Artificial intelligence in the media industry

The symbiosis between media and machines

Author: Yvonne Thomas, Product Manager at Arvato Systems (www.it.arvato.com/broadcast)

Today's media industry survives on data. Texts, photos, videos, graphics, audio files and social media feeds in thousands of different formats and from countless different sources are edited, published and distributed on a daily basis. As a result of technological developments in media and IT technology and a proliferation of digital storage options, more data is being generated today than ever before. According to Statista, the market data and research portal, the photo service company CeWe, for example, developed almost 830 million digital photos in 2005. Just 12 years later this had risen to over 2.1 billion. Statista also provides information about a study by Bitkom from 2017. Bitkom estimates that the number of photos taken worldwide will have almost double between 2013 and 2017. That is a jump from 660 trillion to 1.2 quadrillion images. So it is clear that there are huge amounts of data accumulating around the world, and the media industry is certainly no exception. Although the data is collected and stored, no one – editor or media manager – is in a position to evaluate this flood of data and make it usable: Big Data simply cannot be processed using conventional databases and management tools. But with AI – artificial intelligence – this instantly becomes feasible.

Lifelong learning – even for computers
Artificial Intelligence (AI) is capable of enriching and tagging staggering amounts of data sets and recognizing patterns and correlations. In the media industry itself, there is much talk about pairing analytics with machine learning. In particular, by using machine learning methods, the system improves its own performance and results become more and more accurate.

When the Royals got married – an example of using AI in TV broadcasting
It was probably the media event of May 2018: the British royal wedding. It can be assumed that the various representatives of the press recorded countless hours of video material and delivered these to their respective editorial offices. Reporting also took place simultaneously on various TV and radio channels and on the Internet. This was an advantage for those who were able to communicate news to their audiences – an advantage that the other media representatives did not have. Many fans were curious to see which celebrities were in attendance. The reporters could of course fall back on the official announcements from the palace. But did these include information about every star and celebrity who took part in the celebrations? Not likely. Additional information, however, was provided to an editorial team who had their miles and miles of film analyzed using AI software specifically engineered for facial recognition. Those VIPs, who were clearly captured on film and whose faces were known to the software, could thus be identified. But even better still, the system was also able to include these new images in its "memory" and so will be able to recognize these same celebrities even better and faster in the future.

Visualizing metadata in a user-friendly way

However, this analysis only produces unstructured metadata. For example, you can search through video files for a specific object, such as a logo. The system then finds out in which sequences this logo can be seen and displays the metadata as a list with the corresponding time codes. For the user, this data is pretty difficult if not impossible to work with. This is where a media asset management system (MAM system) comes in, which automatically assumes ownership of the information the data from Connector, a defined REST interface. The MAM system then displays the data in a form that the user can quickly and easily work with. Using our above example, the logo recognition data would be displayed on a video timeline. All other associated data would also be directly displayed in a structured form on the asset. For example, several AI-generated shortlists are visible on the same asset.

Differentiation at a glance

Some users may want to reassure themselves by manually checking the data in the system for accuracy, adjusting it if necessary, and thereby training the system further. This is where a transparent MAM system is an advantage, such as the VPMS from Arvato Systems, which clearly shows which data is AI-generated and which data has been entered manually.

How the machine goes to school when the curriculum is set

Machine learning can be divided into Supervised Learning, Unsupervised Learning and Reinforcement Learning. With Supervised Learning, the system uses its ability to recognize characteristics and thus classify data. A model is built using example data. The system learns that different typical properties of the data define their affiliation to a certain group. If new data records are added to the system, the computer recognizes their properties and assigns the data to groups. In practice, for example, texts can be classified automatically. The system recognizes certain buzzwords or groups of words and assigns the text to a genre. It becomes particularly interesting when the system is not limited to one text format, but rather can analyze and thus cluster documents as well as audio and video recordings.

Training without a timetable

In Unsupervised Learning, on the other hand, AI will uncover the as yet unknown relationships between the data, find recurring patterns and create a structure for the data itself, which are called clusters. The number and type of clusters will most likely change as new data is added. Typical applications for Unsupervised Learning in media are speech recognition and speech-to-text transcription. To be useful, however, the system itself must first of all understand what language is being used. The better and more extensive the test material that is fed into the system, the easier it will be for the system to do this. If, for example, the system recognized that French was being spoken in this video, this is where the real difficulty begins: In Paris, for instance, French sounds a little different than it does in Provence while the dialect spoken by North African immigrants differs from that spoken in Belgium. However, by using suitable algorithms, the system learns to assign different pronunciations to the same written word.

Learning according to the reward principle

Reinforcement Learning is a highly complex process by which Artificial Intelligence performs defined actions in a particular environment as soon as an exactly defined state occurs. The environment reacts to this with a positive evaluation – a "reward" – or it judges the action negatively. AI remembers this evaluation and knows which action is the right one as soon as the same conditions occur again. Here's a practical example: For online media to function 24 hours a day, 7 days a week, the technology involved must function smoothly. Server outages, unacceptable sound quality or security failures are just not acceptable in this industry. AI can ensure that the equipment is functional at all times by calculating the components' probability of failure and taking timely countermeasures before an emergency actually occurs.

Letting the media experts lend a helping hand

By nature, people who work in media are usually open to new things. However, their industry has been particularly affected by the technological upheavals of recent years, and there seems to be no end in sight. Still, when introducing an AI tool to users, a number of rules must be observed to ensure that the target group warms to this change and wants to use the media asset management system. As a rule, work processes – and this can possibly include the organizational structure of a media company – have to be adapted to meet changing needs. The system must of course fit into the existing IT landscape and the interfaces must ensure a smooth exchange of data between the MAM system and the various data sources.

No fake news floating around in the Cloud

It is recommended to combine AI with a Cloud solution right from the outset. The intelligent analysis of Big Data takes place in the Cloud and, despite the enormous computing processes involved, does not affect in-house IT systems – even during peak loads. However, those who deal with information processing via the cloud must take suitable protective measures to ensure the integrity of their cyber and data security. And this includes "fake news": After all, false messages can also be generated with AI and then spread relatively easily. Fortunately, AI is also able to detect and counter fake news.

Lots of scope available for AI

Artificial Intelligence is already used extensively in facial recognition, speech analysis and photo classification. Other typical applications in the media world include scene analysis, audio track extraction and intelligent raw cuts. The better the use cases are defined, the better the system can be trained. And the more practice-relevant the learning is, the better the results the intelligent system will be able to produce. Also, when deciding to use AI in production work, it is necessary to become well versed in data security and GDPR.

AI is already proving its worth. So it will probably become indispensable in the future. Skeptics should therefore definitely get to grips with the topic and perhaps try analytics services in smaller projects to begin with, because the time has come to get into AI.

The Author:

Yvonne Thomas graduated in Television Technologies and Electronic Media Engineering from University of applied Science Wiesbaden (HSRM), Germany, in Oct. 2010. She received a prominent award of the ARD/ZDF Academy for her thesis in September 2011 at the IFA in Berlin.

Following these studies she started beginning of 2011 to work at the European Broadcasting Union (EBU) in the Technology & Innovation Department as Project Manager. She was responsible for projects and strategic programs on 3D and Future Television technologies, such as UHDTV or LED studio lighting and highly involved in its standardization.

Since September 2015 Yvonne joined Arvato Systems´ Broadcast Solution division. As a Product Manager she is responsible for the journalistic frontend "MediaPortal" of the Video Production Management Suite and follows new trends in the media landscape, like AI and machine learning.

At NAB 2017 Yvonne was presented with the annual Technology Women to Watch Award from TVNews Check.