Title	Annotated Bibliography of AI and Audio-Visual Archives
Working group code	RA01
Study title	Increasing Access to Photos, Videos and Social Media records through AI-generated Descriptive Metadata
Status	Final
Version	
Writers	Jessica Bushey, SHSU; Sarah Joy Maddeaux, IFRC; Kaelan Doyle, GAA, UBC
Date	February 2023

RA01: Annotated Bibliography of AI and Audio-Visual Archives
February 2023

Title	Source Type	Summary	Key Points	Limitations

AEOLIAN Network. "Artificial Intelligence for Cultural Organizations." Accessed August 11, 2022. https://www.aeolian-network.net/	Website	This is the main website for the AEOLIAN Network. The *Artificial Intelligence for Cultural Organizations Network* is a research project funded by the US National Endowment for the Humanities (NEH) and the Arts and Humanities Research Council (AHRC) in the United Kingdom. The AEOLIAN Network aims to investigate the role that AI can play to make born-digital and digitized cultural records more accessible to users. AEOLIAN hopes to contribute to this field by creating workshops and research outputs, and developing an international network of theorists and practitioners working with born-digital and digitized archives.
		
		Note upcoming special issue of Journal on Computing and Cultural Heritage produced by AEOLIAN: 'Applying Innovative Technologies to Digitised and Born-Digital Archives': https://dl.acm.org/journal/jocch/aeolian
			● AEOLIAN has three specific objectives:
			  ○ To make digitized and born-digital collections more accessible.
			  ○ To analyze these collections using innovative research methods.
			  ○ To identify synergies and collaborative avenues between US and UK cultural organizations.
				● The website contains case studies, surveys, and a blog. However, since the project is in its early stages, these sections do not contain a wealth of information and their utility is limited. As the project progresses these sections should hopefully gain more content.

AI4media. "AI4media Project." June 8, 2022. https://www.ai4media.eu/	Website	This is the main website for the AI4Media Project. AI4Media is a 4-year-long project funded by the European Union's Horizon 2020 research and innovation program. The project aims to become a centre of excellence for AI research and to establish a large network of AI researchers from Europe and abroad.
		
		The project hopes to deliver the next generation of core AI advances and training to serve the media sector while ensuring that the European values of ethical and trustworthy AI are embedded in future AI deployments.
			● According to the website, Europe has a human-centric and trustworthy ethical brand of AI to offer to the world
			  ○ One field where such AI is needed is the media
			● The website contains a number of scientific journals, open datasets, and a list of relevant projects and tools
			  ○ CUHE project
			    ■ "The CUHE project aims to develop and demonstrate a web-based application based on AI recommendations that will allow cultural heritage professionals (e.g. museum curators, archivists) as well as (humanities) researchers to explore existing media and cultural heritage digital collections in a more holistic way and allow them to curate new galleries or create digital stories and exhibitions which can showcase and share the new insights gained."
			● Broadcaster archives are a valuable source of info about the past. However, archived broadcast material is frequently impaired and inadequately documented. AI-based tech is needed for video restoration and metadata extraction.
			● AI can be used to detect bias in media
			● AI can be used for organization and content moderation
			  ○ a) Automated (re)organisation of large media collections of photos and video
			  ○ b) Automated and human-in-the-loop moderation of user-generated media content
				● There are few limitations to this source. However, the project is quite recent so many of the case studies and surveys are currently in progress.

Audio Visual Preservation Solutions, Inc. (AVP). "Homepage." July 18, 2022. https://www.weareavp.com/	Website	This is the main website for *Audio Visual Preservation Solutions, Inc.* AVP is an IT Services and IT Consulting company based in Brooklyn, New York. The company offers a number of services to "maximize the potential of digital assets through the innovation of information ecosystems." The AVP website contains several publications, articles, and webinars related to AV archive research that are useful to our study.	- The stated purpose of AVP is to maximize the potential of digital assets through the innovation of information ecosystems - AVP offers a number of digital asset management solutions and services: **avp plan** - AVP will work with stakeholders and user to uncover the true needs, motivations and opportunities of your organization **avp select** - AVP will help clients find the right technology for their data, their users, and their organization. This includes DAM, MAM, digital preservation, or AI tools **avp activate** - AVP will help clients implement, grow, and scale their digital asset management program **avp explore** - AVP will help clients prototype, pilot, and test innovative data solutions	This source is useful and does not contain limitations
The Association of Moving Image Archivists. "The Association of Moving Image Archivists." Accessed August 24, 2022. https://amianet.org/	Website	This is the main website for the Association of Moving Image Archivists. The AMIA is an international nonprofit association dedicated to the preservation and use of moving image media. The AMIA website contains a number of links to events, journals, publications and other resources related to the use of AI for AV archiving.	- The mission of AMIA is to support public and professional education and foster cooperation and communication among individuals and organizations concerned with the squisition, preservation, description, exhibition, and use of moving image materials - The vision of AMIA is to ensure that moving image archivists worldwide shall have adequate support, protection, education, information, funding, and resources necessary to properly and effectively preserve and make accessible the worlds moving image heritage for current and future generations - AMIA has a number of partner organizations: **Preservation** *Iron Mountain Entertainment Services* **Associate Organizations** *National Film Preservation Board* *Coordinating Council of Audiovisual Archives Association* *UNESCO: Memory of the World* *UNESCO: United Nations Educational, Scientific, and Cultural Organization* **Other Organizations** *Association for Recorded Sound Collections* *The Film Foundation* *International Federation of Film Archives* *International Association of Sound Archives* *International Federation of Television Archive* *Society of American Archivists* *SouthEast Asia-Pacific Audiovisual Archive Association* *Northeast Document and Conservation Center* *Oral History Association* *North American Archival Network, International Council on Archives*	This source is useful and does not contain limitations

Beeld & Geluid. "Onderzoek." Accessed August 11, 2022. https://www.beeldengeluid.nl/onderzoek	Website	This is the main research page of the Beeld & Geluid website. Beeld & Geluid is a leading knowledge institute in the field of media culture and AV archiving. The institute works with scientists and cultural heritage professionals to perform research on new and innovative forms of technologies.
		
		The page contains a number of links to current research projects related to using AI to manage AV materials. Furthermore, the website includes a number of research partners, publications, and projects that may be useful to our study.
			● The website contains a number of partners that could be useful for the literature review/study
			   ○ AVA_NET
			   ○ Netwerk Digitaal Erfgoed
			   ○ Podiumkunst.net
			   ○ ClickNL
			   ○ Media Perspectives
			   ○ Canon Van Nederland
			   ○ OorlogsBronnen
			   ○ Den.nl
				● The website is only available in Dutch which makes it less accessible for English speakers. However, Google Translate can be used to provide an overview of the main content on the site.

Beltran, Louis. *"Azure video indexer: Advanced metadata extraction from video and audio content."* YouTube. (2021, June 4). Retrieved October 16, 2022. https://youtu.be/kEkVtyiGJOA	Webinar	This is a video presentation by Louis Beltran, Microsoft MVP and PhD Student at Tomas Bata University in Zlín, Czechia. In the video, Beltran provides an introduction to Microsoft Azure and discusses some of its main features. In particular, Beltran examines the application's ability to automatically generate and extract metadata from AV materials. This is done through a number of APIs such as: Computer Vision API (automatic content analysis of images and videos), Face API (automatic detection of people and emotions), Form Recognizer (automatic text recognition) and more.
		
		Final Thoughts
		
		This source and others appear to confirm that Microsoft Azure is currently the best AI solution for automatically generating metadata from AV materials. This is due to the wealth of features it possesses and the price of the service. Additionally, the platform is quite user friendly and can be utilized without extensive knowledge of programming.
			● Azure includes a suite of services that allow the general public to access and utilize AI technology
			● Azure Cognitive Services
			   ○ Bringing AI within reach of every developer, without requiring machine learning expertise
			● **Azure can use any programming language**
			● You send a file and you get a response back from the system
			   ○ There is computer vision, voice, language, and web search
			● **Computer Vision API**
			   ○ Analyze content in images and videos
			● **Face API**
			   ○ Detect and identify people and emotions in images
			● **Custom Vision**
			   ○ Customize image recognition and tailor it to your business needs
			● **Form Recognizer**
			   ○ Extract text, key-value pairs, and tables in documents
			● Video AI, in particular Video Indexer by Azure is the key to solving the challenges facing modern AV archives
			   ○ Improve content discoverability
			   ○ Increase content value
			   ○ Personalize the viewing experience
			   ○ Uncover hidden content insights
			   ○ Reduce manual labour
			   ○ Increase revenue / drive viewing
			● Video Indexer
			   ○ Extracts metadata and insights from videos
			   ○ It uses machine learning models that can be customized and trained
			   ○ The video insights include face identification, text recognition, object labels, scene segments
			   ○ Additional insights are extracted from audio, such as transcription and emotion detection
			   ○ You can use these results to improve search, extract clips, create thumbnails, and more, thus enhancing user engagement
			   ○ It is available as a service and as a platform
			● To use video indexer you need to create a subscription
			   ○ You can sign up for a free trial
			      ■ Up to 600 minutes of free indeixing using the video indexer portal
			      ■ Up to 2400 free minutes using the API
			● Video Indexer Portal
			   ○ Allows users to upload videos from url, file, etc.
				● This source is informative and does not contain any limitations

| | | | ■ Most common media formats are supported, such as MOV, WMV, MPG, AVI, MP4
● Video Insights
  ○ Can automatically detect languages and allows the user to make adjustments to the automatically generated metadata
● You can find specific topics as well
● You get full captions and can even translate them
● Advancesd audio and video analysis
● Azure can create a keyword breakdown of the video and frames
  ○ Every single frame has a description
  ○ You can combine this with custom vision
    ■ Find specific things that you are wanting
● The video indexer is free to start
  ○ 10 - 40 hours to start
  ○ If you want more advanced capabilities, such as streaming, encoding, storage, you can buy Azure Media Services | |

Cultural AI. "A Lab for Culturally Valued AI." Accessed August 11, 2022. https://www.cultural-ai.nl/	Website	This is the homepage for the Cultural AI initiative. Cultural AI is the "study, design and development of socio-technological AI systems that are implicitly or explicitly aware of the subtle and subjective complexity of human culture." Cultural AI is driven by a number of partner institutions. The website contains a wealth of useful information, including the mission of the initiative, current research projects, preliminary results, and news. Additionally, it contains a number of useful webinars and interviews about the use of AI in the cultural heritage sector.	**Interview with Saskia Scheltjens (Rijksmuseum)** ● The extra data extracted from AV records can be used to create more stories/support different narratives ● Co-operation between cultural heritage institutions and AI researchers needs to be mission-drive, not tech-driven ● 2 main areas for use of AI in cultural heritage: Collection and Organization ● Technology is not neutral, neither is data **Culturally Aware AI - Sub-project 1** ● The way cultural heritage datasets are labelled encodes bias ● How can AI interpret and represent the contents of a heritage collection in a way that captures polyvocality? ● It may help to build a dictionary of contentious and biased terms **Culturally Aware AI - Sub-project 2** ● Transparent Data Stories: Putting heritage data in perspective ● How can AI represent multiple perspectives on cultural objects? 　○ By using a narrative, as it is a human way of learning ● The steps to creating a story is to first create a dataset, then knowledge graph, then a story ● Identify links between terms to identify stories **SABIO - The SociAl BIas Observatory** ● The goal is to detect bias in large AV collections, and make it accessible, quantifiable, and searchable automatically ● 1. Help museums navigate collections ● 2. Facilitate sociological research ● 3. Educate about bias and societal change ● 4. Work towards more inclusive museums **Re-Frame - Analysis of the context and framing in online journalism – Clariah Media Suite** ● Using AI to identify similar themes or topics in different news segments and AV records ● Computer vision 　○ Allows people to conduct frame analysis on data **Questions & Answers** ● Culture for AI vs AI for Culture 　○ Culture for AI means making AI smarter by integrating cultural values into the system 　○ AI for culture means using AI systems to improve culture ● We have to be careful about the information we put into these systems 　○ This can be tricky because many curators do not know what is in their datasets	● While the source is quite useful, there is a lack of discussion about archival theory.
European Commission. "European AI Alliance." Futurium. Accessed August 24, 2022. https://futurium.ec.europa.eu/en/european-ai-alliance	Website	This is the main website for the European AI Alliance. The European AI Alliance is both a gate of resources and a forum created by the European Commission to establish an open policy dialogue on Artificial Intelligence. This includes the legal, technical, and economic implications that AI presents to society. The website contains a number of articles, publications, forums, blogs, webinars, interviews and various events related to the use of AI.	● The European AI Alliance brings together legislators and citizens, academics and practitioners, public authorities, civil society, business and consumer organisations in an open and multidisciplinary community of exchange ● Members of the AI Alliance are able to publish their own resources in the Open Library ● The Alliance has developed a checklist of key requirements for Trustworthy AI. This checklist has been translated into a tool called ALTAI: https://futurium.ec.europa.eu/en/european-ai-alliance/pages/welcome-altai-portal	● The open library and forum limits contributions and posts to citizens of EU Member States.

Europeana Pro. "Empowering Digital Change for the Cultural Heritage Sector." August 24, 2022. https://pro.europeana.eu/	Website	This is the main website for the Europeana Foundation. Europeana is a collaborative initiative based in the European Union that aims to "empower the cultural heritage sector in its digital transformation." Additionally, Europeana hopes to facilitate a Europe "powered by culture" to enhance the economy and resilience of the union. Europeana achieves this goal by functioning as an open web portal for cultural heritage collections across Europe.
		
		The Europeana website contains links to current services and tools, projects, events, webinars, and presentations.
			• According to the website, the mission of Europeana is to empower the cultural heritage sector in its digital transformation
			  ○ The project hopes to achieve this goal by developing expertise, tools and policies to "embrace digital change and encourage partnerships that foster innovation"
			• To achieve this goal, Europeana uses a number of techniques.
			  ○ First, Europeana hosts forums, events, and webinars to help cultural heritage professionals collaborate.
			  ○ Second, Europeana advocates for cultural heritage institutions and cultural heritage professionals.
			  ○ Third, Europeana helps cultural heritage professionals reach new audiences online and shares digitized cultural heritage objects on its website.
			  ○ Fourth, Europeana helps build capacity by providing opportunities for institutions and individuals in the cultural heritage sector to develop their skills and practice
			  ○ Fifth, Europeana helps develop new and innovative forms of technology. The initiative designs their own software solutions, including Europeana, Europeana Pro, and a suite of APIs. The initiative also partners with organizations to produce technological innovations that help cultural heritage professionals work with, improve, and interact with digital culture
			• Europeana has a number of core values:
			  ○ **Usable**: Technologies and tools to support cultural heritage must be easy to use and open to professionals
			  ○ **Mutual**: Innovation must be community based and reciprocal and combine the best of sector knowledge and practices
			  ○ **Reliable:** Europeana ensures that its digital data is always authentic, trustworthy, and robust. It also ensures that digital data is easy to create with and that partners benefit from sharing it
				• This source is useful and does not contain limitations

| EyCon. "Visual AI and Early Conflict Photography." Accessed August 11, 2022. https://eycon.hypotheses.org/ | Website | EyCon is a research project that focuses on using AI-based tools to analyze historical photographs documenting armed violence. EyCon's primary objective is to "assess the usefulness of computation to visualize, navigate and analyze large visual corpora." EyCon aims at harnessing and questioning computation as well as testing new approaches to visualization when it is applied to historical investigation."

According to the team, traditional digitization methods of AV materials have resulted in the creation of archival silos, which reinforce notions of national exceptionalism, international interventionism and expansionism. By using advanced AI technologies, EyCon aims at "connecting, analyzing and commenting on these divided repositories to increase the discoverability and usability of overlooked and scattered material on colonial, imperial and international armed conflicts up to 1918."

The project is supported by a network of archival institutions in France and the United Kingdom and is co-funded by AHRC/Labex Passés dans le Présent and the Université de Paris. | ●   EyeCon has two specific objectives:
  ○   The aggregation of data into a thematic collection on early conflict photography. Image extraction scripts will be used to enrich the database with photo-engravings from digitized published material.
  ○   The development of AI techniques for historical enquiry and data enrichment of a large visual corpus of historical photographs. EyCon aims at providing researchers and other users with integrated computational tools to apply distant vision to its visual database and solutions to augment their close reading capacities.
●   These AI tools will be used for the following purposes:
  ○   To compare and search images for similarities
  ○   To visualize a large visual corpus thanks to image embeddings, topic modelling and clustering
  ○   To isolate photographic tropes, subgenres as well as significant "anomalies" in the database
  ○   To retrain existing datasets on early conflict photographs. | ●   This source is useful and does not contain limitations. |

| F. Tsalakanidou et al., Deliverable 2.3 - AI technologies and applications in media: State of play, foresight, and research directions, AI4Media Project (Grant Agreement No 951911), 4 March 2022 | Report | This report provides an update on the current state of affairs for the AI4Media project. The report is organized according to a series of roadmaps focused on various research topics. Additionally, the report explores future research directions for the project.

Final Thoughts:
This is quite a large report and it is difficult to summarize all of the information contained within it. However, there are a couple of key takeaways.

First, AI technologies can be used to create a wealth of useful tools for the cultural heritage sector. However, these tools must be developed in consultation with cultural heritage professionals who possess knowledge of the humanities in order to ensure that biases are reduced and user friendliness remains a primary consideration.

Second, high-level metadata is required for AI cultural heritage tools to function. The collection of metadata helps optimize content management and facilitate search in large AV archives. | **Roadmap on AI tech and applications for the media industry**
● Metadata creation is essential to optimize content management and search in vast audiovisual broadcast archives
● AI algorithms for video and audio analysis and detection are essential for creating high-level metadata. This includes face detection and recognition, detection of text in video, object detection and recognition, and speech recognition.
● NLP Technology can be used for automatic real-time translation of content and elevated communication experience, free of language barriers.
● ML, DL and computer vision can be used to automatically analyze video content and thus facilitate tasks such as real-time monitoring and trend detection in social media aiming at increased personalization but also enabling monetization of video archive collections.
● AI can be used to transform video and audio in structured data and easily mine volumes of user-generated content for useful information. It can also be used for indexing, archiving and search.
● AI can also automate other content management tasks like data format conversion or sub-title extraction, thus enabling real-time indexing.
● These AI techniques make content discovery easier and more accurate while also enabling automatic content moderation. In addition, they facilitate content personalization and content engagement.
Roadmap "Cross-modal and multimodal representation, indexing and retrieval"
● AI, ML, and DL can be used to develop significant tools for the cultural heritage sector, such as:
  ○ **Search engines with sophisticated cross-modal capabilities:** These search engines can allow cultural heritage professionals to search for images given a textual description of the queried content, provide an image to retrieve videos with similar visual content, or provide audio data to retrieve a corresponding video.
  ○ **AI tools for multimodal analysis of the multimedia content to jointly exploit different modalities:** For instance, a sophisticated joint audio-visual speech recognition solution would allow to efficiently subtitle any conference or discourse and give access to its content by keyword search.
● Searching in multimedia content associated with metadata requires more sophisticated techniques combining reasoning and deep learning.
Roadmap "AI for Social Sciences and Humanities Research"
● In the modern era, media collections are massive and individual items are rich with information. This makes performing even basic analyses of media collections incredibly resource intensive.
  ○ AI can be useful in this scenario
● However, many AI models are generic and not useful to the cultural heritage sector
  ○ Additionally, these models are often solely technology-driven and are created by developers without any knowledge of the humanities. Considerations for biases and user-friendliness often come as an afterthought rather than a primary consideration
Roadmap "NLP and Conversational Agents"
● The transformer model architecture is the most versatile for algorithms in the cultural heritage sector
  ○ This new paradigm has allowed performance increases in various tasks such as named entity recognition and linking, opinion mining and argument mining | ● This source is incredibly rich with information and does not contain any limitations. |

| International Federation of Television Archives. "FIAT/IFTA." Accessed August 11, 2022. https://fiatifta.org/ | Website | This is the main website for the International Federation of Television Archives (FIAT/IFTA) network. Founded in 1977, FIAT/IFTA is a global network and professional association for those engaged in preserving and exploiting broadcast archives.

The website is quite extensive and contains a wealth of information, including the mission statement of the organization, links to webinars, seminars, world conferences, publications, and presentations. | The mission of FIAT/IFTA The main objective of FIAT/IFTA is to "provide a forum for exchange of knowledge and experience between its members, to promote the study of any topic relevant to the development and valorisation of audiovisual archives and to establish international standards on key issues regarding all aspects of audiovisual media management. Four thematic commissions permit to address all the issues related to audiovisual content and to promote the study of specific topics." In order to achieve their mission, FIAT/IFTA organizes an annual World Conference, international seminars, and meetings to discover the future of the AV field and its uses. FIAT/IFTA contains a complex organizational structure The General Assembly is the highest governing body of the federation. The General Assembly meets once every 2 years and appoints the president and elects the Executive Council Members The Executive Council consists of 12 members and appoints positions such as the General Secretary, the Treasurer, and the Vice President The Media Management Commission specializes in current and future trends in metadata and media management and in their causes and consequences when it comes to the archive as an organization, the technologies used, and the skills of the archivist The Preservation & Migration Commission deals with the long term preservation and migration of AV materials The Value, Use and Copyrights Commission provides a forum for members and experts in the AV field to support the valorization, exploitation, and opening of AV collections focused on the end user The Media Studies Commission provides a platform for the international exchange of information and material in relation to archive access The FIAT/IFTA Office manages the operational functions of the organization | This source is useful and does not contain limitations. |

| International Federation of Television Archives. "Implementation of Automation Tools Based on Machine Learning and Artificial Intelligence." YouTube. FIAT/IFTA, July 4, 2022. https://www.youtube.com/watch?v=-CrWSu9OcW0 | Webinar | This is a presentation by Olivio Segura in 2022 for FIAT/IFTA titled "Automate Data Analysis." The presentation discusses the Institut National De L'audiovisuel's (INA) work to redesign the structure and functioning of their audiovisual archive using artificial intelligence technology. Additionally, it provides a detailed overview of an Automated TV News Analysis Workflow.

Final Thoughts:
Segura concludes his presentation with some advice for cultural heritage professionals hoping to use AI to manage AV collections:
● Centralize systems and databases
● Link data models and keep data consistency
● Combine tools to make workflows
● Build user interfaces
● Build cross-departmental teams
● Involve final users
● Know your data- think with datasets and corpus
● The use-case of the AI system should be considered first | **Olivio Segura - Automate Data Analysis Presentation**
● Since the foundation of INA in 1974, the organization has had to deal with new forms of AV materials. They have also moved from analogue to digital over time
● They redesigned their data governance strategy from user-centric to data-centric
● When creating an AI solution, there was a major gap between the engineers and the librarians for what the use case of the system would be
  ○ They needed a person in between the two to link them, someone that had knowledge of both fields
● The team created a flexible toolbox which allows users to mix tools and create workflows
  ○ Using this strategy they are able to create a workflow in a very short amount of time
● The team used the following steps:
  1. Collect
  2. Process
  3. Compare results and data references
● If there is a match between results and data references then we can detect content
Automatic TV News Analysis Workflow: an Overview
● **First Step: Topic Segmentation**
  ○ Topic segmentation is difficult: digital audiovisual content is not strictly delimited by physical boundaries. We have to segment the content before we can proceed with anything else
    ■ **Pictures classification**
      ● Allows the user to have a first cutting of the content
      ● Challenge: TV contents often change graphic styles and identities
    ■ **News segmentation**
      ● Use previous results to apply speech to text and match the different parts of the contents
      ● This includes image classification (seperate pictures of TV sets and pictures of reports), speech to text (extraction of oral speeches in both TV sets and reports), data comparison (build news segmentation by matching vocabulary)
● **Second step: Indexing the Content**
  ○ Identify concepts within the images with object recognition
  ○ Facial recognition and OCR
    ■ Connect OCR with subtitles and on-screen text to try and identify the people in the reports
    ■ Obtain complementary information with OCR (places, producers etc.)
  ○ Music detection with INA speech segmenter
    ■ Distinguish the voices of people and music
    ■ Identify gender
    ■ Use diarisation and transcription to know who is talking and about what
    ■ Use text classification to extract topics
● AI stewardship
  ○ Maintain models and workflows
  ○ Train and sustain references and databases
● Programmes knowledges
  ○ Quality control of data produced | ● Although informative, the webinar does not make reference to the archival field / archival theory. For example, the authenticity of records manipulated using AI technology is not considered. This provides a considerable limitation. |

			○ Documentary watch to maintain models	

Jaillant, Lise. *Archives, Access and Artificial Intelligence: Working with Born-Digital and Digitized Archival Collections*. Bielefeld University Press, an imprint of Transcript Verlag, 2022	Book	This is a book by Dr Lise Jaillant, Senior Lecturer in English and Digital Humanities at the University of Loughborough. The book contains a number of chapters related to the management and preservation of digital records. However, Chapters 1 and 7 specifically discuss the use of artificial intelligence to increase the discoverability and accessibility of AV materials.
		
		Final Thoughts:
		This book is quite informative and provides a couple of useful case studies. One of the key takeaways from the text is the current limitations of AI technology, particularly when applied to the humanities. This is largely due to the subjective nature of the humanities, the multiple forms of media used, and the hierarchical nature of archival collections.
		
		Algorithms used:
		● Stochastic gradient descent (SGD).
		● Hierarchical classification framework using the HCE loss and syntax-aware classifier
			Chapter 1 : Artificial Intelligence and Discovering the Digitized Photoarchives
			● Staff at the Frick Art Reference Library (FARL), in collaboration with Machine Learning experts from Cornell University, Stanford University, and the University of Toronto, are exploring the use of computer vision to enhance the accessibility and discoverability of the FARL collection
			● In particular, the team is attempting to use AI tools to automatically annotate images in the collection with the headings used in the Photo Archives local, iconography-based classification system
			  ○ Additionally, FARL is attempting to increase the efficiency and standardization of its system
			● At first, FARL attempted to crowdsource metadata but the quality was not satisfactory
			  ○ The metadata generated by volunteers increased the discoverability of images in the digital realm but did not provide standardized search results
			● In the end, the team utilized a deep learning neural network
			  ○ However, the FARL dataset presented a couple of interesting challenges
			    ■ This was largely due to the fact that the dataset differed from conventional deep learning classification tasks
			      ● First, each classification heading consists of a series of constituent descriptors
			      ● Second, the Photo Archive classification headings follow a hierarchical structure
			● In order to resolve these problems, the photo archivists developed a decision tree capturing the classification system's syntax
			  ○ The engineers created a novel hierarchical cross-entropy loss and a syntax-aware classifier which resulted in a high performing network
			● Additionally, the deep-learning researchers incorporated this syntax into both the objective with which the network parameters were optimized as well as the algorithm with which the network's decision was made
			● After training, the network was fed images from the unlabeled portion of the Library's Photo Archive and the network predicted a classification heading for each one
			  ○ These images were then annotated with the predictions and shown to the Library's photo archivists using the Zooniverse application
			● At the end of the chapter, the team provides a comparison between the performance of SL, SML, and HML
			  ○ Although HML is more efficient, it can increase the rate of false positives
			Chapter 7: Inviting AI into the Archives: The Reception of Handwritten Recognition Technology into Historical Manuscript Transcription
			● Until recently, transcriptions of manuscripts were costly to generate
			● Handwritten Text Recognition (HTR) technology is now transforming access to our written past, producing increasingly accurate transcriptions for use by individuals and institutions, or providing material for further analysis
			● Although mass digitization of records has taken place over the last thirty years, these records are not easily readable or searchable
			● There are a number of Handwritten Text Recognition (HTR) techniques currently available
			● However, commercial systems can be opaque / black box
			  ○ The algorithms and approaches used are rarely published, and care must
				● This source is quite useful, and provides a concise overview of some of the current AI technologies being utilized in the cultural heritage sector. However, it does not make explicit reference to metadata.

			be taken regarding copyright, image licensing, and long-term storage of digital assets
			● One example of an HTR service is Transkribus
			● The authors conducted a survey on the satisfaction rate of Transkribus
			  ○ 80% satisfaction rate
			  ○ However, only 4% of respondents said the results generated from Transkribus were very accurate and required little correction, 34% said results were quite accurate; 16% said results were disappointing, and 8% found results unusable
			● Transkribus Features
			  ○ Automatic line detection, layout analysis, and segmentation; HTR training; Keyword Spotting; tagging functionality; export functions (including TEI); the ability to manually correct and edit results; the ability to share ongoing work with others including working on group transcription; table region generation; Unicode integration; and the ability to create standardized transcripts.
			● Some of the participants claimed that machines should never be trusted, and that these tasks should exclusively be done by paleographers and historians
			● The participants had concerns about the proprietary nature of the service and whether the information would be stored according to legal and industry best practices
			● However, the participants also commented on the benefits of the technology
			  ○ For example, it is traditionally very difficult to know if a document will have any useful information before reading it. By creating an searchable OCR copy of the record, HTR can increase the chances of finding useful information in historical documents
			● Final Recommendations
			  ○ Recommendations for memory institutions include: digitization processes should produce (and make available) high quality images to facilitate HTR, and also digitize as widely and completely as possible.
			  ○ Recommendations for Transkribus (and other HTR providers) include: further information on how HTR operates should be provided, to allow researchers to understand "epistemic affordances"; users should be pointed to resources and training on how best to utilize the results of HTR; and individual users and projects should be encouraged to share their models, results (and where possible, data), to benefit the wider historical community	

Klein, Lauren. *Why AI Needs the Humanities*. Loughborough Institute Of Advanced Studies, 2022. https://www.youtube.com/watch?v=iR4uJqoTgys	Webinar	This is an online presentation delivered by Lauren Klein in 2022 for the Loughborough Institute Of Advanced Studies. In the presentation, Klein discusses the value of the humanities for informing the development and use of artificial intelligence.
		
		Final Thoughts:
		The humanities is essential for developing ethical and unbiased AI systems. When considered solely through an analytical / technical perspective, AI solutions often fail to perform in cultural heritage institutions.
			● Language model sizes have been increasing massively each year
			● Data by Design Project
			  ○ The act of visualizing data is never theory or context free
			  ○ It needs the humanities
			● Art historians are wary of AI taking over all cultural heritage work
			  ○ Klein claims that art historians should not be concerned about this, as it is not an either or situation
			  ○ Art historians also worry about AI prioritizing formalism over the social context of the art
			    ■ This is not the case, pattern recognition is a big part of art history
			● AI is not just packaging information and reusing it
			  ○ The questions and answers are not prepackaged
			  ○ There is no one size fits all AI
				● Although the webinar is incredibly informative, it does not provide much information about the use of AI in the archival field.

LLC, Brandeis. "CLAMS Computational Linguistics Applications for Multimedia Services." CLAMS. Accessed November 2, 2022. https://clams.ai/.	Website	This is the main website for the Computational Linguistics Applications for Multimedia Services (CLAMS) project. The goal of the project is to "develop an open platform with state-of-the-art AI tools for archivists and media researchers to improve access, search, and exploration of archival audiovisual material." In particular, CLAMS tools aim to automatically generate metadata and information from audiovisual materials. The CLAMS project is funded by the Andrew W. Mellon Foundation.	- CLAMS contians three primary components - Multi-Media Interchange Format (MMIF) - The goal of MMIF is to "have an open serialization format for computational analysis tools that support interoperability between such tools and software, so that users of the tools can create and customize different pipelines to extract meaningful information and insights from digitized audiovisual material" - MMIF has two different meanings 1. An annotation format for audiovisual media as well as associated text (transcripts, captions, on-screen text) 2. A collection of linked-data components that specifies such a format syntactically and assigns semantics to elements in the format - MMIF is based on Python - Workflow Engine - Easy to adapt - In active development - Includes a toolshed of CLAMS compatible apps - Software Development Kit - Includes a collection of APIs to develop CLAMS apps - Built on python - In active development, pre-alpha stage	The CLAMS project is in an early stage of development and many of the tools are currently pre-alpha. As a result, there is a lack of documentation related to the project.
Loughborough Institute Of Advanced Studies. AI & Cultural Heritage - Roundtable Discussion. 2022. https://www.youtube.com/watch?v=ROXVJ-TUhTA.	Webinar	This webinar is an AI and cultural heritage roundtable discussion hosted by the Loughborough Institute Of Advanced Studies. The webinar provides a general discussion about the use of AI in the cultural heritage sector, including its strengths and weaknesses. **Final Thoughts:** It is important to consider the broader societal impact of AI systems, as they can reinforce biased/colonial/racist narratives. Additionally, in the context of archival institutions, we need to consider how these AI systems will impact the archival nature of the records they manipulate. These points should be considered before implementing AI systems.	- One of the current uses of computer vision technology in the cultural heritage sector is to identify archeological ruins and types of paintings - Can we use those tools to disseminate information to the general public? - Will AI help include the general public in the heritage field more? - Potentially, but it is complicated - It is important to question the metadata found in AV collections because they can uphold colonial / racist / biased narratives - Archivists need to be concerned about what AI is doing to the records - What does AI do to authenticity? At the end of the algorithm do we still have the original record?	The webinar is quite short, and due to the discussion format does not contain sources/citations. However, it provides a good starting point for further research directions.

| MeMAD. "Methods for Managing Audiovisual Data." March 10, 2021. https://memad.eu/ | Website | MeMAD is an EU-funded H2020 research project (2018-2020). MeMAD aims to develop methods for the efficient re-use and re-purpose of multilingual audiovisual content targeting to revolutionize video management and digital storytelling in broadcasting and media production.

MeMAD produces detailed, rich descriptions of moving images, speech and audio. MeMAD integrates the latest research achievements in Deep Neural Network techniques in computer vision with knowledge bases, human and machine translation in a continuously improving machine learning framework.

The MeMAD website is quite substantial and provides links to webinars, publications, and study results. | MeMAD claims that they make the AI system "learn from the human" to create descriptions that are not only a time aligned semantic extraction of objects but also utilizes audio and recognizes action sequences MeMAD utilizes deep and recurrent neural networks, machine learning, artificial intelligence, and big data analysis **The MeMAD project contains a number of primary objectives:** 1. The first objective of MeMAD is to "Develop Novel Methods and Tools for Digital Storytelling" Current methods for fiction production and factual storytelling rely on human interpretation. Adding captions and descriptions to this content is slow and expensive. 2. Deliver Methods and Tools to Expand the Size of Media Audiences 3. Develop an Improved Scientific Understanding of Multimodal and Multilingual Media Content Analysis, Linking Media and Consumption 4. Deliver Object Models and Formal Languages, Distribution Protocols and Display Tools for Enriched Audiovisual Data **MeMAD also contains a number of potential use cases:** 1. Content Delivery Services For the Re-Use by End-Users/Clients Through Media Indexing and Video Description 2. Creation, Use, Reuse And Repurposing Of New Footage And Archived Content In Digital Media Production Through Media Indexing And Video Description 3. Improving User Experience With Media Enrichment By Linking To External Resources 4. Automated Subtitling/Captioning And Audio Description. Speech And Sounds To Text And Also Visual Content To Text, Both With Multiple Output Languages, For General Purpose Use And For The Deaf, Hard-Of-Hearing, Blind, And Partially-Sighted Audiences | The project is still in its early stages and further research needs to be done. |

| MeMAD. "Human-Driven AI Solutions to Access and Manage Audiovisual Content: A Research Perspective [Memad]." YouTube, March 9, 2021. https://www.youtube.com/watch?v=iI-ryBy3UGQ | Webinar | This webinar includes two presentations for the MeMAD project. The first presentation is by Jorma Laaksonen from Aalto University, Finland, and is titled *Recent Advances in Automatic Description of AV Materials*. This presentation provides an overview of the current AI research being conducted by the MeMAD project. This includes visual video captioning, face recognition, speaker diarization, speaker naming, spoken language identification, audio background recognition, scene recognition, and automatic multimodal content description.

The second presentation, titled *The Role of Human Input in Automating Audiovisual Content Description* by Sabine Braun, Kim Starr and Jaleh Delfni*, examines how human AV descriptions function and then compares them to machine learning generated descriptions.

Final Thoughts:
Although AI systems and tools can be incredibly useful for managing AV collections, they are not without their issues. First, all data inputs in an AI system are prone to errors. Data must be carefully curated before being entered into the system to avoid perpetuating biases. Second, it is incredibly hard to get all parts of a record's description correct simultaneously using AI. Third, formalizing the concept of "interestingness" is very hard when it comes to using AI to manage archival materials. For example, who gets to define what is interesting or valuable? | ***Recent Advances in Automatic Description of AV Materials - Jorma Laaksonen***
● There have been a number of recent advances in automatic description of AV materials
  ○ **Visual video captioning**
    ■ Aalto University's inhouse development
    ■ LSTM model with self-critical reinforcement learning
    ■ Both static image and visual and aural video features
    ■ Training with MS COCO, TGIF and VATEX datasets
    ■ Evaluation in TRECVID 2020 Video to Text Task
  ○ **Face recognition**
    ■ EUROCOM'S inhouse development
    ■ Automated finding of reference images for known persons
    ■ MTCNN face detection followed by faceNet face recognition
    ■ Evaluated with MeMAD project's TV programs
  ○ **Speaker Diarization**
    ■ Lingsoft/LLS' inhouse development
    ■ Separates speakers from each other and background music
    ■ Based on the sparse auto-tuning spectral clustering by using normalized maximum eigencap
  ○ **Speaker naming**
    ■ Ad hoc method used to associate names from faces to voices
    ■ Works very reliably for some program types
  ○ **Spoken language identification (SLID)**
  ○ **Audio background recognition**
    ■ VGGish NN audio features and NN model with multi level attention
  ○ **Scene recognition**
  ○ **Automatic multimodal content description**
  ○ When you combine the inputs from all of these technologies you are able to create sentences that describe the video source
● The webinar also discusses further considerations and future work including the following:
  ○ All inputs are prone to errors
  ○ It is hard to get all parts of the description correct simultaneously
  ○ Some part of the description is likely to be correct in any moment
  ○ The description should address shots longer than 3-10 seconds ideally
  ○ The topic of the discussions should be named from ASR output
The Role of Human Input in Automating Audiovisual Content Description - Sabine Braun
● The multimodal content revolution has many different end-user requirements
● Main issue with Machine description:
  ○ **Methodological issues:** Crowd sourced training data: purpose, quality, mostly still images / single frames; availability and size of data sets
  ○ **Computer vision issues:** Character recognition, gender ID, Object ID: scale, perspective, non iconic imagery, action labeling
  ○ **Linguistic and sequencing issues:** Lexical choice, accuracy; Storytelling: cohesion
  ○ **Integration of other modes:** audio cues; vocal profiling speech comprehension
● Modeling of human AV content description:
  ○ Discourse processing as bottom-up / top-down processing
    ■ Mental modeling, knowledge activation, framing (conceptual modeling)
● Capturing the finer detail: | ● This source is useful and does not contain limitations. |

- ○ From saying/seeing to meaning
- ○ Relevance and saliency, explicature and implicature
- ○ Semantic / pragmatic modelling
- Capturing the whole story:
  - ○ Schemata, story grammar and coherence
  - ○ Narrative sequencing/schematic modelling)
- Next Steps for machine descriptions: recommendations
  - ○ Efficient character identification
  - ○ Intelligent object recognition
  - ○ Informed action labelling
  - ○ Temporal sequencing
  - ○ Narrative saliency
  - ○ Sensitivity to narrative paradigms of storytelling
- MeMAD content providers:
  - ○ MeMAD datasets (human-made metadata)
  - ○ EBU core-based data model
    - ■ Visualizing the KG: MeMAD explorer
    - ■ To Mod API
  - ○ How to segment and categorize media content to be searched or re-used? In other words, what's a unit of content?
  - ○ How can we leverage text to enrich media content?
- Temporal content alignment
  - ○ Chunk the program into smallest units of content
  - ○ Generate vector representations for text units and content descriptions using SentenceBERT
  - ○ Compute a similarity matrix between each description and chunk vector representations
  - ○ Cluster successive units of high similarity into segments
- Unsupervised topic switch detection
  - ○ Use AI to detect when the topic changes
- How to extract video interesting moments based on multimodal info
- How to treat the different modalities?
- Formalizing interestingness is hard
  - ○ Hard to define, interesting to who?
  - ○ Survey on interestingness and related concepts
- MediaEval predicting media memorability
  - ○ Audiovisual feature extraction
- Machine translation in MeMAD
  - ○ Multilingual MT with multimodal input
  - ○ Neural machine translation
  - ○ Discourse oriented MT
  - ○ Cross-lingual content retrieval
- Neural machine translation in a nutshell
- Multilingual neural machine translation
- Neural interlingua model
- Transfer learning and zero-shot translation

MeMAD. "Industrializing Media Production: A Producer's Perspective [Memad]." YouTube, March 9, 2021. https://www.youtube.com/watch?v=nI0AF8zqT50	Webinar	This webinar examines the ways in which MeMAD can be used to enhance media production. In particular, the webinar looks at political debates and newscasts and examines how AI technology can be used to enhance user access to these forms of media. **Final Thoughts:** The presentation concludes with a number of key takeaways. First, although Machine Translation is faster than human translation, it is not pleasant or creative. Second, the benefit from automated data is different depending on the kind of task. Third, automatically generated metadata and record descriptions can be improved by having humans curate them. Fourth, starting with inadequate auto generated metadata can have a detrimental impact on usability and acceptance of descriptions. Finally, records professionals must be aware of an overabundance of data when using AI tools- less is more.	**AI can be used to analyze political debates:** - Facial recognition, Text recognition and speaker ID, Natural language processing, Speech to text, Vocabulary analysis, Named entity recognition, Topics modelling **MeMAD created microservices to pull crucial information from debates:** - Audio segmentation, Audio classification, Automated speech recognition (ASR), Face detection and recognition, Shot-cut detection, Video captioning, Machine translation (MT), Named entity recognition and disambiguation, Semantic enrichment, Subtitle generation, Content description generation, Content segmentation, Relevant tv moment detection, Spoken language segmentation and classification, Object classification - Data crossing is used to empower analysis **AI use cases:** **Business Case** - Case 1: Subtitles and translations - Which method is better? Faster? Most comfortable? - How does the work process change with added automation? - Which areas are subject to improvement? - Automated titling: MT subtitles are faster, but they are not pleasant or creative: - Language pairs differ - Tedious, annoying, limiting - Spotting and segmentation should be improved - Not all inputs are the same - Tangible spotting improvements have already been made during the project **For exploration and discovery** - Case 2: Searching archive material - Search in combination with machine translation tech to find specific clips - Can search by subject or theme too - Tasks are familiar enough to the participants allowing them to comment and compare - The benefit from automated data is different per type of task - Object and face recognition is useful when there is not dialogue, otherwise dialogue is better **Solution: reconciliation of Metadata** - Human curation of automatically generated metadata - Improving content description using AI sources as input instead of starting from a blank page - Summarize and reduce metadata volume - How many layers of data do we want to store? - How do we keep the metadata that is correct in a tool that is pleasant to use - Case 3: Using rich metadata in existing production tools - Use rich metadata in the existing tools - Too much data is useless - Need to provide less metadata that is more accurate **Conclusions** - Some tools are at the stage of measurable productivity improvement (subtitling) - Saves time, but still needs work to improve quality: to make consumer reception and human curation more pleasant	This source is useful and does not contain limitations.

			▪ Productivity depends on the data input: starting from suboptimal auto generated metadata (ASR) has a critically negative impact on usability and acceptance
			○ Other solutions are moving from exploration to implementation
			    ▪ Automatically created metadata can tentatively replace human-curated descriptions (for search), but this depends on multiple dimensions of metadata completing one another
			    ▪ Beware of too much data! Less is more
			    ▪ Challenges with combining large volumes of rich A.I. - generated metadata with existing production tools that people use for everyday work	

MeMAD. "Making Content Available for All: An Accessibility Perspective [Memad]." YouTube, March 9, 2021. https://www.youtube.com/watch?v=laat2U9ZMuA	Webinar	This webinar discusses the ways in which MeMAD can support access to AV materials.
		
		According to the presenters, MeMAD combines various natural language processing technologies into one product to ensure user accessibility to materials. This includes automated subtitling, captioning, and speaker segmentation.
		
		The webinar concludes by presenting the results of a user satisfaction study for MeMAD.
		
		Final Thoughts:
		The use of MT for automated subtitling and translation has mixed results. For some translators, these technologies improved productivity, while for others it decreased productivity. In general, the user experience of these technologies was negative, and the individuals surveyed had mixed opinions on incorporating these technologies into their workflows. This could be due to the individual preferences of the people surveyed, or the limitations of the technology at the present time.
			● MeMAD combines various NLP, ML, and DL technologies into one user-friendly product.
			● MeMAD has a number of use cases on accessibility
			  ○ Automated subtitling, captioning, and audio description of AV content
			  ○ Speaker segmentation
			  ○ Speech recognition
			  ○ Automatic translation
			● There are two viewpoints on accessibility
			  ○ 1. Broadcasters and producers of AV materials have a huge amount of materials coming in every day and the manual subtitling and translation they are doing is becoming extremely expensive and resource intensive
			  ○ 2. Viewers who require support to participate need accurate and quick information (not 2 weeks later)
			● MeMAD results:
			  ○ For some translators, MT improved productivity, while for others it decreased productivity. A fast translator does not benefit as much from MT
			  ○ User experience was negative
			  ○ Mixed opinions about incorporating the service into their daily work
			  ○ There were many concerns about the impact of the technology on the quality of translations
				● Although the webinar is comprehensive, the final results are inconclusive. The study appears to indicate that cultural heritage professionals see limited utility / benefit from using AI technologies. However, the study also claims it can help improve productivity in certain use cases. Future exploration is required.

Soni, Maulik. "Video Encoding and Streaming Using Azure Media Services." YouTube, June 28, 2019. https://youtu.be/ZDBCGP0j_Bw.	Webinar	This webinar is a presentation by Maulik Soni, Cloud Solutions Architect at Microsoft.	- Azure includes a number of features for media files - Media & Entertainment - Content creation, - Rendering, video editing - Content management - Media asset management - Archiving - Machine learning - Correlation - Metadata enrichment - Search - Digital Supply Chain - Content distribution & monetization - Live, linear, and on demand OTT delivery - Publishing - Audience Analytics - Personalization - Ad targeting - Offer management - Public safety - Surveillance & Intelligence - High scale ingest - Machine learning - Metadata enrichment - Search - Forensics - Archiving - IoT Cameras - Block Chain - Smart Cities / Stadiums - MSFT 1st Party - Cloud video and intelligence - Stream - Teams - Skype - Yammer - Xbox - MSN - Azure media services - Multiple processes and steps - Cloud upload & live ingest - Transcoding - Indexing (Media AI) - DRM Encryption - VoD, Live Streaming & Multi-CDN - Player clients - Media file basics - Container format - using codec technology - Containers contain a header information, a video codec and an audio codec - Why is video compression used? - Compression saves a large amount of storage space - Codec is for encoding and decoding a source file - It is software or hardware based

| | | | | ○ There is lossless and lossy
    ■ Lossless keeps the file as accurate as possible
● Soni then provides a live demonstration of Azure media services
● Files are stored as a blog
    ○ Video and audio are combined in a container | |

Stevenson, Adrian. "Employing Machine Learning and Artificial Intelligence in Cultural Institutions." Archives Hub Blog, July 9, 2021. https://blog.archiveshub.jisc.ac.uk/2021/07/09/employing-machine-learning-and-artificial-intelligence-in-cultural-institutions/	Blog Post	This is a blog post by Adrian Stevenson which discusses the author's experience at the Aeolian Project's workshop titled *1: Employing Machine Learning and Artificial Intelligence in Cultural Institutions*. Adrian Stevenson is the Technical Innovations Manager at Jisc and focuses on using AI/ML to support the Digital Resources Directorate of the organization.
		
		The blog post breaks down the various presentations delivered during the workshop and provides the author's comments on each solution.
		
		Final Thoughts:
		There are a few common takeaways from each of the presentations discussed by Stevenson. First, as was seen in other sources, AI requires well-curated data for it to function adequately. Second, tools for annotating data are just as important as classifiers. Third, although ML excels at suggesting possibilities, identifying trends and gaps, visualizing content, and enhancing catalogues with new terminology, it needs framing and context to reduce errors and false positives. Fourth, AI tools need to be integrated into day-to-day workflows and user interfaces if their benefits are going to be realized. Fifth, AI developers need to be aware of the false dichotomy between maintenance and innovation, and that maintenance is sustained innovation. Finally, cultural heritage professionals need to consider heterogeneity, diversity and whether or not AI is just and fair.
			'Visual AI and Printed Chapbook Illustrations at the National Library of Scotland' – Dr Giles Bergel (University of Oxford / National Library of Scotland)
			● Gile's team is using ML on data from data.nls.uk to detect illustrations in manuscripts
			● The team uses a three step approach
			  ○ First, the team finds illustrations in manuscripts using Google's EfficientDet object detection convolutional neural network
			  ○ Second, the group matches and groups illustration using their "state of the art" image search engine
			  ○ Finally, the team applies classification algorithms with the VGG Image Classification Engine which uses Google as a source of labelled images
			● The team learned the following lessons from using ML technology on AV materials:
			  ○ AI requires well-curated data
			  ○ Tools for annotating data are no less important than classifiers
			  ○ Generic image models generalize well to printed books
			  ○ 'Classical' computer vision still works
			  ○ AI software development benefits from end-to-end use-cases including data preparation, refinement, consulting with domain experts, public engagement etc
			'Machine Learning and Cultural Heritage: What Is It Good Enough For?' – John Stack (UK Science Museum)
			● In this presentation John Stack gives an overview of how AI is being used by the UK Science Museum to collect data into a central knowledge graph
			● The team hopes to use AI for name disambiguation as a first step before adding links to wikidata and using entity recognition to enhance their own catalogue
			● According to John, ML provides a number of main benefits:
			  ○ ML excels at suggesting possibilities and identifying trends and gaps
			  ○ ML is also useful for visualization and identifying related content, as well as enhancing catalogues with new terminology
			● However, there are limitations to ML
			  ○ ML content needs framing and context
			  ○ False positives are not always apparent and require specialist knowledge
			● Final thoughts:
			  ○ We need to consider how AI may drive new ways/modalities of browsing that we haven't imagined yet
			  ○ AI tools need to integrate into day-to-day workflows for their benefits to be realised. Integration into UI is crucial
			'AI and the Photoarchive' – John McQuaid (Frick Collection), Dr Vardan Papyan (University of Toronto), and X.Y. Han (Cornell University)
			● In this presentation, John McQuaid provides an overview of how the Frick Collection is using the PyTorch deep neural network to identify labels for their photo archive collection
			● The team then compares the ML results with internally crowdsourced data from their staff and curators captured by the Zooniverse software as a validation exercise
			● In the end, the team determined that 67% of the ML labels matched with their crowdsourced data
			  ○ As a result, ML is most useful for curatorial amplification but still requires significant human effort
				● The blog post is a summary of an online workshop that was delivered in 2021. As a result, the post is not as comprehensive as the original source.

			'Keep True: Three Strategies to Guide AI Engagement' – Thomas Padilla (Center for Research Libraries)
			● In this presentation, Thomas Padilla discusses the opportunity for GLAMs to distinguish themselves in the AI space
			● Padilla also covers a number of themes related to AI
			  ○ 1. The Non-Scalability Imperative
			    ■ Scale is everywhere with AI
			    ■ However, there is a current problematic dependency where scalability is made possible by non-scalable processes, resources and people
			    ■ We need to consider heterogeneity, diversity, and whether or not AI is just and fair
			  ○ 2. Neoliberal Traps
			    ■ Who gets to decide what is considered ethical AI?
			    ■ According to the author, although GLAMs are trying to do the right thing with AI, this is often done in the context of "Neoliberal moral regulation" which is unfair and ineffective
			  ○ 3. Maintenance as Innovation
			    ■ There is a false dichotomy between maintenance and innovation
			    ■ Instead, maintenance is sustained innovation
			Roundtable discussion with the AEOLIAN Project Team
			Dr Lise Jaillant, Dr Annalina Caputo, Glen Worthey (University of Illinois), Prof. Claire Warwick (Durham University), Prof. J. Stephen Downie (University of Illinois), Dr Paul Gooding (Glasgow University), and Ryan Dubnicek (University of Illinois).
			● In this presentation Stephen Downie talks about the need for standardization of ML extracted features so they can be reused across GLAMs in a consistent way
			● We may need to work more on the data rather than wasting time and resources re-creating expensive ML models	

Stevenson, Jane. "Assessing Machine Learning Outputs." Archives Hub Blog, June 29, 2022. https://blog.archiveshub.jisc.ac.uk/tag/labs/	Blog Post	This is a blog post by Jane Stevenson, the Archives Hub manager at Jisc. The author discusses the current use of AI technology for the Archives Hub Labs project. Additionally, the author compares different types of AI solutions and discusses the limitations of AI technology for managing AV materials.
		
		Final Thoughts:
		There are a number of off-the-shelf AI tools currently available for purchase. However, these tools are often not customly designed to be used by the cultural heritage sector. As a result, many of these ML tools are unable to present their results in an easily digestible manner.
			● One of the current challenges with AI projects is the presentation of Machine Learning results
			  ○ Unfortunately, there are not many out of the box tools that can help with this either
			● The author then describes their experience using AWS Rekognition and other AI tools
			  ○ AWS Rekognition offers a number of features
			    ■ **Image Moderation:** This tool identifies and censors offensive or controversial images
			    ■ **Face Comparison:** This tool allows users to identify individuals throughout a collection of AV materials
			    ■ **AWS Text Rekognition:** The team used this tool to extract text from images. The team also compared Text Rekognition to Transkribus Lite and concluded that both work equally well. However, these services often struggle to detect handwriting. Finally, Transkribus struggled to detect text on posters while AWS Rekognition experienced no issues.
			● The blog post concludes with a number of questions:
			  ○ How do we present the results back to the project team?
			  ○ How do we present the results to the participants?
			  ○ Do we ask participants specific questions in order to get structured feedback?
				● The blog examines a select number of AI solutions and does not conduct a comprehensive survey.

			○ Will we get text that is useful enough to go to the next step? ○ Which images provide good text and which don't? ○ How might these results be used on the Archives Hub to help with discovery?	

Stevenson, Jane. "Images and Machine Learning Project." Archives Hub Blog, January 27, 2022. https://blog.archiveshub.jisc.ac.uk/2022/01/27/images-and-machine-learning-project/.	Blog Post	This is a blog post by Jane Stevenson, the Archives Hub manager at Jisc. In this blog post, the author discusses the addition of a new sub-project under the Labs project titled "Images and Machine Learning."
		
		The purpose of this subgroup is to examine the use of AI to improve the recordkeeping practices of the organization, facilitate user collaboration, and make the work of researchers easier.
		
		Final Thoughts:
		There are a number of current issues with using DAO Store and the IIIF. First, preserving digital objects in solutions such as DAO Store is challenging as archives are hierarchical in nature. Additionally, archival metadata has its own standards and requirements. Second, using frameworks such as the IIIF can be challenging, as archives often do not know what kinds or sizes of digital objects they will receive beforehand. Furthermore, different contributors may have different preferences for the interface and the functionality of the IIIF.
			● The Images and Machine Learning Project at Jisc has three main themes
			  ○ DAO Store
			  ○ International Image Interoperability Framework
			  ○ Machine Learning
			● These themes are being worked on with 10 participants, including: Cardiff University, Bangor University, Brighton Design Archives at the University of Brighton, Queen's University Belfast, the University of Hull, the Borthwick Institute for Archives at the University of York, the Geological Society, the Paul Mellon Centre, Lambeth Palace (Church of England) and Lloyds Bank.
			● According to the author, the objective of the project is to acquire a "realistic sense of what is required to implement digital storage and IIIF display, and to see how machine learning tools work with a range of content"
			● DAO Store
			  ○ Jisc is looking into providing a solution to host the content of its contributors. However, there are a number of challenges with hosting digital objects.
			    ■ One of the main issues with preserving digital objects is the nature of archives as hierarchical multilevel collections
			    ■ Additionally, archival metadata has its own particular requirements
			  ○ The team is considering allowing contributors to upload content to their current administrative interface and allowing them to attach metadata
			● International Image Interoperability Framework
			  ○ IIIF is a framework that enables images to be viewed in any IIIF viewer
			  ○ IIIF allows researchers to group images and easily compare them
			  ○ However, there are a number of challenges associated with using this technology:
			    ■ The archive does not know beforehand what kinds of digital objects it will receive. For example, images could be of varying resolutions
			    ■ Contributors may have different preferences for the interface and functionality of the IIIF
			● Machine Learning
			  ○ The main issue with machine learning lies in the data
			    ■ If the data is inadequate, biased, or suspect in some way, then the outcomes are not likely to be good.
				● Due to its medium, the source does not contain any citations or links to further resources. Additionally, it is written strictly from the perspective of Jisc and may have limited use outside of the organization.

Stevenson, Jane. "Machine Learning with Archive Collections." Archives Hub Blog, February 28, 2022. https://blog.archiveshub.jisc.ac.uk/2022/02/28/machine-learning-with-archive-collections/.	Blog Post	This is a blog post by Jane Stevenson, the Archives Hub manager at Jisc. In this blog post, Stevenson provides an introduction to machine learning algorithms and examines their use in the archival field. Additionally, the blog post provides an insight into the current use of AI at Jisc.
		
		Final Thoughts:
		
		There are a couple of main takeaways from this blog post. The first takeaway is the distinction between supervised and unsupervised machine learning and their unique use cases. In the context of our study, supervised ML is ideal as it allows us to carefully select the training data. The second takeaway is the challenges associated with creating datasets. Creating datasets for archival institutions is often challenging as the records are frequently unlabeled and the collections are often too small to create adequate training data. Finally, the blackbox nature of ML models poses a risk to accountability and transparency.
			• The Jisc Archives Hub is primarily concerned with using machine learning to aid archival discoverability and to identify patterns and bias in cataloguing
			• ML can be carried out as either supervised or unsupervised.
			  ○ Supervised learning is the most reliable and produces the best results.
			    ■ Supervised requires a set of data that contains both the inputs and the desired outputs - this is called training data.
			    ■ Creating datasets is challenging because archive collections may not be labelled or be at the size where this process is optimized.
			• The ability for ML to aid in discoverability depends upon how well the image is already catalogued, whether training data can be provided to improve the algorithm, and how well ML can then pick out features that might be of use.
			• It is sometimes more time consuming to train the algorithm than to simply catalogue manually.
			• It can be very hard to understand what is happening in a ML algorithm and why. Lack of transparency and prevalence of black box models.
				• The blog post does not provide much detail about ML and functions more as an introduction to the summary.

| Stevenson, Jane. "The Archives Hub and IIIF: Supporting the True Potential of Images on the Web." Archives Hub Blog, October 26, 2021. https://blog.archiveshub.jisc.ac.uk/2021/10/26/the-archives-hub-and-iiif-supporting-the-true-potential-of-images-on-the-web/. | Blog Post | This is a blog post by Jane Stevenson, the Archives Hub manager at Jisc. In this blog post, Stevenson provides an introduction to the IIIF model and discusses what archives hub is planning to do with the technology.

Final Thoughts:
IIIF can be used to solve a number of current challenges associated with managing AV materials. First, it can help prevent data silos by allowing different repositories to share their content through the IIIF API. Second, it can enable searching within the text of an object, which facilitates discoverability and access. Third, it helps researchers conduct their work by allowing image comparisons, external connections, sequencing, deep zoom, and more. However, in order for these tools to function adequately, good quality, high-level metadata is required. | IIIF is a model for presenting and annotating digital content on the web, including images and audiovisual files Digitized AV materials are often in silos, which makes it hard for users to discover, gather together, compare, analyse and manipulate images. IIIF enables a number of core features: **Deep Zoom**: view and zoom in closely to see all the detail of an image **Sequencing:** navigate through a book or sequence of archival materials **Comparisons:** bring images together and put them side-by-side. This can enable researchers to bring together images from different collections, maybe material with the same provenance that has been separated over time. **Search within Text:** work with transcriptions and translations **Connections:** connect to resources such as Wikidata **Use of Different IIIF Viewers:** different viewers have their own features and facilities. To put it more simply, IIIF is like a layer above the data stores that hosts content These images are accessed through the IIIF layer, which allows agents to create viewers and tools for the data held in all the stores Different repositories have their own data stores but can share content through the IIIF API This helps reduce data silos The IIIF has the following APIs: **Image API** This API delivers the content (or pixels). The image is delivered as a URL, and the URL is structured in an agreed way. **Presentation API** This delivers information on the presentation of the material, such as the sequence of a book, for example, or a bundle of letters, and metadata about the object. **Search API** Allows searching within the text of an object **Authentication API** Allows materials to be restricted by the audience. This is useful for sensitive images or images under copyright that may have restrictions **IIIF Viewer** IIIF images are served in a standard way, and any IIIF viewer can access them Some common IIIF viewers include the following: The Universal Viewer: https://universalviewer.io/ Mirador: https://mirador-dev.netlify.app/tests/integration/mirador/ Archival IIIF: https://archival-iiif.github.io/ Storiiies digital storytelling: https://storiiies.cogapp.com/#storiiies How the Archives Hub Will Use IIIF **Hosting Content:** They are starting a 15 month project to explore options for hosting and delivering content. This will also include the use of an IIIF Image API **Creation of IIIF Manifests:** These are part of the Presentation API. Manifests contain a sequence, for example the ordering of a book, as | As with the other blog posts, one of the limitations of this source is its format. The blog post does not provide citations and does not go into depth. However, it is a useful starting point for the study. |

			well as metadata such as a title, description, attribution, rights information, table of contents, and any other information about the objects that may be useful for presentation.	
			○ **Providing an Interface to Manage Content:** This is useful for any image store, however, the author claims it may have specific implications around the metadata and records put into an IIIF manifest	
			○ **Integrating an IIIF viewer into the Archives Hub:** This will allow users to access the images hosted in the Archives Hub IIIF	
			○ **Assessing Image Quality:** Archives Hub will access various resolutions to determine which will strike the ideal balance between size and image quality for the end user.	
			○ **Looking at Rights and IIIF:** The team will examine how IIIF complies with copyright and other legal rights	