MDN WORKSHOP 2016

TUESDAY 7 JUNE

Time	Session	Speaker	Organisation
09:15	*Registration & coffee*		
9:45 – 10.30	**"Sparql as Reporting Tool"**	Robert Engels	NRK
10:30 – 11.15	**NRK Origo, Practical CCDM and EBUCore**	Tormod Vaervagen	NRK
	NRK Origo project is a total reorganisation of the radio- and television infrastructure in NRK. Powered by a common self-developed metadata layer using the CCDM and EBUCore standards, third party products will be integrated to build the total infrastructure. The metadata layer will get inputs every stage of the production face from planning to play-out, and will be the real integration layer in the infrastructure, keeping the loss of metadata to its minimum. In addition semantic technology will be used to link additional data sources. One of the goals with Origo is to make publishing to the web first as easy as to traditionally play out content on TV, making Norway's largest broadcaster becoming a publishing house. The project is now in it first year and the presentation will give an overview over the architecture and metadata structure that powers the project.		
11:15 – 11:45	*Tea, coffee*		
11:45 – 12:30	**Content and Metadata Workflow for UGC in Live Production**	Werner Bailer	Joanneum Research
	Coverage of sports and music events can be improved by inclusion of user generated content (UGC), which provides additional angles otherwise not covered by the production team. Recently, apps for live streaming for UGC have emerged, but the integration of the content into a live production is still challenging. Also, content needs to be filtered automatically in order to only suggest a few streams to the editor. This filtering is based on a broad range of metadata, including for example location and quality of the content. Some of the metadata is available from sensors of mobile devices and needs to be streamed to the production system, while other metadata is extracted from the content. Another important type of information are time stamps of all content and metadata, in order to ensure synchronization in such a challenging scenario.		
	We present a system consisting of a capture app, which streams both content and metadata into a processing system that performs automatic metadata extraction in real-time. Metadata are available as a stream for live consumption and are indexed in a metadata store. Based on the metadata, the available UGC streams are filtered, so that the selected streams can fed into an editing system.		
12:30 – 13:30	*Lunch*		
13:30 – 14:15	**Data Driven Journalism**	Alberto Messina	RAI
	Data – Driven Journalism is the art of making up newsworthy stories by analysing and extracting sense from data. Many efforts have been spent worldwide in this area, leading to several noteworthy success cases. This speech will illustrate some insights about RAI's approach at the matter, with specific emphasis on automated data gathering and analysis tools, reference components and system integration aspects, data models, and the role of semantic data		
14:15 – 14:45	**BRIDGET**	Miroslaw Bober	University of Surrey
14:45 – 15:15	**Tagging Content at the Finnish Broadcasting Company YLE**	Pia Virtanen	YLE
	Yle has developed a series of content description processes and established a "Yle-vocabulary" to semantically tag all Yle content - articles, images, tv and radio programmes - in a common, consistent and language independent way.		
	The presentation will give an overview of this development and how the tags are used for recommendation, personalization, navigation, internal search and automatic linking of content on the web and in applications.		

	Yle uses external vocabularies as primary sources for concepts; those vocabularies, especially Wikidata which was recently implemented, and e.g. challenges with partially overlapping vocabularies are discussed. The tags are produced both manually and automatically. In all content description. Yle is aiming at computer-assisted processes; the progress in automatic metadata production at Yle will also be briefly discussed.		
15:15 – 15:45	*Tea, coffee*		
15:45 – 16:30	**RAI Rights Management System Using MPEG-21's Media Contract Ontology**	Laurent Boch	RAI
16:30 – 17:00	**Opportunities for big data in media and data project at VRT**	Mike Matton	VRT
	Big Data is a buzzword with many definitions. What is a trend however in many industries is the move to more data-driven business processes: the transition to a data-driven economy. This presentation will have 2 parts. First of all, we will provide an overview of which and how media production processes become more data-driven. Furthermore, we will introduce the data project that has recently started, which aims at making several of the media business processes at VRT more data-driven.		

WEDNESDAY 8 JUNE

09:15 – 10:00	**Sport Metadata Workflows – from Ingest to Visualisation** The presentation will go through a representative sport metadata workflow from data capture to data visualisation. Data can be delivered in different formats from different sources and is transformed into semantic data using the EBUSport ontology. Data is then queried for use in various applications and for data visualisation. Different sparql query strategies will be addressed as well as different ways to interact with a triple store. The different formats in which query results can be exported will also be discussed.	Jean-Pierre Evain	EBU
10:00 – 10:45	**A Study about the Organisation of Video Contents Based on the Educational Curriculum** NHK STRL has developed a curriculum ontology which establishes the relation between brush-up / preparation work and items of study. The connections between each item in the ontology are generated based on their new or existing natural words utilizing the order of items of the curriculum as published by the Ministry of Education. This ontology enables us to find the path of study based on specific terms and to make videos aligned with each respective study path. In summary, this work provides a way of constructing/enriching a ontology and structuring video contents.	Makoto Urakawa	NHK STRL (Science & Technology Research Laboratories)
10:45 – 11:15	*Tea, coffee,*		
11:15 – 12:00	**Semantic Fingerprinting**	Maarten Verwaest	Limecraft
12:00 – 12:45	**New Language Technology Tools** BBC News Labs has just launched a second pilot to test new language technology tools in Russian and Japanese. The experiment combines computer-assisted translation (Google Translate) and text to speech voice synthesis to translate news videos. The presentation will report on preliminary results.	Susanne Weber	BBC
12:45 – 13:45	*Lunch*		
13:45 – 14:30	**YLE's metadata in 2020 – How to Answer future Unknown Needs** Online publishing requires more metadata about the content than traditional broadcasting. However, identifying the most valuable metadata is difficult due to rapidly changing business and customers needs. Due to limited resources, only the most valuable metadata elements can be produced. Also implementing metadata changes to existing systems and processes tend to be slow due to complicated architectures and human	Kim Viljanen	YLE

organizations. We propose an agile, just in time approach for identifying and addressing metadata needs in a media company. The individual content producers around the company should be creating the data that is most valuable at the moment - and when the needs changes, the data creation should be changed immediately. In addition to metadata, we also discuss the need for optimizing content for online use and the consequences for metadata creation.

14:30 – 15:00	**The "WITH" Platform – Facilitating Creative Re-use in the Cultural Heritage Domain** The presentation will introduce the "WITH" platform available for cultural institutions, professional users and third party developers to search for cultural resources in particular to promote innovation and demonstrate the social and economic value of cultural content. The platform is used to aggregate content from multiple sources, curate and annotate collections, as well as facilitate interoperability across data models and standards using the services of NTUA's MINT tool. An example will be presented using EUSCreen video archive material.	Vassilis Tzouvaras	NTUA
15:00 – 15:30	*Tea, coffee*		
15:30 – 16:00	**How to Datify Media Content to Win Friends and Influence People: Your Next Metadata Strategy** The advice from Dale Carnegie from his 1936 seminal book *How to Win Friends and Influence People* has never been more relevant than today as evidence-based decision making has shifted the spotlight to the customer. Metadata models/ schemas (EBU) and media ontologies (CCDM) were primarily designed to standardise and streamline production but when the media product(s) were completed, so was the metadata. In today's omni-channel multi-streaming media ecosystem, market drivers have inspired many businesses to rethink their traditional ways of working and adopt a more customer-centric approach. it's no longer enough to merely push content to the masses and hope for the best. From education, healthcare, and media, businesses are re-designing their backend systems and transforming their workflows to run data-to-insight methods that help target more personalised and extended offerings. Harnessing customer data will require new metadata extensions that go beyond production efficiencies to include behaviours *around* the content or products. In this session, temporal metadata, storyline metadata, social activity, and ratings data will be explored to demonstrate how the closing of the "virtual metadata circle" will inform future investments.	Madi Solomon	Optimity Advisors
16:00 – 16:30	**FIMS – A View from the Trenches** An analysis on a strategy to map a 'naive' domain model ("mine") to the FIMS model. How does a developer perceive the FIMS proposition and what should be done to facilitate adoption?	Sean O'Halpin	EBU/BBC
16:30 – 17:00	**Memoways Comet** Comet connects a video to a stream of video fragments. On one hand, you have an edited movie, telling a story in a closed and linear form. On the other hand, you have a stream of video fragments, giving access to additional information in an open and dynamic way. The head of the Comet is the edited movie, while the trail is the stream of video fragments. The use of the video player is very simple: watch a video, select interesting topics just by clicking on keywords that will dynamically appear under the video and create a personal playlist of additional content, you can then remix and share.	Ulrich Fischer	Memoways
17:00 – 17:30	**Wrap-up and conclusions**	Tormod Vaervagen, MIM-MDN Chair	NRK