined or managed as a component of a preservation system. The task was to develop a metadata dictionary that includes both essential technical characteristics of audio-visual files as well as preservation metadata, which focuses pri-
marily on digital provenance: that is, metadata documenting the creation, chain of custody and change history over time. Rights related metadata, strictly referring to the rights to preserve, also form an essential element.

For the technical metadata, a study was made of a variety of AV specific metadata schemas (PBCore, EBUcore, AES, LC VideoMd, AudioMD and NARA’s reVTMD). Metadata schemas developed by digital repositories operating primarily in the academic library domain (Rutgers University and University of Texas) were also studied primarily because of the access they provide to well-defined dictionaries based on existing standards. And, they offer examples of actual, implemented metadata systems in institutions that consider audio-visual collections an essential part of their digital preservation responsibility. PREMIS was chosen as the standard to use for the digital provenance and (preservation related) rights metadata.5

The dictionary consists of technical metadata for audio, video and still images. It also includes PREMIS digital provenance attributes and a complete defined set of ‘events’ to be carried out as defined in the Information Model workflow. The biggest challenge in compiling the dictionary was trying to reconcile a set of diverse metadata schemas. The schemas differed not only in their definitions of which attributes were relevant at what level (file/stream), they also had different attribute names and slight differences in definitions that made comparison difficult. In addition, differing opinions in the field as to which attributes were essential and which were ‘nice to have’ added complexity to the decision making process.

Ultimately a balance had to be found between including every possible characteristic and including what was considered essential at this point in time, with the underlying assumption that future versions of the dictionary may contain additional attributes.

---

5. PREMIS was also chosen as the standard in the PrestoPRIME project whose focus was on the preservation of digital audio-visual collections. www.prestoprime.eu
The current dictionary does not yet include extensive metadata that reflects documentation of re-formatting related processes for files created as a result of a digitisation project. Although PREMIS has a place for ‘creating application’ and some creation related information, it does not include the full level of re-formatting metadata that NARA’s reVTMD offers (in fact that is why NARA created the schema); or that offered by the AES 57-2011 schema.

These schemas offer much more granularity and detail in areas such as what transfer machine was used, calibration, needles employed in disc transfers, etc. This extensive digital provenance information, along with a clear link to the technical characteristics of the original analogue source, are considered by some experts to be essential in ensuring that the resulting digital file can be considered authentic by users.6

---

**Fig 6: Excerpt from Sound and Vision’s Preservation Metadata Dictionary V1.0: examples of technical metadata for AV files.**

<table>
<thead>
<tr>
<th>Attribute of</th>
<th>Name</th>
<th>Definition</th>
<th>Value type</th>
<th>Obligation</th>
<th>Repeatable</th>
<th>Data Constraint</th>
<th>Values allowed, or link to CV-list.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moving Image</td>
<td>duration</td>
<td>The elapsed time of the entire item or track in playback</td>
<td>Text</td>
<td>M</td>
<td>NR</td>
<td>Structured form.</td>
<td></td>
</tr>
<tr>
<td>Audio</td>
<td>dataRate</td>
<td>Also known as bit rate; the rate at which data is presented within the code; Data rate of the compressed data over time expressed in bytes per second.</td>
<td>Numeric</td>
<td>M</td>
<td>NR</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Moving Image</td>
<td>dataRateMode</td>
<td>Indicates whether the stream data has been processed to achieve a fixed (constant) or variable bit rate.</td>
<td>Binary</td>
<td>M</td>
<td>NR</td>
<td>CV</td>
<td>Allowed values (LC): Fixed; Variable.</td>
</tr>
<tr>
<td>Audio</td>
<td>timecodeInitialValue</td>
<td>Starting value for timecode.</td>
<td>Text</td>
<td>M</td>
<td>NR</td>
<td>Structured form.</td>
<td></td>
</tr>
<tr>
<td>Moving Image</td>
<td>timecodeRecordMethod</td>
<td>Method for recording timecode on the video source item</td>
<td></td>
<td></td>
<td></td>
<td>See also: <a href="http://vocabs.dublincore.org/terms/timeCode/">http://vocabs.dublincore.org/terms/timeCode/</a></td>
<td></td>
</tr>
<tr>
<td>Audio</td>
<td>timecodeRecordType</td>
<td>Type of timecode recorded on video source item, e.g., SMPTE droptime, SMPTE nondroptime, etc.</td>
<td>Text</td>
<td>M</td>
<td>NR</td>
<td>CV</td>
<td>[LTC]; vertical interval (VITC); Other</td>
</tr>
</tbody>
</table>
Fig 7: Excerpt from Sound and Vision’s Preservation Metadata Dictionary V1.0 with examples of preservation metadata based on PREMIS.

Finally, a more extensive set of technical metadata for analogue carriers needs development and inclusion in the dictionary as well. Here again, both reVTMD and AES 57-2011 offer solutions. In fact, a metadata expert at Rutgers makes a good case to extend the AES 57-2011 standard to cover all multi-media carriers and is working on such an implementation at Rutgers.7


IMPORTANT FINDINGS

1. SYNCHRONISING THEORY WITH PRACTICE

The OAIS model is an abstract one; it needs precise definition when applying it to a specific archive environment. It is not an “all or nothing” concept. Its requirements are actually high level principles and provide plenty of room for interpretation. This makes the model on the one hand, a nuanced and flexible one. Yet on the other hand, it does not provide ready-made implementation solutions. In the end, what is desirable and can actually be realised in a running archive will determine what specific solutions are chosen.

At Sound and Vision, the TDR project was running at the same time that the process to acquire a new MAM-system began. In order to enable that a new MAM would contribute to OAIS compliance goals, steps in the Information Model workflows were quickly translated into detailed and concrete MAM requirements. This meant that some requirements outside of the IP workflow needed to be analysed. It was at this point that it became clear that OAIS compliance requirements themselves offer little concrete or technical solutions. It was not always clear if requirements were to be met by way of workflow implementations or through technical requirements on an application level. Each workflow requirement was then identified as one to be met by either a MAM application or elsewhere. In the end, those OAIS compliant lifecycle management requirements, that became part of the overall MAM requirements, actually represent a sub-set of the overall OAIS requirements.

2. PRESERVATION FROM AN ENTERPRISE IT ARCHITECTURE VIEWPOINT

The need to identify which OAIS requirements were relevant in a MAM
requirements list clearly illustrated that not all preservation requirements could be met by one application. It is recommended that different requirements be fulfilled by different components within the total Enterprise IT architecture. To realise this, much more insight is needed into the differences as well as the underlying relationship between the types of (meta) data, the different systems, the workflows and the functions of the Enterprise IT architecture, in relation to OAIS compliant digital lifecycle management. For Sound and Vision there’s still a lot of work to do.

When moving from the analogue preservation world to the digital IT domain, it helps to consider preservation from an enterprise IT architecture perspective. Such a perspective requires a clear understanding of the difference between general IT processes and those processes associated with the “business” of preservation. Confusion may stem from the fact that there is overlap in some OAIS functional areas (take the functional OAIS area Administration for example, where issues about security and risk are covered, issues that form a part of regular IT management processes.) IT processes for example, focus on whether specific applications within the IT infrastructure are working properly – is it secure, are there errors, is it managing capacity properly? An Enterprise IT architecture viewpoint focuses on the lifecycle management of the data/content: the business activities that need to be carried out on the data; in the TDR context, the business processes that ensure the long-term preservation of the object. One illustration of this is mistakenly equating ‘storage’ with preservation.

In fact, storage is simply one of a set of IT applications within an Enterprise IT architecture system. It is concerned with such things as “writing/recording the data, retaining the data for an undetermined length of time, reading and accessing the data. More advanced storage may include replication of files through techniques such as RAID, erasure coding, or the simple generation of multiple copies on different media or at different sites (geo replication), or the optimisation of data placement, e.g. in hierarchical storage management (HSM) systems. A preservation system
on the other hand involves processes that need to take place during ingest and access to content to support preservation; it includes carrying out preservation actions performed on content. Preservation is a process that includes the storage, but is not limited to it. A preservation strategy must be defined and translated into business processes that are then associated with IT functional areas. Applications are then found to ensure that such functional processes are carried out. The TDR project set out to define these strategies and processes.

Fig 8: Distinctive levels within the Enterprise Architecture of Sound and Vision and their corresponding functions and domains.

3. LEARNING, COMMUNICATION AND COLLABORATION

It was clear at the start of the TDR project that digital preservation processes are not confined to what the IT department manages; implementing OAIS compliant lifecycle management involves the entire archive institution: at ingest, during storage, from description to dissemination. This means that the entire organisation play a role in providing long-term preservation and access guarantees: the acquisition and selection staff, the cataloguers as well as the access services staff. OAIS compliance makes more explicit the important relationship the archive has with its producers/donors as well with its Designated Communities: all incoming and dissemination activities play an integral role in the archival lifecycle of the material.

Therefore, one of the most important goals of the TDR project was to try and involve the entire organisation in digital lifecycle management decision making and to share knowledge about digital preservation as broadly as possible. A project structure was chosen as the best approach, even though the work primarily focused on policy development. It was thought that by involving different staff members (catalogue, system application and access services staff as well as policy advisors) everyone would better understand their own role, which would in turn strengthen the responsibility they feel in the digital lifecycle management process. This goal was only partially achieved. It was too difficult for most project members, alongside their daily work responsibilities, to spend time researching and translating theoretical processes into archive policy. In the end, the work was carried out by Sound and Vision’s IT Information Policy division, who had originally been assigned this task. The project approach did succeed in raising the level of consciousness about different aspects of digital preservation and digital lifecycle management: what these concepts mean and how important it is to collection management. The OAIS standard, its terminology and its process framework are no longer foreign to the organisation. In
addition, certain technical processes such as validation and fixity checking have gained prominence.

It’s important that the knowledge gained during the OAIS requirement trajectory remains the reference framework within Sound and Vision’s IT department, who in the meantime, needs to start finding solutions to support the newly developed business processes. It is hoped that the defined lifecycle management requirements demonstrate to IT staff that preservation processes and workflows first need mapping to functional areas in the Enterprise IT architecture system; only then can applications be identified that may be able to carry out the needed functionality.

**TO CONCLUDE**

There are many essential questions and issues that remain after having defined the OAIS compliant audio-visual archive requirements. Besides working out where within the Enterprise IT architecture the preservation functions need to land, budget and cost modelling needs to start. What are the cost implications of this workflow and information scenario? Who will pay for which preservation service within the digital lifecycle management process; the archive itself as “Trusted Repository” or the producer/donor as deliverer of digital collections or the Designated Communities for whom the files need to be kept technically up to date forever?

Another burning question concerns all the previously ingested files and metadata: how, for example, to ensure that what is now essentially ‘dark metadata’, generated during earlier digitisation processes, is brought into the preservation workflow and data management system? What do you do with all the files that were ingested earlier, into a non-OAIS compliant repository, and thus never underwent the fixity checking and validation processes an OAIS compliant repository requires, especially given the nature of the amount – 400,000 hours? Further, the preservation business
processes themselves need further development. Are we going to apply the normative workflow and metadata processes to all the collections that are ingested or are we going to apply them in only some cases? If so, how: are we going to define different preservation levels for broadcast production material and cultural heritage material considered to be of national importance? And are we going to apply to all the different types of content files (metadata, photos, contextual documents, etc) the same preservation level or will this only apply to Sound and Vision’s core collection?

All of these questions must eventually be fully answered and translated into policy to ensure a rational and cost effective business operation. However, the Information Model and the other normative policy documents produced in this project provide an important reference framework against which the AV archive can measure how far its current operations
consciously reflect preservation lifecycle management, in the way it approaches the ingest, storage and dissemination of the collections, whether broadcast production or cultural heritage material. There lies a solid, theoretical basis for establishing a technical and organisational preservation structure in the audio-visual broadcast domain that can be considered OAIS compliant. All those who play a role in the preservation process have been identified. Sound and Vision now knows how to operate according to the standards and can thus — by implementing the systems — demonstrate how the standards have been implemented and why. All this allows the organisation to prove to its producers/donors and users, how, by having implemented the basic requirements of a ‘trusted’ repository, it operates responsibly.

Based on the project’s achievements, Sound and Vision plans to seek a sort of certification for digital archives, the so-called Data Seal of Approval9. The normative documents that have been delivered can serve as sound, documentary proof that some of the DSA requirements have been fulfilled. Thus the first step towards the status Trustworthy Digital Repository has been realised.

AV PRESERVATION WORKFLOWS

REFERENCES:

[1] Kwaliteitseisen Digitaal Archief Beeld en Geluid 1.0
[2] Informationmodel Digitaal Archief Beeld en Geluid 1.0
[4] Handleiding Submission Agreements, 1.0
[5] Designated Communities Beeld en Geluid, typering en uitleverreisen, 1.0
Eva Lis-Green is Media Quality Controller and Group Manager at the Archives and Rights Department of SVT. She has worked at SVT since 1984 and has a background as Head of Archives and Transmission, Head of Documentation, Information System Coordinator at the Engineering Department and Project manager for different IT and Media Management systems. Eva-Lis is a specialist in archiving workflows and information systems in the field of media asset management.

Kaisa Unander worked for 20 years in public service at the SVT, the Swedish national broadcaster, the first 10 years in production of news and current affairs. After that, Unander was employed as manager of the newsroom, playout and MCR. In all these positions she functioned as change leader of the digitisation process. Since 2011, she is working as a manager for technical services at the audiovisual department of the National Library of Sweden.
FROM THE 1ST of April media companies, like broadcasters and newspapers shall deliver all web unique content that is electronically published on the Web to the National Library due to a new legislation in Sweden. The Media Companies have to deliver different types of files (media objects) with additional metadata. SVT is one of the media companies and will in this paper give examples how this can be executed. The National Library has to take into account everything that will be delivered from different suppliers and the need for a structured way of taking care of this information so it can be preserved for the future. When the legislation is fully implemented in 2015, deliveries will also be mandatory for all official websites in Sweden. The principles from the OAIS-model will be used in this process. The challenges and possibilities will be discussed.
PROCESS FOR LEGISLATION

The Process started in 2003 when the question was discussed how to pre-
serve the web content in a more flexible way than sweeping the web i.e.
taking snap shots of web pages.

In February 2009 an investigator got the assignment to examine the ques-
tion of legal deposit for electronic documents. On the 16th of November
a draft was presented and sent to various interested parties for consider-
atation. According to the draft the propositions are to be regarded as an
amendment to the present legal deposit act. Integration lies further ahead
in the future. The act came into force the 1st of July 2012. According to
the Act there is a transition period of two and a half years including a se-
lected group of suppliers of about 40 suppliers. In collaboration with the
National Library they are going to develop a technical and administrative
solution. Their period of delivery starts on the 1st of April 2013. The de-
livery period of all suppliers start on the 1st of January 2015.

During this period the National Library and Sveriges Television has taken
part in hearings and also has been able to give consultation responses.
Some of the feedback given to the investigator was taken into account,
others not. The cost for the supplier to build the delivery system and also
the receivers cost for storage was discussed. The difficulty to deliver cor-
correct metadata was another issue discussed.

There were few guidelines or help to understand the law when the Na-
tional Library started to create the system for receiving material, and there
were huge expectations among the partners that the National Library
should make it easy for them. In the first meetings the law was really criti-
cised. For the moment the practise is there are different metadata schemas
for the governmental organisations, mass media companies and “others”.
Surely there will be different standards in the future to make it easier for
both but we have to create those together.
E-LEGAL DEPOSIT – WHO, WHAT, HOW?

WHO?

The legislation has stated that any producers with automatic constitutional protection according to the Freedom of Speech Act, i.e. editors of newspapers, journals; radio and TV programmes; news agencies, as well as professional producers, universities and publishers, shall deliver web unique content to the National Library. The selected suppliers during the transition period 31 March 2013 – 31 December 2014 are 15 authorities, 10 daily newspapers with the largest circulation in Sweden, 10 journals with the largest circulation in Sweden, Companies with nationwide radio and TV transmission in Sweden and local radio over the Gothenburg area.

WHAT?

The following contents are defined to be delivered: electronic material that is completed and of permanent character, electronic material that is made available to the general public by ”transmission over the network” and has the general public in Sweden as main target group, or the material has a Swedish author or creator, or the material is presented in Swedish to a substantial amount. This means that articles from newspapers and journals (reviews, debate, blog etc.), parts of local websites, advertisements produced and published by the supplier, leaflets, guides, reports, e-books etc. moving images (web-casts, videos), web radio, pod radio, pictures, photos, i.e. editorial electronic material that are web unique, shall be delivered.

However entire websites and databases, programme codes/applications structuring the databases, live broadcasts, wikis or other continuously updated sites, information on the intranet, private published pictures, music, films, blogs, comments etc., material of more ephemeral character
(calendars etc.), articles by news agencies, short commercial films etc. introducing web TV features from the supplier are not part of this legislation.

HOW?

In the legislation it’s stated that the content can be delivered on USB or CD as a carrier and that alternative delivery methods can be online network – ftp, Web forms, RSS. Most likely the deliveries will be the last method and in this case you can see how fast technologies changes and a legislation recommendation can become obsolete.

According to the e-legal deposit act the National Library of Sweden may stipulate certain metadata. Regardless the choice of delivery there must be information about the delivered files concerning;

- Where and when the file was first published

- The file format

- Codes that are needed to enable reading a password protected document relations to other files (for example: an article is structured in text, pictures etc.)

- The relation between an e-legal deposit file and an object that is part of a legal deposit delivery e.g. a printed book with password protected exercises on the web. All this is necessary for the long term preservation and accessibility.

The National Library aims to use the OAIS-model but there is still work to do to become a Trusted Repository. The six main principles are to be
followed but the question of access makes it complicated as well as the change of long-term storage in the near future. It is easier to find the understanding for the model than the money. The organisation is surprised about the costs, we are slowly moving towards the cyber economy that makes is more difficult to estimate the needs than before.

**CASE STUDY SVT**

The deliveries can’t be done in one way since the different objects shall be delivered separately together with additional metadata, and we are not prepared to deliver everything at once. There are a lot of new techniques and workflows to be implemented. Therefore together with the National Library a delivery plan was decided.

It’s stated in the legislation that the content shall be delivered at latest three month after publication. We have divided the deliveries into four categories:

- Web unique video
- Articles and stills
- Live streamed video
- Open Archive

The on-demand web unique video will be the first part to be delivered. The workflow is built and we have started to collect the content for delivery from the 1st of April. The tricky part here is to find out what is web unique since the service, SVTPlay, does not differentiate. To find additional metadata in this case is easy since we can use the information from our scheduling and planning system. We can also to some extent track the web unique programs in the planning system. The suggested National Library XML-format is used.
The Articles and Stills delivery (svt.se) will be implemented at the earliest in September 2013. We have to set up a RSS-feed possible for the National Library to collect via FTP. The suggested National Library XML format can also be used in this case. Subtitles have to be delivered separately as well.

For the live streamed video the system will be built during the summer and will hopefully start to record the video in August. The same reference recording system as used for broadcast will be enhanced and used for this purpose. For this delivery the live streamed video programs also is needed to be planned and scheduled in the planning system which means new workflows for the entire productions environment. It has taken some time to get this workflow accepted and in place. But in this case the same XML-format used for Broadcast metadata will be used. The Open Archive video will be manually handled since just a selection of program and clips not published before will be delivered.

E-LEGAL DEPOSIT – RIGHTS MANAGEMENT

There is nothing in the legislation that tells anything about access. Since the copyright legislation is complicated it was stated not to take this issue into consideration.

- To whom – we don’t know
- When – we don’t know
- Where – we don’t know

The rights management is of course one of the real challenges for the system too, not technically, but because we don’t know about the future use it is difficult to do the right thing from the beginning and it makes it frustrating.
CONCLUSION

There is no clear picture concerning the metadata needed for preservation reasons since there is no knowledge about the reuse. New standards, new formats will constantly occur and new information is needed to come along with all these changes. Even if the National Library can request certain file formats and XML formats for metadata there is no standard format decided in the legislation and the National Library has to be prepared for several different types of delivery.

The National Library has a plan for long term storage and preservation but for the e-legal deposit it is obviously even more difficult than for the existing collections in the National Library today. From SVT alone four different deliveries are needed to describe the object.

The National Library is looking forward to be able to search and test how well the whole system works. We are still building the workflows and IT systems; we need to validate a lot and maybe change our minds!
4. AUTOMATICALLY GENERATED METADATA
Cees Snoek is currently an assistant professor at the Intelligent Systems Lab of the University of Amsterdam and head of R&D at Euvision Technologies, one of the lab’s spin-offs. Previously he was affiliated with Carnegie Mellon University (2003) and UC Berkeley (2010-2011). His research interests focus on video and image retrieval. Dr. Snoek is the lead researcher of the MediaMill Semantic Video Search Engine, which is a consistent top performer in the yearly NIST TRECVID evaluations.

He serves on the editorial boards for IEEE MultiMedia and IEEE Transactions on Multimedia. Cees is recipient of an NWO Veni award (2008), an NWO Vidi award (2012), and the Netherlands Prize for ICT Research (2012).
A CLASSICAL IMAGE-UNDERSTANDING problem is the deciphering of Egyptian hieroglyphs. Hieroglyphs ceased to be the alphabet of choice in favour of Demotic script and ancient Greek, around 400CE. Soon after, understanding of hieroglyphs was lost completely. In the centuries that followed many scientists tried to decipher hieroglyphs\(^1\), but it took until 1799 before real progress could be achieved. In that year, savants sent to Egypt by Napoleon discovered an inscription carved on stone that provided a translation of the hieroglyph images into Demotic script and ancient Greek. Eventually, this Rosetta Stone was the key to complete

decipherment by Jean-François Champollion in 1822. Image understanding in modern times shares many similarities with deciphering hieroglyphs. In this paper we discuss recent research efforts at the University of Amsterdam that aim to translate an image into its most descriptive summary on concept and sentence level.

**UNDERSTANDING IMAGES BY RECOGNISING CONCEPTS**

For humans, understanding and interpreting the visual signal that enters the brain is an amazingly complex task. Approximately half the brain is engaged in assigning a meaning to the incoming image, starting with the categorisation of all visual concepts in the scene. Thanks to several computer vision and machine learning breakthroughs, categorisation of images at the concept level is also within reach for machines. However, similar to deciphering hieroglyphs, the most important contributor for machine understanding of images is the availability of written image translations in the form of image labels.

The standard approach to machine understanding of images consists of four basic steps. The first step is to gather positive and negative examples of a certain visual concept, say a boat. The set of labelled images is divided into a training set and testing set. The training set is used for algorithm optimisation and learning of a so-called statistical model that captures the visual representation of the concept of interest into a mathematical formulation. The test set is used to evaluate the recognition abilities of the model by comparing its predictions with the original label. The second step builds up a model of a concept. To do so, each image is analysed by extracting up to one million visual features from the content. These features are invariant descriptors that cancel out accidental circumstances of the recording caused by differences in lighting, viewpoint, or scale. The third step is to project the descriptors per pixel onto one of 4,000 words. These aren’t real words, but
rather summarisations of one local patch of the image describing a single detail: a corner, texture, or point. In the fourth step, a machine-learning algorithm converts the visual words into a probability of a concept being present in an image. These probabilities are used to rank all available images in terms of concept presence. The process to understand images by recognising their most descriptive concepts is summarised in Figure 1.

Crucial drivers for progress in image understanding are international search engine benchmarks such as the TRECVID (TREC Video Retrieval) benchmark organised by the National Institute of Standards and Technology².

Fig 1: General scheme for detecting visual concepts in images.

---

TRECVID aims to promote progress in content-based retrieval from digital video via open, metrics-based evaluation. With the support of 50 teams from academia and industry, including the University of Oxford, Tsinghua University, and IBM Research, it has become the de facto standard for evaluating video retrieval research. Benchmarks’ open character ensures the rapid convergence of effective image understanding approaches.

Recently, we assessed progress in understanding images by recognising their most descriptive concepts by comparing a state-of-the-art search engine from 2006 with one from 2009. We took into account both the situation where the training set data was visually similar to the testing set, i.e., both contained video from the same genre, and the situation where the training set data was visually different from the data used for testing, i.e., both contained video from different genres. As Figure 2 shows, search engine performance doubled in just three years. For recognising concepts, detection rates degenerated when applied to data of a different origin, yet still doubled in three years. Progress may be greater than expected, but it doesn’t imply that the general problem of visual search is solved. Our experiment used only 36 concepts, whereas broad image understanding would require thousands of detectors approaching the vocabulary of a common user.

Expanding the concept vocabulary to one approaching human language demands a complete ‘Rosetta Stone’ that provides hundreds of image translations per concept. In obtaining these translations, the common approach in the literature is to rely on human-expert labelling, but such labelling is costly and hence restricted. We studied whether it is feasible to shift the effort from human experts to amateur-consumers who exploit Internet services like YouTube and Flickr. At these sharing-websites, expensive label quality is substituted for free at an unprecedented scale, but the labels are known to be uncontrolled, ambiguous, and overly personalised. Hence, a fundamental problem is how to interpret the relevance of a user-contributed tag with respect to the visual content the tag is describing. Intuitively, if different persons label visually similar images using the same
tags, these tags are likely to reflect objective aspects of the visual content. Starting from this intuition, we have proposed an automatic algorithm, which accurately and efficiently learns tag relevance by accumulating votes from visual neighbours. The main principle behind the automatic image tagging software is quite simple: take an image, for example the one in Figure 3, which is labelled with the tags, bridge, bicycle, perfect and MyWinners. Then look at other images that share a visual similarity with the image of interest. If the visually similar images are also tagged with the same labels most of the time, it is quite likely that this tag is the most descriptive label for the image. In case of Figure 3, the label ‘bridge’. Human-provided tags on the web act as ‘Rosetta Stone’ for image understanding on the concept-level.

Fig 2: Visual-concept search progress as evaluated on 36 detectors (•).
Fig 3: Collecting free labelled image examples from the web with tag relevance by neighbour voting.

**UNDERSTANDING IMAGES BY RECOGNISING SENTENCES**

While results for image understanding using single concepts are impressive, little has been achieved for sentence-level event descriptions. This comes without surprise, as an event is not a concept. The ideal event detector must be able to provide a human understandable recounting on what visual information in the video is decisive for relevance. Detecting chair, chair, chair is not informative, theatre is. Nevertheless, some researchers treat events the same as concepts, thus the same recognition process can be applied. However, from the sheer number of highly correlated features and projected words, it is not easy to derive how these detectors arrive at their event classification. Moreover, events are often characterised by similarity in semantics rather than appearance. Our goal is to find an in-
formative representation able to recognise, and ultimately describe, events in arbitrary video content. We argue that to reach that long-term goal a more semantic representation is urged for.

As a first step, we analyse the vocabulary that humans use to web videos containing events. Figure 4 illustrates some videos and their corresponding textual descriptions. We process textual descriptions for 13,265 videos in total. After basic textual pre-processing like removing stop words and stemming, we end up with 5,433 distinct terms. Looking into the human vocabulary, we observe that the terms used can be mapped into five distinct concept types as typically used in the multimedia and computer vision literature: objects, actions, scenes, visual attributes and non-visual concepts. We manually assign each vocabulary term into one of these five types. After this exercise we observe that 44 percent of the terms refer to objects. Moreover, we note that a considerable number of objects are dedicated to various types of animals and people; i.e., lion, and teen. About 21 percent of the terms depict actions, like walking. Approximately 10 percent of the concept types are about scenes, such as kitchen. Visual attributes cover about 13 percent of the terms; i.e., white, flat, and dirty. The remaining 12 percent of the terms belong to concepts, which are not visual; i.e., poem, problem, and language. Moreover, we observe that the vocabulary contains both specific and general concepts. This analysis provides guidelines for constructing vocabularies for understanding images at sentence level.

Having defined the concept vocabulary humans use to describe events, we are ready for machine understanding of images on the sentence level. Again the first step is to gather positive and negative examples of a certain visual event, say a person changing a vehicle tire. Similar to concept detection the data is divided into a train set for development and a test set for evaluation. The second step builds up a model of an event. To do so, we decode videos by uniformly extracting one frame every two seconds. Then all the concept detectors from our vocabulary are applied on the extracted frames. Concatenating the detector outputs, each frame is represented by
a concept vector. Finally the frame representations are aggregated into a video level representation by averaging and normalisation. On top of this concept vocabulary representation per video, we use in the third step again a machine-learning algorithm that converts the concept vocabulary into an event score. These probabilities are used to rank all available videos in terms of event presence. Similar to recognition of concepts, the labelled event examples are crucial for describing the image at sentence-level.

The question what automatic detectors to include in the vocabulary for automated sentence description of imagery is a topic of on-going debate. Recent research at the University of Amsterdam has investigated vocabulary construction for automated description of images on sentence-level. Of course one could rely on a vocabulary containing as many concept detectors one can think of, but one can also learn from examples what concepts

---

Fig 4: Examples of videos and human-added textual descriptions on sentence level.
in a vocabulary are most informative per event. In a recent paper we mod-
elled finding this vocabulary of informative concepts out of a large set of
concept detectors as a rare event search. Our proposed approximate solution
finds the optimal concept vocabulary using a cross-entropy optimisation.
Our experiments reveal that 1) some concept vocabularies are more infor-
mative than others for specific events, 2) event detection using an automati-
cally obtained informative concept vocabulary is more robust than using all
available concepts, and 3) the informative concept vocabularies make sense,
without being programmed to do so. See for example the informative
custom concept vocabulary for the sentence a person landing a fish in in Figure 5.
So it seems that for video event detection using concept vocabularies it pays
to be informative. However, much more progress is needed before sentence-
level image descriptions can be generated as accurate as concepts.
CONCLUSION

In this paper we discuss recent research efforts at the University of Amsterdam that aim to translate an image into its most descriptive summary on concept and sentence level. We believe that the progress in artificial intelligence, by means of computer vision and machine learning, coupled with the widespread availability of human image descriptions on the web, act as ‘Rosetta Stone’ for image understanding. It is our belief that automatic metadata generation for images jumps from single words to complete sentences, soon.

3. Open-access publications available at http://www.ceesnoek.info
Sam Davies joined BBC R&D in 2007 working on a variety of projects including high frame rate television, object tracking in sporting events and image recognition.

Since 2009 he has been working on the Multimedia Classification project, identifying new techniques for metadata generation from audio and video content in the BBC archive. This work has resulted in prototypes which offer unique ways to analyse the semantic and affective, or emotional content of audio, visual and text documents.

He is currently studying for a PhD in Information Science at City University, London, looking at federated search for broadcast archives incorporating semantic and affective metadata.
AS DIGITAL MEDIA archives grow, there is a need to produce metadata to allow for more effective classification and retrieval not only for broadcast professionals but also for consumers who just want to find something to watch. The British Broadcasting Corporation (BBC) has a large broadcast archive, comprising some 1.5 million hours of audio and video. This archive does not contain all BBC output, only that which has been deemed culturally important by editorial, production and archiving professionals. These professionals also provide metadata with each programme, allowing for searching on programme title, synopsis (if available) or from a tag associated to the programme based around the London Classification system, a BBC devised extension to the Universal Decimal Classification system. Users can then request a copy of the content from the BBC archive. Since 2007, the BBC has also been recording all of its programming off-air, in a project called BBC Redux [1]. This records the Digital Terrestrial Televi-
sion (DTT) signal in the London area, storing it digitally and allowing for programme search based upon the programmes title, date and time of transmission or on any keywords present in the programmes Electronic Programme Guide (EPG). Radio programmes are also stored. In 2011 BBC R&D produced an extension to this entitled “Snippets”. This service indexes subtitles which accompany the broadcasts, allowing users to search through this transcript of the broadcast to find the programmes, or programme clips they want. Users can then either watch the off-air copy of the programme, or order the broadcast quality tape from the archive. A screenshot of this is show in Figure 1.

For non-professionals, the BBC launched the iPlayer service in 2007 which allows users in the UK to receive BBC programmes via the internet. This is primarily designed as a catch-up service, allowing viewers to

Fig 1. Screenshot of the BBC Snippets system
watch or listen to programmes broadcast within a specific time scale. As of May 2013 this window was usually set to seven days, allowing viewers to catch up on any programmes they may have missed in the last week. As a catch-up service it is based around known item searching: users know about what they want from the available collection before they search. As such they will most likely know certain facets about the programme such as the programme title, when it was broadcast or who contributed to the programme. The current BBC iPlayer service is set about to serve this need. Users are able to search based on these facets or can browse through available programmes by day, channel or genre. With only a selection of the last seven days of content available, this is a feasible task. The BBC also provides “programme” pages on their website: a programmes home on the BBC website which collects together all information publicly available from the BBC about a programme or series. However, as more programmes are made available from the BBC’s archive, new methods for searching and browsing the archive may be required. With an increase in content unknown to viewers, more metadata will be required to help them find interesting and entertaining content to watch or listen to.

MOOD BASED CLASSIFICATION

One solution to this is to classify a programme based on its mood. We have been developing our mood based classification system to offer an alternative solution to traditional programme search and retrieval. This content based approach (in that we analyse the content of the programme rather than any existing metadata) aims to identify the mood of the programme.

1. Available at www.bbc.co.uk/programmes
by analysing key emotional components and classifying accordingly.

Our mood based system comprises of three main sections, an overview of which is given in Figure 2. This work was originally presented in [2]. Initially a classification extraction system performs low level signal analysis on the video and audio signal. We then use these characteristics to identify semantic features from the programme. Finally we use a combination of the raw characteristics and semantic features in order to identify the emotional component by a comparison to a pre-labelled ground truth data set. In this section we describe each of these sections along with an overview of our ground truth data collection method.
MOOD BASED CLASSIFICATION: SYSTEM OVERVIEW

The Characteristic Engine analyses both the video and audio signal.

VIDEO ANALYSIS

Two characteristics of the video signal are extracted in our initial framework: tiny grey images (reducing the image to a 32 x 32 greyscaled version) and summation luminance value (summing the brightness of a tiny grey image).

AUDIO ANALYSIS

Five different audio characteristics are extracted from the audio: Formant Frequencies (identifying components of the audio which relate to human speech), Power Spectral Density (the relative strengths of different frequencies in a complex sound); Bark Scale Filtered RMS amplitudes (the amplitude of different frequency bands, filtered using the Bark Scale); Spectral Centroid (a measure of brightness of the sound) and Short Time Frequency Estimation (an estimation of the pitch of the sound over a short time).

We then used these low level characteristics to identify semantic features:

- Bright Sounds
  - Spectral centroid is used to determine the overall “brightness” (or high frequency content) of the programme.
• Big Sounds

  • This was calculated as the overall intensity of the amplitude, calculated from the average of the Bark Scaled RMS amplitudes coupled with the variance between the more intense sounds and the average.

• Shot Changes

  • Using the techniques in [1], this looked to identify where the shot changes (different camera shots) occurred.

• Motion Feature Value

  • Using the techniques in [1], the motion feature value is produced by the difference between non adjacent tiny grey images separated by 0.4 seconds.

• Changes in Luminance

  • Using the techniques in [1], this is an indication of the change in lighting values between two adjacent tiny grey images luminance levels.

• Amount of Luminance

  • Using the techniques in [1], this is calculated from a histogram of the luminance bins that are above a threshold.
MOOD BASED CLASSIFICATION: GROUND TRUTH COLLECTION

In order to match these low level characteristics and features to moods we need a labelled ground truth dataset to test against. In [3] 200 participants watched short excerpts from various programmes and assigned three different mood labels to the excerpt: a continuous mood label as they were watching it, one after watching the excerpt based on the overall mood and finally one of their memory of the entire programme or series, if they remembered it based around what are termed Affective Adjectives taken from the theory of affect.

THEORY OF AFFECT

Affective theories place emotional response from a stimulus on a n-dimensional vector space. Theories based around affective vector space identify non-exclusive components of emotion and aim to classify affect accordingly. In our work we used the research of Osgood [4].

This model proposes that there are nine components of affect, but that the majority all fall within one of three main components; Evaluation, Activity and Potency (EPA). Evaluation relates to if an induced feeling is perceived as positive or negative. For example, positive evaluation could incorporate stimuli which are perceived such as happy, joyous or funny. Conversely, stimuli which are associated with negative evaluation include anger, pain or sorrow. Osgood’s research found that emotions in this category accounted for around half of all perceived feelings. Potency relates to the size or power of an emotion. Stimuli with high potency would have a very strong impact and instill strong emotions: the opposite to stimuli with a low emotion. The final main component is that of activity. If a stimulus contains much action – movement, power or speed – then it would have a high activity element. Sedate, weak or slow stimuli would have a low activ-
Building Models of Mood

Users were asked to rate the clips on sliding scales between two adjectives. These were: happy/sad, serious/humorous, exciting/relaxing, interesting/boring, slow/fast-paced and light-hearted/dark. From an analysis of these results, it became apparent that the mood of the clips could be defined by two dimensions; light-hearted/serious and fast-paced/slow-paced.

These results were used as the ground truth for the machine inference section of the system identified in Figure 2. The clips themselves were analysed using the characteristic and feature extractions engines and models for fast/slow paced and humorous/serious programmes built. These were then used to classify other programmes. In [3], we achieved an accuracy rate of over 95 percent.

The two dimensions of mood identified were also used as the basis for a Graphical User Interface (GUI), built to allow users an initial experience of browsing and finding programmes based upon mood.

Mood Based Classification: Graphical User Interface

The GUI built allowed users to navigate through and search for programmes which had been classified by our system. This GUI was not built to offer an ideal user experience, more to as a means for demonstrating our work. The GUI is shown in Figure 3.
This classifies the programme based on a two dimensional axis: serious/humorous and fast/slow-paced. Highlighted at the top of the display is a programme which we classified as being one of the most humorous and fast paced. This programme is an episode from a BBC comedy series “Never Mind the Buzzcocks”, a fast paced panel comedic show. Along with browsing programmes by mood, users can also search for specific programmes (allowing them to find programmes similar to one they know with a similar mood) and by actor (allowing users to find their favourite actor in a variety of different programmes). Viewers are also able to search

2. More information on “Never Mind the Buzzcock” can be found at: http://www.bbc.co.uk/programmes/b006v0dz
by genre, which has been shown in [3] to have a high correlation to mood. Another feature of our initial demo is the ability to order programmes against their mood and year of broadcast. As the BBC digitises more of its archived content, we run this through our classification system. This is shown in Figure 4. Here, an increase in pace since the early 1960s can be seen through to 2013, though more programmes from pre 2007 would need to be classified before any definite trends could be established.

This GUI is not intended to be public facing instead to enable BBC R&D to do preliminary tests on the accuracy of the mood classification system and offer a framework for future developments around mood navigation of an archive.

Fig 4. BBC R&D Mood GUI showing programme pace against year of broadcast
MUSIC ANALYSIS

In our initial frame work system (Figure 2), we analyse audio and video features. However there are other aspects to the programme which we are looking to include. One of these is the music within a programme. In [5] we presented methods in which a programmes theme tune could be used to help identify the overall mood of the programme. We devised an experiment, similar to that in [3] where members of the public could listen to excerpts of theme tunes taken from the last 60 years of the BBC’s archive, and rate these on a differential scale using affective adjectives. We also asked participants whether they liked the theme tune they heard, if they remembered the theme tune and programme it accompanied and also if

**Fig 5. User interface for Musical Moods**
they knew (or could guess) the genre of the programme which accompanied the theme tune. We ran this experiment as part of the British Science Association’s National Science and Engineering Week, a UK based week-long series of events aimed at encouraging public engagement in science and engineering. The experiment, named Musical Moods was available online for several months and during that time over 13,000 participants listened to one of the 144 available theme tunes over 51,000 times. The user interface for this is show in Figure 5.

As with [3], we primarily ran this experiment to gather ground truth data to allow us to build an automatic system to recognise the mood of a TV theme tune. We presented this work in [6], where we used audio analysis techniques and support vector machines to identify a theme tunes mood where we achieved an accuracy of over 80 percent.

**TEXT BASED ANALYSIS**

All BBC programmes are currently broadcast with subtitles. These are generally accurate transcriptions of what has been said in a programme and can include indications of other aural events such as sound effects or music. In conjunction with the School of Informatics at City University London, we began to research ways in which we could extract the mood of the programme, or programme segment from subtitles, or indeed any text which accompanies a programme. In our initial work, we devised a system which classified an entire programme. This work was originally presented in [7]. Using a system of language processing and affective analysis, we developed a technique which allowed programmes to be classified and categorised based upon their affective content. In our initial investigation we used genre as our ground truth. As was found in [3] there is a strong correlation between genre and overall mood. We ran the subtitles through our system and the result was a three dimensional vector with axis
relating to the EPA scale. This is shown in Figure 6. Here, clear groupings of programmes by genre can be seen with ‘thrillers’ tending towards exciting and negative emotions and children’s programming more positive and exciting. For this, we trained created SVM models for each genre (thriller, factual, children, comedy, life and soaps) and found precision scores of 0.95 and recall scores of 0.91.

We are currently working on integrating this in our classification framework and trialling with transcripts from automatic speech transcripts.

**COMBINATION OF AFFECT AND SEMANTICS**

One area of interest to us is in the combination of semantic and affect in
programme classification: combining existing information about a programme (who was in it, genre, what’s the programme about) with the mood of the programme. We have begun looking at this and created an initial system, the GUI of which is shown in Figure 7. In this system, we use the affective text classification mentioned previously to identify a programme’s mood and then a Named Entity Recognition and Disambiguation (NERD) system to identify named entities: people, places, organisations etc. In this initial test we used the AIDA-YAGO system [8]. We wanted our system to be controlled from a standard television. Since digital television remote controls are fitted with four control buttons as standard (up, down, left and right), we wanted our system to use these to find new programmes. As such, we used this semantic and affective data to offer viewers four options for an onward journey when they have fin-
ished a programme: semantically and affective similar content (something similar), semantically similar (something informative), affectively similar (something entertaining) and semantically and affectively different (completely different). This GUI is shown in Figure 7. We are currently working on refining the recommendations for this system.

**HIGHLIGHT GENERATION**

Another area we are interested in is highlight generation of a programme. This would allow viewers to watch only the key section of a programme allowing them to catch-up on programme if they do not have time to watch the full version. Our initial work in this area focused on sporting events. Using audio analysis only, we developed a system in [9] which allows for identification initially of when a sports match is being played (as opposed to studio commentary to analysis before, at an interval or after the match) and then during passages of live play identifies semantically interesting sections. This allows the user to either just watch the analysis sections (which often focus on key points in the game though may not be present of older, archived programmes) or else to watch a précised version of the game. Using various audio features, we were able to create a timeline of “interestingness” in a match and use a threshold to identify which sections of play were more likely to be interesting than others. This is shown for a section of a rugby union match in Figure 8. We tested this system against a hand created timeline of interesting events, created by the BBC Sport department and found we had an accuracy of 78 percent in detected an interesting event.

Our current work around this for sports is to see if the inclusion of social media can help identify an interesting event, and also to see if these techniques can be brought over for other genres such as drama or factual.
CONCLUSIONS

We have presented an overview of the BBC R&D department’s work on Multimedia Classification, which primarily aims to develop novel methods for extracting metadata from a programme and using this to allow for programmes classification and retrieval. We have discussed our work on extracting affective, or mood, based metadata from the audio and video of a programme, along with some of our other work on identifying the mood of a programmes theme tune, the mood of a programme from subtitles, the combination of mood and semantics and finally our work on automatically identifying the highlights, or key points, of a programme.

The main proposed usage of this work is in identifying new ways to open up the BBC’s archive to the general public. Whilst there are good
methods already available for professional archivists to find programmes and for non-professionals to find a programme they know about and know exists in this collection, we are looking at ways to enable non-professionals to find programmes that they didn’t know existed, didn’t know anything about but would find interesting or entertaining to watch.

REFERENCES:


Sarah-Haye Aziz is the Head of the RSI’s television archives and in charge of the department’s internal training programs. Sarah-Haye has been involved in many projects and publications as co-author on SSR and RSI history books, with research interests in the fields of media history and relationship between media and cultural identity.

Lorenzo Vassallo is a technical advisor at the RSI and supports archivists as an interface with the IT department and external suppliers. He coordinates all the projects to maintain consistency and harmonise the Archives’ Systems with the other RSI Departments. He contributed significantly to the development of CMM, the cataloguing system of RSI.

Francesco Veri is archivist at RSI’s TV archive. He works on the documentation of historical video sources, films and documentaries with the goal of retrieving, recovering and preserving Swiss Italian’s audiovisual cultural heritage.
IN 2011 RADIOTELEVISIONE della Svizzera Italiana (RSI), the Swiss Italian Broadcast, introduced an Automatic Indexing System in its archive’s Multimedia Catalogue (CMM). The CMM catalogue is expanded each year by digitally storing RSI’s daily production and by recovering past material: about 300,000 hours are available for consulting in a catalogue with the corresponding written documentation. RSI’s focus was to facilitate information and data management in its television and radio archives. Since then RSI has regularly used automatic techniques in its daily workflow.

The aim of this article is to share our archivists’ experience of implementing an Automatic Indexing System into RSI’s archiving system. Despite the fact that SSR represents a small reality inside the broadcast panorama – and this applies all the more to RSI given that it is part of SSR; we consider that all the issues that RSI had to face based on its experience, follow certain patterns that could easily emerge in other larger broadcast organisations.
The aim of this paper is not to give a particular emphasis on “why” RSI decided to use an Automatic Indexing System but rather to explain “how” the project has been developed inside our archives by our IT staff and our archivists.

In other words, our aim is to understand what the introduction of automatic techniques means for archivists. In fact, we consider that it is interesting to see what all this has meant for the archivists when the workflow shifted from manual to automatic.

**AUTOMATIC INDEXING SYSTEM IN RSI**

Throughout this paper, the term Automatic Indexing System (AIS) will refer to both the automatic transcription of the audio source (or Speech to Text) and also the automatic semantic analysis (or feature extraction).

Since 2008, RSI has adopted a modern sustainable archiving system which includes a stable and consolidated workflow. By and large our catalogue multimedia system (referred to as CMM) was set up in order to support improvement and innovation, and so allows us to implement new technology such as AIS’s transcriptional and semantic analyses engine.

AIS was introduced for a practical reason, that being to improve RSI’s archiving system and compensate and assist archivist activities. As you can imagine, RSI as a national Radio and TV archive, has a huge quantity of audio-video documents. Due to mass digitisation, archivists are faced with a huge amount of content which needs to be catalogued. But unfortunately RSI does not have enough human force available to reflectively cover the management of all our video and audio documents. Therefore AIS is on the one hand an opportunity to support the manual process of recording, while on the other hand it is also of great use to increase archive productivity in data retrieving, handle the large amount of content and to make data accessible to various audiences. Finally, automatic tech-
niques are also useful for documenting large amount of digitalised old raw material.

**THE IMPLEMENTATION OF AIS**

The AIS project was set up between 2011 and 2012. By and large, before the introduction of AIS in the cataloguing system, archivists had to manually manage documents in every aspect of documentation until the documents were published in the CMM.

RSI advanced the implementation of the automatic indexing system in two distinctive steps: a) through the implementation of the new tool in the daily workflow after the accomplishment of a series of selective tests, in both Radio and Television documents; and b) through the archivist’s direct participation during the tuning phase.

In order to explain the implementation of AIS in CMM we will focus on “how” RSI has dealt with this project. The AIS project was developed in two different phases.

**PHASE 1: ANALYSIS AND START UP**

In the first analytical and start up phase, IT staff worked on the language model and on the workflow. A prototype was developed, and it became the basis for tests and analyses.

The language model followed two specific linguistic patterns: a general journalistic one and one based on sport journalism jargon. In order to build a coherent general model the prototype was adjusted on the Italian spoken in Switzerland, using a substantial number of written texts and articles coming from local newspapers. At the same time the daily workflow has progressively adapted in order to integrate AIS into the standard system.
At this stage the archivists were a bit concerned about this new technology, owing to the first test taken on the system RSI catalogue did not show observable graphic distinction between human and automatic indexing and the overall quality was not at the best because of the lack of tuning. For this reasons, archivists were worried that the future catalogue users would not be able to distinguish between what the machine did and what the human did (as a consequence of a negative impact on archivist job).

Moreover, the automatic cataloguing was a bit inaccurate in respect of human indexing. In relation to this, one of our major concerns was that archivist role would eventually become marginal in respect of the Automatic Indexing System, and be reduced to a AIS “quality checker”.

Hence, archivists observed that this prototype was not related to functionality, and was definitely not ready for their requirements.

PHASE 2: TEST AND TUNING

In the tuning phase the whole system, and in particular transcriptions, were tested and tuned. At this stage IT staff and archivists pondered whether or not to modify the normal workflow, and identified which programmes were eligible for speech recognition. At this point archivist started to play a crucial role in selecting material and evaluating what was best for an Automatic Indexing System. The Archivists new role was to decide which TV and Radio programmes should be processed with AIS.

Once this process was set up, archivists had been directly involved in the project, not only during the selection of documents, as mentioned above, but also through a proactive and synergic collaboration with technical and computing staff in finding new innovative solutions. The aim was to satisfy user requirements and overtake archivists’ fears towards the new technology.

For the specific cases previously mentioned, archivists suggested that the IT staff use different colors in order to distinguish between human and
automatic indexing. So the automatic indexing was highlighted in orange and human in blue.

Furthermore, archivists also wanted to have more decision making power on the workflow with the possibility to change some automatic indexing parameters; such as to decide whether or not to proceed with automatic credits extraction of the text. This allowed them to reassess the central role of archivists. As a result, the fear that archivists would simply become “quality checkers” completely vanished.

It is important to point out that the project would not have been successful during the second phase, if there was not a synergic collaboration between archivists and the IT staff. Aside from the practical reasons of this collaboration, an essential point was the psychological aspect. In fact, when archivists acquired a participating role in the development of the
project they also had the opportunity to experiment and gain valuable hands-on experience of AIS, in addition to achieving awareness of their upcoming co-existence with the new system. In other words, archivists fully realised that an Automatic Indexing System does not represent a threat to their role and they better understood the instrument’s boundaries (what it could and could not do).

**KEY LEARNING POINTS**

So at this stage one is naturally led to wonder: does this system work well? The answer is: yes but...

The system definitely works from a workflow perspective. The imple-
mentation of AIS did not raise any particular problems, and archivists are able to easily learn this tool because it is incorporated in the previous workflow (the one with which the archivist is already familiar). However, it is important to consider that (such as is with every technology), AIS systems need maintenance and constant updates. So it is necessary to invest money, time and experience for these purposes.

Notwithstanding the positive results, it is important to point out some difficulties faced in respect of the choice of the document typology (TV or Radio).

Firstly, Radio and TV documents are different in programmes setting. Often radio documents have noise or music in the background, which interferes with the automatic transcription. Moreover there are differences between radio and TV on logical automatic sequences creation. The key frames in video help to locate a change of context and so the creation of logical sequences in TV documents. Contrarily, the sequences in Radio are engendered from silences, with the result of the creation of too many sequences.

Secondly, Radio and TV are different typologies of media with different narrative language, due to a diverse significance of speech pauses during programs.

These two differences affect the automatic analyses of the audio in Radio and TV. AIS in TV uses different logical schemes than in Radio. TV and Radio are different and sometimes it is hazardous to use one common technical solution for both media. So, whether or not working with Radio or TV, it is important to have different approaches and find different solutions.
CONCLUSION

As a result of this experience we have achieved a good level of awareness and knowledge about the advantages and disadvantages of AIS, as well as further hypothetical areas of improvement.

From a technical perspective, this system could allow us to open up new horizons. In fact there are a certain number of specific applications that can be engendered by the incorporation of an automatic indexing system to an archiving catalogue. For instance, it could be possible to realise a new workflow that can exploit a semantic engine for automatic metadata extraction starting from written texts (i.e. editorial texts, web texts) to the main database CMM. Once this procedure has been completed, it would be possible to integrate semantic analysis of written texts to their related video/audio object.

From a human perspective it is essential to involve archivists in the project. Despite the importance of technological aspects we could assert that computing scientists have not the full control of the technical developments. Archivists can contribute to new solutions and ideas. Involving archivists is also important to them because it allows them to better redefine their roles inside of the archive.
5. USER GENERATED METADATA
Yves Raimond holds a PhD from Queen Mary, University of London. His thesis was entitled ‘A Distributed Music Information System’, and defined a framework for applying a range of Semantic Web technologies for managing and distributing music-related information. As part of his thesis, he contributed extensively to what would become the ‘Linking Open Data’ community project.

Since 2008, he has been working for the BBC, first on the bbc.co.uk/programmes service, publishing structured data about all BBC programmes, and then in BBC R&D on the ABC-IP Technology Strategy Board.
5:1 THE BBC WORLD SERVICE
AN ALTERNATIVE APPROACH TO PUBLISHING LARGE ARCHIVES

THE BBC HAS been broadcasting from 1922 and has accumulated a very large archive. Opening this archive and enabling search and discovery within it remains a big challenge.

BACKGROUND

Items within this archive have been catalogued using a number of systems throughout the years. The BBC classification system is called Lonclass, and the cataloguing effort has been geared towards re-use. Unfortunately the coverage of the catalogue is not uniform across the BBC's archive. For example it excludes the BBC World Service which has been broadcasting since 1932.
Traditionally, when it has come to publishing archive content online, the BBC focused on specific topics or brands. An example of that is the Radio 4 In Our Time archive giving access to all In Our Time episodes online. Another set of examples are the BBC Four collections, focusing on particular topics rather than brands.

In order to drive topic-based navigation within these collections of archive content, programmes are manually tagged with identifiers from DBpedia, a large database extracted from Wikipedia. The resulting tags are web resources, pointing to datasets published within the Linked Data cloud. A benefit of using Linked Data identifiers as tags is that they are unambiguous and that we can retrieve more information about those tags when needed. For example, programmes tagged with places can be plotted on a map, or topic-based aggregation pages can be enriched with information about the corresponding topic.

The process of manual tagging is very time-consuming and would take a considerable time to apply to the entire archive. The problem is compounded by the lack of availability of textual metadata for a portion of the archive, and programmes for which we have no data or worse, wrong data. These programmes will be left out of such brand or topic-specific publications of the archive.

In the rest of this article, we describe an alternative approach to publishing archives online. This approach consists in purposefully exposing large archives with little or no pre-existing metadata using machine-generated data that can be inaccurate and involving users in helping us correct that data.

An interesting test-case for this approach is the BBC World Service archive. This archive has been left out of the cataloguing and tagging efforts mentioned previously. However all pre-recorded content within this archive has been fully digitised. For the English language part of the World Service, the archive holds around 70,000 programmes since 1947, which amounts to about three years of continuous audio. The data around those digitised programmes is very sparse and sometimes wrong. But as mentioned before,
AUTOMATICALLY TAGGING ARCHIVE CONTENT

Can we use the content itself to bootstrap search and navigation within such an archive? Of course listening to and annotating all of this archive would take a considerable amount of time, but can we automate this process? In particular, can we automatically tag content with DBpedia web identifiers?

The first step towards understanding what programmes are about is to figure out what was said in the programme. We use the Open Source CMU Sphinx toolkit for performing automated speech recognition. The automated transcripts are very noisy. However, they include useful clues as to what the programme is about e.g. proper names, locations, organisations...

We need to isolate these useful clues and infer the topics of the programme from them. Most of the existing concept tagging tools are designed to work on text that was manually written and rely on punctuation, capitalisation, etc. We therefore developed our own tool for automated tagging from noisy transcripts. We use the structure of the DBpedia graph to disambiguate and rank keywords spotted throughout the transcript. For example, if a programme mentions Paris and Tour Eiffel a lot, we will pick Paris in France, as it is closer in the DBpedia graph to the Tour Eiffel resource. If a programme mentions Paris and Texas a lot, we will pick Paris in Texas, as it is closer in the DBpedia graph to the Texas resource. For each programme, we get a ranked list of DBpedia tags, describing what the programme is about.

Figure 1 (next page) shows a couple of example results. The first programme is a 1970 profile of the composer Gustav Holst. The second programme is a 1983 profile of the Ayatollah Khomeini. The third programme is a 1983 episode of the Medical Programme.

We evaluated our algorithm on a dataset of 132 manually tagged programmes. We use the TopN measure. This measure was originally devel-


<table>
<thead>
<tr>
<th>TAG</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programme 1</td>
<td></td>
</tr>
<tr>
<td>d : Benjamin_Britten</td>
<td>0.09</td>
</tr>
<tr>
<td>d : Music</td>
<td>0.054</td>
</tr>
<tr>
<td>d : Gustav_Holst</td>
<td>0.024</td>
</tr>
<tr>
<td>Programme 2</td>
<td></td>
</tr>
<tr>
<td>d : Revolution</td>
<td>0.037</td>
</tr>
<tr>
<td>d : Tehran</td>
<td>0.032</td>
</tr>
<tr>
<td>d : Ayatollah</td>
<td>0.025</td>
</tr>
<tr>
<td>Programme 3</td>
<td></td>
</tr>
<tr>
<td>d : Hepatitis</td>
<td>0.288</td>
</tr>
<tr>
<td>d : Vaccine</td>
<td>0.129</td>
</tr>
<tr>
<td>d : Medical_research</td>
<td>0.04</td>
</tr>
</tbody>
</table>

*Fig 1. Table of example results*

oped to evaluate automated music tagging systems. This measure will be 1 if the top-N tags returned by our algorithm exactly match the N manually added tags. The measure will decrease as manually applied tags appear further down the list of automated tags. We compared our algorithm with a baseline random tagger and with a couple of off-the-shelves concept tagging tools and our algorithm performed best. This is understandable as those concept tagging tools are meant to work on manually written text, not noisy automated transcripts. We also published our evaluation dataset and scripts on our Github page, including the results of our automated transcription for that evaluation dataset. Other institutions are currently working with this dataset to evaluate other automated tagging algorithms. 

Processing large archives remains a challenge. It would take more than 4 years to automatically tag the entire World Service archive with this
algorithm on commodity hardware. We therefore developed a framework to process very large archives using Amazon Web Services. It took around two weeks to process the whole archive, for a low and predictable cost ($3000 for three years of audio).

THE WORLD SERVICE ARCHIVE PROTOTYPE

After running that process we had a set of descriptive topics for all programmes in the World Service archive. This automated data can be used to bootstrap search and navigation within the archive, letting users discover and listen to programmes that often weren't listened to since they were last broadcast. We therefore built the World Service archive prototype.¹
However, as with all automated processes, the resulting data can be wrong. In order to deal with this noise we built some features within the prototype enabling users to validate or invalidate the automatically extracted data. In particular, users can vote tags up or down, to approve or disapprove them. This feedback is used to make the search and navigation within the prototype more reliable, and can be fed back directly in the automated tagging algorithm to evaluate and refine it.

We also segment the audio depending on who is speaking. This enables users to get a quick overview of a particular programme and to jump to specific points in the audio.²

---

1. The World Service Archive is available at: http://worldservice.prototyping.bbc.co.uk.
However after segmenting the audio we know that individual speakers are contributing to the programme, but not who they are. We built a simple mechanism enabling users to name those speakers. This mechanism is also built on consensus. The name displayed by default will be the name chosen by the most users. We are also able to recognise speakers across programmes. Therefore the names added by users can be propagated to other programmes in the archive detected as featuring the same speaker.

As for user feedback on tags, we can use the speaker names as a basis for evaluating our speaker identification algorithm. For each pairs of speakers with the same added name, did our algorithm think they were the same?

Fig 4. Screenshot of the World Service Archive Prototype website

---

2. An example is: http://worldservice.prototyping.bbc.co.uk/programmes/X0558256.
This enables us to evaluate our speaker identification algorithm on a representative task. With the speaker names added by users as of May 2013, we get a precision of 85 percent and a recall of 54 percent.

Since the launch of the prototype in August 2012 we have been progressively adding more and more users. We have had more than 60,000 user edits on the prototype as of May 2013. We are currently monitoring the average quality of the data to see how it improves with more and more user edits. A detailed report on the evolution of data quality remains future work. A large number of topics will be covered in some form by the programmes held in this archive. In particular, the archive may hold programmes that could provide context around current news events. For example a 1983 “Medical Programme” episode on techniques for measles immunisation could help put in context a recent epidemic.
We built a dynamic visualisation of content within the archive that relates to current news events. This visualisation constantly listens to the BBC News channel and pulls out content from the archive related to the topics being discussed. The big blue dots are topics that were discussed on BBC News in the last five minutes. The small dots are programmes within the archive that relate to those topics. The redder a dot is, the more likely it is to be relevant to a current news event.3

CONCLUSION

In this paper we described an alternative approach to publishing large archives with little or no pre-existing metadata. This approach consists in purposefully exposing large archives using machine-generated data that can be inaccurate and building mechanisms for users to help us correct that data. We are currently evaluating this approach on the BBC World Service archive. We derived automated tags for all pre-recorded programmes in the English language part of the archive. We also derived speaker segmentations and identified speakers across programmes. We used the automatically derived data to bootstrap search and navigation in a prototype website: http://worldservice.prototyping.bbc.co.uk. As automatic data can be incorrect, we built features within the website enabling users to validate or invalidate the data. This user data enables us to continuously improve search and navigation within the archive and to evaluate and refine our algorithms.

3. Visualisation available at http://worldservice.prototyping.bbc.co.uk/visualisations/current
Liam Wylie is the senior curator and content producer at RTÉ Archives where he is responsible for the website www.rte.ie/archives. He has worked as a film and television archivist and is a former head of collections at the Irish Film Archive.

Through his own production company Red Lemonade Productions he has independently produced and directed archive based documentaries for RTÉ television.
THROUGH THE WEBSITE “www.rte.ie/archives” the archives of the Irish national broadcaster RTÉ look to engage with their audience in a number of different ways. RTÉ Archives have been combining use of their own website and Twitter to ask the public for help in identifying people and places from the photographic collections. If crowd sourcing is thought of as a way of presenting your audience or users with a problem and asking for their help to solve it, then that is what RTÉ Archives are doing. This article looks at the thinking and the results behind this approach to user engagement.

RTÉ ARCHIVES

RTÉ Archives house the largest audiovisual collection in Ireland. The ar-
chives are responsible for collecting, preserving and making accessible the creative and documentary output of the national broadcaster. Combining hundreds of thousands of hours of moving image and sound recordings together with significant collections of photographs, manuscripts and administrative documents, RTÉ Archives contain a unique record of Irish life.

Our role is to manage, develop and open up the national resource that is RTÉ Archives.

USE THE CROWD

Technology is providing organisations charged with looking after large audiovisual collections with the potential to improve the information about their content. Elsewhere in this book you can read about clever combinations of technology and human interventions that look at improving the creation of new metadata and adding value to information that has been created in the past. If we are linking, creating or enhancing metadata the ultimate aim is to help improve the management and access to audiovisual collections.

Crowd sourcing is a term that has been with us for a number of years now. Usually an online public is asked for some kind of help or input to a project. When it comes to improving metadata within a catalogue, some information (or lack of information!) that would usually be private, is exposed to the public in the hope their knowledge will improve or correct the recorded information.

Appealing to the crowd for help creating catalogues or improving existing records has proved successful for a number of organisations. The National Library of Australia asked the public for help with tagging and corrections in a newspaper digitisation project, while the Netherlands Institute for Sound and Vision created ‘Waisda’, an annotation game that
collects user generated metadata for television programmes.

At a time when resources are tight, large scale crowd sourcing projects maybe as out of reach as some of the wonderful, but expensive technological solutions being offered to enhance metadata.

A more obvious problem is that although broadcasters have audiences most public service broadcasters are not public facing. The archives and collections of these organisations have been created to serve their programme makers. So how do you use the crowd if your collections are not available to the public?

RTÉ ARCHIVES ONLINE

RTÉ as a public service broadcaster is aware of the interest in the content that is held in its archives and is committed to finding ways to increase public access to these collections. This process began with the creation of the website www.rte.ie/archives.

The website works as an access point for the public to the RTÉ Archives. It is where we tell people what we do, provide contact points to our sales teams and have begun to make content available. There are over 100,000 photographs available to the public to browse and search online.

All collections are drawn on by a small team to create online exhibitions offering visitors the opportunity to look at and listen to content from the RTÉ Archives. To date there are 24 of these exhibitions, a recent example is a look at the 50th anniversary of the visit of President Kennedy to Ireland.¹ It also allows a place for RTÉ to share some of its own history.²

We also try to offer our audience new content on a daily basis by creating

¹ www.rte.ie/jfk
² http://www.rte.ie/archives/exhibitions/681-history-of-rte/
stories in a features section. Here the content can tie into something that is date specific. On this day in 1971 members of the Irish Women's Liberation Movement travel to Belfast by train to buy contraceptives as a protest against the law which forbids the importation and sale of them in the Republic of Ireland.³

Where we can link archive to contemporary news stories is another way to try and connect with the public. Horse meat is rarely sold or eaten in Ireland so when in January 2013 traces of horse meat were found in beef burgers we posted a story about a butcher selling horse meat in Dublin in 1966.⁴

---


*Fig 1. Screenshot from the RTÉ Archive website*
This is how we have begun at RTÉ Archives to make more of the content we look after available to the public. Figure 1 shows the website, the public face of RTÉ Archives. Having created that public front, we want to connect and build our audience. Like any broadcaster one of the things we talk a lot about in RTÉ is finding ways to let people know about what we have and what we do. We are no different in the RTÉ Archives – we want people to discover the great things we look after for them. Creating a connection with our audience or the public is what challenges and interests us.

Social media is essential now to connect with the public and to help spread the word on what our work is about and what we are trying to achieve. As we have a small team to carry out this work Twitter @RTEArchives has been our chosen social media tool.

**CAN YOU HELP US WITH...?**

So what has all of the above to do with user generated metadata? Having created a public front for RTÉ Archives, we want to engage with our audience and to carry on growing that connection. We cannot have a “build it and they will come” approach, we need to continue to talk to our audience and offer new and interesting content as often as possible.

One regular item we have been running is to take a photograph from the RTÉ Archives which has incomplete or no catalogue information, and ask the public if they can help? Every Tuesday a photograph is published as feature story on “www.rte.ie/archives”.

We then tweet a link back to the story. In a second tweet if we don’t have information we will include the photograph itself. If people have information, they can tweet it or send us an email.

Figure 2 is the first photograph we posted as a feature with the missing name Gerry Donovan, which was supplied by members of the public.

Figure 3 is another early example of a photograph that created a con-
conversation. We asked for help in identifying this pop band playing in the RTÉ television studios in 1983. We tweeted about it on 3 September and for days after people were engaged through Twitter trying name this band.

A week later and as you can see below plenty of Twitter chat but still no answer. Actually, to this day we still do not know the name of the band. At the time @RTEArchives was new to Twitter but this photograph was getting people talking and was being re-tweeted by some large accounts, all helping to make people aware of RTÉ Archives.

We do not get an answer every week. Indeed some weeks we get nothing. Any information received is passed on to the cataloguing team who update the photographic records. In January 2013 we asked if anyone could help us identify the people with the Reagans in a photograph taken during the American president’s visit to Ireland in 1984.
In early February we got this great email sent to RTÉ Archives giving almost complete information. Only housekeeper Lil’s surname is missing, and this was given to us soon after!

**BENEFITS**

This approach has worked on two fronts for RTÉ Archives. Asking the public for their help means we avail ourselves of a wider pool of knowledge, while at the same time connect with our audience. This interaction allows us to add or amend information within the photographic catalogue and the public get to see some of what we do. From the 40 photographs we have asked for help with, we received information for 22 of them.
These figures may seem small, but additional traffic has been driven to the RTÉ Archives website and to the online photographic collection. Interestingly, members of the public have contacted RTÉ Archives with information about photographs that are available to view online but we had not directly sought help with.

A further benefit harder to quantify is the engagement from archive staff who are seeing the advantage in making images available in this way and recognition from the public for the work they carry out on a daily basis.

RTÉ Archives sit within RTÉ Digital which is responsible for the broadcaster’s online and digital platforms. The creation of an online presence for the archives has allowed a direct engagement with the public. The digital space gives the public access to content from the archives and gives RTÉ Archives the chance to let our audience know what it is we do. The approach to facilitating some user generated metadata maybe small, but is resource light and relatively inexpensive.

The key challenge for public service broadcast archives used to serving their programme makers, is to find ways to connect directly with the public while continuing to carry out their core activities. If an archive can find a way to do this, then it can tap into the knowledge held by the public. An online presence, combined with a regular engagement with your audience, offers the possibility of availing ourselves of user generated metadata. Clearly this is not the panacea for all the ills of information deficits within our catalogues, but it is an interesting additional resource that works.
6. THE EPILOGUE
Brecht Declercq holds a Master’s Degrees in Contemporary History and International Relations. Since 2004 he works for the archives department of Flemish public broadcaster VRT. He assisted in the development of the MAM and in several archival disclosure projects. In 2008 he took a half year break at VRT to research the future of contextualisation for FARO. Since 2008 he leads the digitisation of VRT’s radio archives in the DivA and VErDi projects.

Main tasks now are also on the international level, coordinating VRT’s contribution to EUscreen.eu and the strategic one, following-up Flanders’ AV heritage policy. As a FIAT/IFTA MMC member, he’s an active voice on the international AV-archival scene.
6:1 CONCLUSIONS
METADATA IN AUDIOVISUAL ARCHIVING:
THE HOURGLASS MODEL

AT THE END of another very inspiring FIAT/IFTA MMC Seminar the moment has come for some final remarks. In this sixth edition of the series 'Changing sceneries, changing roles' we were inundated again with fascinating findings, provocative statements, and creative, novel ways to deal with the challenges that are the daily struggle for audiovisual archive staff.

At the end of two fertile days of pitching ideas, it's good to canalise the maelstrom of thoughts. To help you doing so, dear reader, I'd like to take you back to the preparations of the conference.

In the spring of 2012 Annemieke de Jong, a longstanding commission member to whom I owe so much of my knowledge about audiovisual archives, posted an inspiring document, featuring a large, red brick. Each of the four corners of the brick held a particular type of metadata: user-generated metadata, linked metadata, preservation metadata and automatically generated metadata. The document was entitled “Metadata.
Cornerstone of digital archiving.” Like all good ideas it owed its strength to its simplicity.

**PEACE OF MIND IN THE METADATA REALM**

In the constructive atmosphere so typical for the MMC, Annemieke’s proposal gained prompt and wide acclaim. I as well saw the merit of this rock solid idea, even though I somewhat struggled with it and it took me quite a while to get used to. Was it the linked metadata, so new and promising, of which I had realised the impact only recently? Or was it the preservation metadata, which I was less familiar with, but which undeniably take a place of utmost importance in the audiovisual archivist’s field?

Comparing the four types of metadata, it occurred to me that they aren’t created at the same time. Limping further on that train of thought I tried to come up with a new scheme. Perhaps inspired by a visit to Paris, who knows, a strange crossover between the Eiffel Tower and the Grande Arche de la Défense came to mind. A triangle merged with a square, which under gentle pressure of the graphic designer of this book, I prefer not to show. Like all bad ideas it owed its weakness to its complexity.

It was Oscar Rishede, archivist at Danmarks Radio (DR), who first brought logic and peace of mind. Whilst preparing a workshop in Copenhagen, he asked me to talk about metadata: “you know: how they are made, what they do.” With the consent of the Media Management Committee I would like to share this logic with you.

The scheme in figure 1, in which someone sometimes recognises an hourglass, describes the place that metadata holds in contemporary audiovisual archiving. It grew out of the question how all these different types of metadata that are being talked about so much, relate to each other. Preservation metadata for example, are sometimes automatically generated. Still no one will ever argue that they are the same. The somewhat
straightforward confrontation of trendy terms seemed to have caused my confusion. Therefore I hereby put them on different levels:

- The creation level at the top describes the way metadata are created.
- The purpose level at the bottom describes the goals that metadata serve.
- The processing level in between is where the connection between the two aforementioned is made.

*Fig 1. Creation, purpose and process*
At the creation level four sources of metadata can be distinguished: the so-called production metadata, user generated metadata, automatically generated metadata, and the old school full manual annotation. Evidently these sources correspond with four metadata creators: the maker of the essence, its user, the archivist and his computer.

At the purpose level we find the objectives for which these metadata are used: preservation and collection management (evidently connected to the term ‘preservation metadata’), search and retrieval (a classic focus in the archival world) and enhancement and contextualisation (where the process of linking metadata can also be situated).

**NEW COMBINATIONS AHEAD**

Any of the four ways of creating metadata can be combined with any of the three purposes. The model thus results in twelve interesting combinations, some of which have been known for centuries (using full manual annotation for managing the collection for example), and others were born only recently. Most of the case studies that have been presented at the Seminar easily fit in this model.

For example in the “Semantic Linking for Second Screen” project by VARA and the Amsterdam University, keywords drawn from automatically generated metadata are linked with external information, to disclose it in an enriched, contextualised form. In Switzerland a long predicted innovation that seemed to be forever stuck in the research phase, has finally reached daily practice: speech recognition, a form of automatically generated metadata, is used to make RSI’s archive better searchable.

In Sweden, SVT is sedulously searching for a viable way to generate metadata that can assist the Swedish Royal Library (KB) in preserving the web content, a task rooted in the renewed legal deposit legislation. It is obvious that production metadata will be invoked in that process. As the
final presenter at the Seminar, RTÉ’s Liam Wylie reminded us that high tech is not always at stake. RTÉ Archives simply use very accessible means to involve the user (in casu Twitter and Vine) to help cancel old question marks from their catalogues, thus increasing searchability. A closer connection with their audiences is a positive side effect proving to be just as important as the original scope.

This hourglass scheme also allows to think about new and creative ways of creating metadata in order to preserve and manage the collection, to search through it, to contextualise or to enhance. Allow me to illustrate this with only one example of a combination rarely used in audiovisual archives: given the high cost to have files coming from mass digitisation projects checked manually, and also given the fact that technology is not yet far enough to do this automatically, one could consider involving the users (e.g. the general audience) in this task. This way user-generated metadata could be connected to quality control, an important part of collection management and preservation.

But let’s take this even further, by making multi-combinations. Indeed, one can also connect one of the ways of metadata creation with multiple purposes. Or one can have multiple ways of metadata creation combined to serve one (or even more) purposes. Yves Raimond's presentation on the BBC World Service Archive Prototype was a wonderful example of the latter (which from a practical view left the seminar programming committee with an easy job, as his presentation could be embedded in multiple sessions). In its essence this BBC project uses no less than four types of metadata creation to serve two purposes: making the collection searchable and enhancing it with related information.
A TOOLBOX FOR THE FUTURE

It is my firm belief that an important part of future audiovisual archivists’ jobs will be the quest for the right combinations: which purpose to serve and which way to create the appropriate metadata. The nature of the collections and the resources available will be the main side conditions in this process. It will be up to the archivist to connect sources and purposes, by inventing flexible solutions. Although it is likely that traditional manual annotation will still survive for a while, good decision making process management and quality checking will become ever more important skills.

The archivist used to have only the full manual annotation method available, which I’d like to compare with just a screwdriver sometimes even used as a hammer or a knife. Soon however, this multi-purpose goody will be exchanged for a whole fancy toolbox. Then the archivist will come to know and fully appreciate the strengths and weaknesses of each tool and be able to make informed decisions which one to use. Since the archivist will manually intervene less, it will also be more important to monitor the quality of the processes well. Ultimately a good overview of the overall management of the processes will be required, because the ever faster growing collection not only has to be made searchable, enhanced and contextualised, but also has to be managed and preserved.

That’s why I would like to join in with the final words of the conference, as pronounced by FIAT MMC chair Eva-Lis Green: “it's about advocating the long term in a short term world.”
“The future isn’t either traditional or digital: it’s a feedback loop between the two. Television fans want to get involved and be counted. It’s how creative we are in engaging those fans – and keeping them connected even as they may move away from the traditional network – that will determine how potent and profitable we will be in the future.”

– Kevin Reilly, President of Entertainment, Fox Broadcasting (as quoted by Mainwaring, 2012)

With the proliferation of digital media in both the private and public arenas, a great tide of digital content is upon us, one that requires careful management and preservation. Without early intervention, valuable digital assets may be lost. It is common knowledge that benign neglect (i.e., the strategy of shelving a book for a long period) is not a valid strategy
for digital resources; this plan eventually places digital media at risk of technological obsolescence. Much effort has gone into developing realistic strategies to combat this risk, including emulation and migration. Metadata supports these strategies by explicitly describing the context, content, structure, and management of an object to ensure its access over the long term.

Foremost among the seminar’s trending subjects was the issue of how to bridge the gap between archivists and Information Technology (IT) staff. Many participants were concerned over the cultural differences between the two groups; IT is known to embrace pioneering technology while archivists are reputed for relying on tried-and-tested tools for solving problems. Florian Delabie (RTBF) embodied one possible answer to this concern, being both skilled in IT and media archiving. During the seminar, Delabie asserted the important role of the archivist in ensuring quality metadata while claiming that early user adoption is far more likely when IT involve archivists at an early stage and demonstrate the benefits of new application software. Although Beth Delaney (audiovisual collection management consultant) and Sandra Collins (Digital Repository of Ireland) agreed that communication between these groups is vital, Collins claimed they must jointly set goals and responsibilities at the beginning of a project. Many of the presentations raised the pertinent question of who should be responsible for metadata input. The various approaches presented (automated, user-produced, and semantically linked) are testament to the degree of thought the field is giving to this issue. The proceedings also demonstrate that the archivist's role is much more intricate than simply managing media. It continues to be about preserving at-risk digital assets while ensuring they continue to be accessible into the future.

Jan Müller, president of FIAT/IFTA and head of the Institute for Sound and Vision, opened by welcoming the attendees to “metadata heaven”. In response to the momentum of Linked/Social TV, Müller surmised that all television would soon be interactive, shared, and created by amateurs.
The key to collecting and managing this material, he argued, is metadata. Next, Seth van Hooland (Digital Information Chair, Information and Communication Science department of the Université Libre de Bruxelles) offered a comprehensive keynote about LinkedData, in which he argued for ‘clean’ RDF (Resource Descriptive Framework) – structured metadata, but cautioned archivists against sharing this metadata with the public unless practitioners intend to maintain it for the long term. He emphasised the importance of moving towards standardised global taxonomies and to explore the business value of added-value metadata on the web. Mine unstructured data and identify what you can get for free but always question the sustainability.

Video distribution must change in response to the way its users behave. To meet this challenge, Rutger Verhoeven (VARA) and Maarten de Rijke (University of Amsterdam) guided the audience on a tour of the Social Video Live Player, a LinkedTV player that suggests related programs for viewers to watch after analysing program subtitles, generating keywords and linking them, by means of the semantic web, to similar content on the web. TV becomes social and fragmented as people share and decide what to watch from friends using social media networks on multimedia devices, mobiles and tablets. The Second Screen is opening up all sorts of opportunities for content producers to push related live and archive content to audiences in new and novel ways. This is the start of a new journey with many unknowns of how the end user will interact and consume related content. One certainty is that the proliferation of content in new interactive formats means more pro-active archiving!

Xavier Jacques-Jourion (RTBF) walked us through a semantic based multimedia browser called GEMS, a proof of concept currently under development. As with the Social Video Live Player, GEMS extracts conceptual terms from the program audio and maps them to the semantic web using automated tools. GEMS is a unique tool which maps spoken terms to broader and narrower concepts, allowing multiple paths to dis-
covery. During the question period, audience members asked whether the solution would scale with large amounts of content. Jacques-Jourion posited that tailoring the number of conceptual terms to the user, using a profile, would help match users to the content they desire. Another question focused on how one could quality check the veracity of the links and domains with the response that you can moderate the input. There were some positive sound-bites “Metadata is not interesting it is what you do with it that is interesting” and “Opens eyes of the documentalist to see what technology can do”. During the first panel discussion of the day, Seth van Hooland stated that the Linked Metadata ecology is in an extremely fragile state; methods have yet to be developed to ingest and preserve Linked Metadata. Therese Nilsson (SVT) explained that while the Swedish public service television company's user tags usually developed into folksonomies, the broadcast station's journalists vet terms so that some measure of control is imposed on the user-generated vocabulary. Jennifer Wilson (BBC Scotland) argued that archivist-generated metadata should be enriched with Linked Metadata, but acknowledged that the needs of archivists and users are not always aligned. She used the analogy of a “Bionic Man” in that librarians add the value of proper data but also can have a symbiotic relationship with Linked Open Data as both need each other. Trust and Quality is the basis of trusted datasets and certification.

Seth van Hooland added that archivist metadata and Linked MetaData should be kept separate; the way information is presented to users should be up to the interface designer. Brid Dooley (RTÉ) offered that linking cultural heritage metadata from different institutions may be a worthwhile endeavor. The question of opening up broadcast archives to the public was debated with the conclusion that everyone recognised the benefits and would like to make available all of their content. We were reminded of some of the issues of data cleansing and the mass volumes of content in our collections which needed to be digitised.
Karin Bredenberg (National Archives of Sweden) gave a keynote about preservation metadata, offering attendees a broad description of the PREMIS (Preservation Metadata: Implementation Strategies) guidelines and data model and the OAIS (Open Archival Information System) reference model. Summarising these complex frameworks is difficult, but Bredenberg’s presentation was generally successful; she used the metaphor of food storage in a refrigerator to describe the OAIS concepts of SIP (Submission Information Package), AIP (Archival Information Package), and DIP (Dissemination Information Packet). Concluding with a brief description of areas that merit further research, Bredenberg conceded that the audiovisual domain is relatively new to the PREMIS group, but encouraged broadcast media archivists in this field to implement the guidelines because it is a widely used standard in the digital preservation sector. She also stressed the importance of IT and Archives departments working together, the need for technical and structural metadata for audio/video and to increase the use of preservation metadata.

Hot on the heels of Bredenberg’s keynote, Daniel Steinmeier (Beeld en Geluid) launched the second group of case studies by describing Sound and Vision’s work to become a Trusted Digital Repository, an international certification based on the OAIS reference model that has yet to find an official independent body to perform an audit. His presentation concluded with several lessons learned from the experience so far: involve other staff in your organisation, articulate staff roles and responsibilities, acknowledge that no single IT application can meet all of an archive’s functional requirements, conduct a risk assessment, and administer a cost/benefit analysis.

Kaisa Unander (National Library of Sweden) and Eva-Lis Green (SVT) co-presented a case study in which they discussed the strategy employed by the National Library to comply with new federal legislation that obligates it to archive web content from national media syndicates (i.e. newspapers, broadcasters). Unander and Green explained that the 'designated
community' was unspecified in the legislation, making the development of preservation plans more challenging than it would have otherwise been. SVT adopted an option strategy (see Blue Ribbon Task Force Report) to preserve the digital objects until further work can be done to define the intended future users. Due to the fact that there is no legislation for either Re-use or Rights it will be a “Prison for Files!”

In the day's second panel discussion, speakers tackled the question of establishing what constitutes sufficient metadata capture. Elena Brodie Kusa (FIAT/IFTA MMC) proposed mapping existing metadata to the set of PREMIS elements to determine the completeness of existing metadata. The issue of bridging the gap between IT and archive departments arose again and there seemed to be a consensus that collaboration is essential and must be supported by management. Sandra Collins echoed an earlier statement made by Steinmeier that both groups must first define their goals and efforts at the beginning of any collaborative effort. While archivists were well represented in the conference program, computer scientists comprised less than a quarter of the total participants. Beth Delaney asked the question “How much preservation metadata do you need?” For Integrity and Authenticity you need to build in all the actions and create an audit trail for long-term access. Some pertinent sound bites concluded the first day:

“You cannot decouple Preservation and Access” and “Education and Training is important to create awareness and understanding of requirements.”

**OBSERVATIONS – DAY 2**

During the second day of the conference, participants continued to explore computer-driven and crowd sourced approaches to tag audiovisual assets. Although there is tension between these two methods, many par-
Participants argued for a combination of both. In spite of the challenges, solutions are already underway and were introduced as case studies at the conference.

In the day’s first keynote, Cees Snoek (Intelligent Systems Lab, University of Amsterdam) described artificial learning whereby a computer is trained to automatically generate accurate sentence-level descriptions of video content derived from shapes contained in the video frame. This research project offers an evolutionary next step in future automatically generated metadata. Just imagine a machine automatically generating whole sentences describing the scene or event! Although further academic research is required, no doubts in a few years’ time this could be a reality.

The keynote was followed by a case study from Sam Davies (BBC R&D), in which emotional terms are mapped to video content, effectively increasing user accessibility when combined with conventional discovery tools. This research prototype is currently being developed to provide the general public new ways of searching open archive collections to find programmes and related content based on mood. The second case study, from Radiotelevisione Svizzera – RSI Archive, documented a successful collaboration between archivists and IT staff to implement software to automatically tag television and radio content by means of speech to text computer audio analysis. Sarah-Haye Aziz highlighted the work of the archivists in the design, testing and implementation phases which helped to identify the advantages and disadvantages of using speech recognition technology in the cataloguing of television and radio programmes. In a lesson in Change Management – her team of Archivists were involved from the start of the project that helped to design the new workflows, and more importantly to redefine their own roles.

In the first panel discussion, the participants acknowledged the challenges disruptive technologies pose to archivists. While Brecht Declercq (VRT) suggested addressing archivists' reticence at adopting new tools through better interface design, Alberto Messina (RAI) argued that cul-
tural differences between IT and archivists hampers project development. Hypothetically, the IT team might scope out software applications without involving the archivists and the archivists may not be willing to compromise on software functionality. The panelists agreed that automated tools, by freeing up archivists' time, would increase the quality and quantity of metadata. However, Brecht argued that careful choices must be made to match the best tool to the collection. Google was identified as one source of rapidly developing technology that could possibly disrupt current audiovisual workflows (N.B. During the conference, Google announced its automatic tagging service, which linked identifying terms to its Knowledge Graph).

The second half of the day focused on user-generated metadata, often described as crowdsourcing. In a keynote address, Lora Aroyo (Intelligent Information Systems Web and Media Department of Computer Science, at the Free University Amsterdam) posited that, due to the symbiotic relationship between humans and machines, a new infrastructure must be built around the ways users consume audiovisual content. Aroyo suggested a holistic approach incorporating user experience design, quality metrics, user motivation studies, and the use of flexible tagging schemes. Yves Raimond (BBC) presented a research pilot project on the BBC World Service Archive, now out of beta, that automatically tags radio content using linked data and then exposes it to a selected group of users, who are then invited to evaluate and validate the tags' integrity. Since August 2012, the project has attracted more than 2,000 users and one person tagged 200 programmes over a weekend! Over 70,000 programmes with little or no metadata have been digitised and tagged in a short space of time at a relatively low cost. This proves the value of engaging with interested parties to use new technologies to contribute to collections. Liam Wylie (RTÉ) presented the final case study of a grass-roots initiative, which engages the public through Twitter. Wylie argued that only by gaining the attention of the public can collection managers use crowdsourcing as a means to
realise economies of scale – "To use the Crowd you need to have the attention of the Crowd." This is a very simple way of using social media in a truly interactive way where audiences learn what archives do and provide new metadata to enhance the collection; it is a very subtle form of marketing increasing access to the collections with new audiences.

The final panel discussion focused on the value of user-generated tags. While Julia Vytopil (Beeld en Geluid) argued that because the general public often develops non-specific, overlapping, and redundant tags a controlled vocabulary is needed to productively constrain user choice. Aroyo claimed that attention should be paid to niche-sourcing, arguing for tagging initiatives to be geared towards particular user groups, who often require different incentives than general or casual users. The risk of user mis-tagging can be mitigated through the use of an array of different approaches, including filtering (Aroyo), the use of controlled vocabularies (Vytopil), and expert review by metadata specialists (Svein Prestvik from NRK). In response to another question, Aroyo suggested capturing the provenance of user tags in much the same way that Wikipedia documents user edits to its entries. Prestvik conceded that integrating broadcast metadata from many disparate sources has been very challenging, though progress is being made. Both Camilla Roesen (Danish Radio) and Prestvik agreed that production users should generate rights metadata, but acknowledged the difficulties this posed. Elena Brodie Kusa questioned whether relying on social media excluded an important part of the public that had yet to cross the digital divide. Though the panelists agreed this was an important consideration, some believed that people who have been digitally illiterate are rapidly gaining computer literacy skills.

Computer driven innovations seem ubiquitous; from social media tools that allow us to connect online to IBM's Watson, an artificial intelligent computer that won a quiz show. It was not a surprise to hear so many presentations about automatic video and audio analysis, however, the number of computer generated tag projects that were out of beta phase was
impressive. According to one delegate’s predictions (McKernan, 2013), machines will do 90 percent of routine audiovisual tagging, while humans will do the remaining 10 percent. While, Thomas Prehn (Asunto), who co-moderated the event with Elena Brodie Kusa, remarked that humans should be at the center of any tagging initiative, Cees Snoek's presentation supported the trend toward a mostly automated cataloguing workflow. As Aroyo stated in her keynote, humans and machines have a symbiotic relationship. Perhaps computers will grow to understand us, while we learn to embrace the advantages they bring. One thing is certain, as our society's appetite for media continues to grow, so does our need to archive and open up this audiovisual content for the public.

Some concluding Sound-Bites from the delegates on metadata approaches and new technologies:

“A good embrace when technology collaborates with Archives.”
“There will always be Archivists and there is a place for both.”
“We are entering a world of new relationships with IT/Business/Technology Providers and Universities.”

Ian Matzen  
Research assistant, University of Oxford

Jacqui Gupta  
MMC Committee / BBC

London, July 2013
REFERENCES:

<table>
<thead>
<tr>
<th>PAGE</th>
<th>TITLE</th>
<th>DATE</th>
<th>INFORMATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>6669 – DATAMASKINER</td>
<td>1950</td>
<td>Dr. E.U. Condon and Dr. S.N. Alexander at the national bureau of standards, Washington, USA, examining copy from the bureau's high-speed electronic automatic computer. © SVT Bild</td>
</tr>
<tr>
<td>14</td>
<td>169782 – THE STACK</td>
<td>19??</td>
<td>Los Angeles, USA. No further information. © SVT Bild</td>
</tr>
<tr>
<td>54</td>
<td>775686 – KONSERVERAD MAT</td>
<td>19??</td>
<td>Sweden. No further information © SVT Bild</td>
</tr>
<tr>
<td>55</td>
<td>1922967 – LITET KYLSKÅP</td>
<td>19??</td>
<td>German presentation of a miniature icebox for household use. © SVT Bild</td>
</tr>
<tr>
<td>90</td>
<td>998005 – STOPP I LAGENS NAMN</td>
<td>19??</td>
<td>Sweden. No further information © SVT Bild</td>
</tr>
<tr>
<td>91</td>
<td>2048878 – KVINNLIG ADVOKAT</td>
<td>19??</td>
<td>“Miss Dr. Lawyer” in court, Berlin, Germany. Due to economic recession women were for the first time allowed in all professions. © SVT Bild</td>
</tr>
<tr>
<td>169</td>
<td>78721 – SÖKANDET EFTER MAT</td>
<td>194?</td>
<td>Children looking for food in hotel waistbins, Bad-Gastein, Austria. © SVT Bild</td>
</tr>
</tbody>
</table>