

dpp[®]

THE CLOUD FOR MEDIA:

Automating Media

Enabled by

VU BIQUITY[®]
An Amdocs Company

FOREWORD

Over the last year we've seen an unmatched proliferation of D2C (direct-to-consumer) video services vie for consumers' attention. Many media companies have made huge investments in their D2C offerings in order to compete in an ultra-competitive marketplace.

One of the critical components for achieving marketplace success is the implementation and execution of media automation. Content supply chain workflows have been steadily moving towards more advanced automation, and now more than ever, its importance can't be underestimated. More automation equals less manual intervention. Yet, a greater reward is now understood. With automation, companies can launch into the marketplace more quickly and efficiently, thus delivering a cost benefit too.

For many media companies, automation progress has been achieved via cloud-enabled, software-defined media supply chains, often to handle large content libraries. The cloud environment offers ideal flexibility, allowing automated tasks to flourish and enabling companies to derive maximum benefit from their usage. Media automation possesses a variety of competencies, and those who best understand them will create the most effective game plans for success.

An avid supporter of the DPP, Vubiquity is proud to have contributed to this report. We hope you benefit from the informative use cases and valuable input from experienced leaders. Cloud-based automation is here to stay, and its growth will continue to surge in the future.









Raman Abrol
Chief Executive Officer, Vubiquity

OCTOBER 2021

THE CLOUD FOR MEDIA

Automating Media

Introduction	4
Executive Summary	6
Contributors	7
1 Laying the Groundwork	9
<p>CASE STUDIES</p> <div style="display: flex; justify-content: space-around;"> <div style="border: 1px solid black; padding: 5px; text-align: center;">  <p>11</p> </div> <div style="border: 1px solid black; padding: 5px; text-align: center;">  <p>12</p> </div> </div>	
2 Defining Automation	18
3 The Progress of Media AI	19
<p>CASE STUDIES</p> <div style="display: flex; justify-content: space-around;"> <div style="border: 1px solid black; padding: 5px; text-align: center;">  <p>23</p> </div> <div style="border: 1px solid black; padding: 5px; text-align: center;">  <p>26</p> </div> </div>	
4 Automating the Media Factory	30
<p>CASE STUDIES</p> <div style="display: flex; justify-content: space-around;"> <div style="border: 1px solid black; padding: 5px; text-align: center;">  <p>31</p> </div> <div style="border: 1px solid black; padding: 5px; text-align: center;">  <p>35</p> </div> </div>	
5 Learnings About Machine Learning	38
Conclusion: Delivering Value Faster	45

Introduction

The processes of media production and distribution have been on a path of increasing automation over many years. But the progress has been slow.

A 2018 DPP workshop and report, [AI for Real](#), found increasing deployment of rules based automation, and some real-world use of machine learning solutions. But considering the hype around AI at the time, it was sobering to uncover the extent of the challenges faced in the further advancement of such technologies. They ranged from skills gaps to commercial constraints; from ethical considerations to matters of intellectual property.

In 2018 the scale of hype around AI was matched only by the extent of the challenges

Nonetheless, it felt as though the industry was on the cusp of major steps forward in the development of useful and business focussed AI tools.

In the years that have followed, the biggest technological transformation for many media companies has been the move to software defined media supply chains. Almost without exception, these are enabled by the cloud.

No wonder, then, that [The Cloud for Media](#) created such engagement and interest among DPP members in 2020.

So will these cloud based supply chains enable the potential of machine learning to be fully realised? Can automation of the media factory deliver greater business effectiveness? Has the time for an automation revolution come?

As 2021 closes, we'll plot where we are now on the automation roadmap

As 2021 closes, we'll plot where we are now on the automation roadmap, based on input from experts across the media industry. The real-world experiences of media companies are used throughout to illustrate the realities of cloud based automation today, the successes that have been achieved, and the challenges that remain.

THE CLOUD FOR MEDIA SERIES

Automating Media is one of four reports in the DPP's 2021 series, *The Cloud for Media*.

Each document in the series examines a different aspect of media in the cloud; the others focus on [Post Production](#), [Playout](#), and [Streaming at Scale](#).

These reports do not assume deep technical knowledge, but a high level understanding of key cloud technologies and terminology may be useful. We have therefore provided a brief accompanying guide, [Cloud Technology](#).

Executive Summary

1

AI now solves specific business problems

A clear change has occurred in industry attitudes to AI since we examined them in 2018. Hype has subsided, and with it the unrealistic expectations that lead to disappointment. Instead, companies have moved beyond generic AI tools and are instead solving specific and tangible business problems with focussed solutions.

2

Automation at scale requires cloud

Automation is often most effectively applied to large bodies of content, to which a great deal of computing power can be applied. This makes the cloud an ideal home for automated media factories.

3

Partnerships enable automation

Vendors and customers now recognise that their complementary expertise must be harnessed to develop effective tools. Commercial agreements that incorporate shared risk and reward are the foundation of effective partnerships, helping to overcome earlier concerns about intellectual property and training data.

4

Automation is powering the D2C revolution

Major global media companies have made huge investments into direct to consumer offerings. But preparing the vast libraries of content would not be practical without automation. From deduplicating versions to conforming audio, manual processes would be economically unviable.

5

The automation dividend is reduced time to revenue

It goes without saying that companies want to save costs, and automation can help in this regard by reducing manual effort. But in almost every use case and conversation, a more important factor emerges: automation speeds up processes, reducing time to market – and therefore time to revenue.

Contributors

The content for this report has been gathered through workshops and interviews with subject matter experts from across the industry. Valuable input has been provided by our Expert Sponsors: Codemill, Pikel, and Vubiquity; and by our Contributing Sponsor: Base Media Cloud.

Although the content of this report has been informed by these discussions, it should not be assumed that every contributor shares all the views presented here.



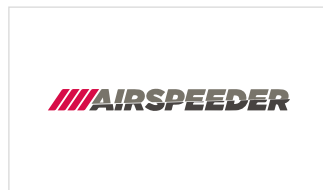
TIM BURTON
Managing Director



DAVID KLEE
VP – Strategic Media Solutions



MICHAEL DUFFY
CTO



STEPHEN SIDLO
Head of Media



CLAIRE FEENEY
Head of Technical Operations



LEE ATKINSON
Worldwide Tech Lead –
Broadcast Media



SANDY MACINTYRE
Vice President News.
Director Key Initiatives



JASON PERR
VP – Software & Product
Development



BEN FOAKES
Founder & Managing Director



PHILIP ROBINSON
Lead Architect



NEIL ANDERSON
UK Country Manager



JONAS SANDBERG
Senior Solution Architect



JOSH DERBY
Group Vice President, Advanced
Technology & Development



DAVID KELLER
Regional VP Delivery Acceleration



RAINER KELLERHALS
Managing Director, Media &
Communications EMEA



SHYAM VISAMSETTY
Managing Director



JIM KARCHER
Master Control Operator



MARK CHRISTIE
CTO



CRAIG SEIDEL
CTO



HEWSON MAXWELL
Head of Access Services
Solutions & Platforms



TANIA HOQUE
Director – Digital Products



ANTONY JOYCE
Head of Technology Architecture



JITEN MISTRY
Senior Solution Architect



MELISSA FECHER
Media Engineer



DEE WASSELL
Global Director – Media Support



KYLE KREOLL
Senior Post Production Engineer



GORDON PRINCE
Senior VP – Operations &
Technology Solutions



STEWART CURTIS
Director – Projects & Delivery

Laying the Groundwork

KEY INSIGHTS

- ▶ Automation is at home in the cloud, and basics such as Media Asset Management (MAM) are as important as ever.
- ▶ Cloud content repositories enable multiple processes to operate on media without moving the content around.
- ▶ Major broadcasters are now reaping the benefits of the agility that has long been promised by cloud based software architectures.
- ▶ Agility is also enabled by automated pipelines for building, testing, integrating, and deploying software.

The Cloud for Media explored the benefits offered to media organisations by moving operations to the cloud. While each company seeks to achieve its own specific objectives, a majority identified agility as the most significant benefit of the cloud. Software defined infrastructures and supply chains enable technology teams to better keep up with ever changing business needs.

A majority of companies identify agility as the most significant cloud benefit

Another central feature of cloud workflows is consumption based pricing, offering a much more granular understanding of the costs incurred in the operation of individual business processes. As identified in [Next Gen Supply Chain](#), the eyes of media executives fall increasingly on these detailed costs of creating and delivering each piece of media.

So any media company moving to the cloud is likely to seek greater efficiency, delivered in part through automation. It is fortunate, therefore, that the agile and flexible nature of the cloud makes it an ideal environment for automation. Experiments can be performed, new techniques can be explored, then workflows can be scaled up once they are proven.

DATA GRAVITY

As companies move more of their workflows and content to the cloud, there is a tendency for that content to form its own gravitational pull. That is to say that once the 'mass' of a large data set is sufficiently great, it becomes natural for processes and applications to be drawn to the data, rather than data being moved for processing.



Content in the cloud creates its own gravitational pull

The analogy with planetary gravity was first made by David McCrory [in 2010](#), and his resultant term 'data gravity' seems especially relevant to the media industry. The large file size of video, and the expansive and diverse libraries of content held by many media organisations, mean that there are significant efficiency gains to be made by processing content where it resides rather than moving it.

The first step towards automation is usually to centralise media in the cloud and catalogue it, as we will see in the first of six case studies in this report.

CASE STUDY

AIRSPEDER

FOUNDATIONS FOR THE FUTURE



It is telling that even those on the leading edge must begin with seemingly basic steps: make media accessible in the cloud, with appropriate catalogue and access management, such as a MAM system.

BASE has worked with a roster of brand new sports leagues and producers, including Formula E, Extreme E, and upcoming flying car series Airspeeder. In each case, the cutting edge technology of the sport is matched by aspirations of cutting edge technology in their media production and distribution. But in each deployment, a managed media repository is the first step.

With a three dimensional race track in the air, around which manned octocopters will race, the production requirements for Airspeeder are leading them towards mixed reality using gaming engines. But none of that could have been approached before they got the basics right, by implementing their cloud based MAM.

Synchronisation agents were installed to monitor ingest folders in Australia and the UK, enabling content to be automatically moved into their central cloud storage. It is ingested in the MAM, with proxies generated to allow teams in both continents to browse and work on the media.

“ I’m a big proponent of the foundation work being done before we move on to a more advanced system, because it just sets up good habits within the business.

STEPHEN SIDLO, AIRSPEEDER

Although the league is new, the organisation has been creating content for five years, and so AI powered automated transcription and video analysis have been applied to make that back catalogue searchable. Further automation will be added over time, such as for syndication of video to global outlets, or the association of vehicle telemetry data with related media.

The principle of data gravity is reflected in this connection of video and non-video data, as well as in remote production and editing of the content in the cloud.

The Airspeeder team never considered deploying their infrastructure anywhere but the cloud. It’s the obvious choice for a new media team. But what about a major broadcaster with many years of existing on premise investment? For UKTV, the cloud offered the chance to reimagine large parts of their business.

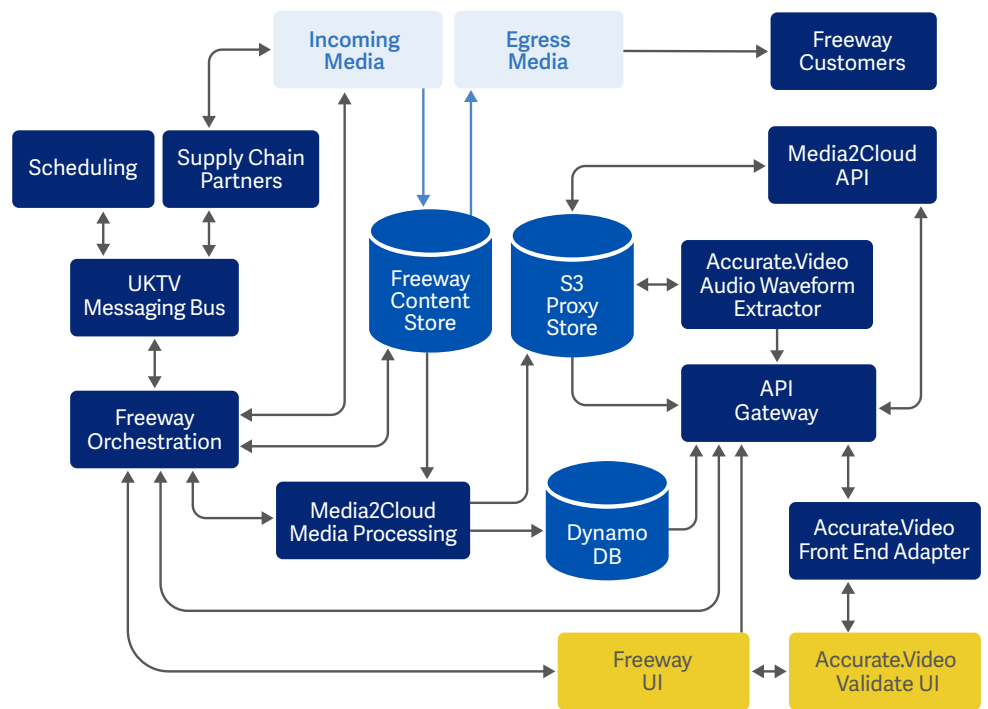
CASE STUDY



CRUISING ON THE FREEWAY

For British broadcaster UKTV, the move to the cloud has been central to reengineering its media supply chain. It moved from a largely outsourced model of media management and processing, to an in-house managed cloud based system known as Freeway that went live in September 2020.

The cloud has been central to reengineering UKTV’s media supply chain



Working with its partners Codemill and eCreation Media Technology, UKTV implemented an off-the-shelf MAM system in the AWS cloud. Workflows and user interfaces (UIs) were built to UKTV's specific needs, integrated with the MAM's APIs. The Freeway UI, for example, has custom components and also makes extensive use of AccuratePlayer and other AccurateVideo components for video playback and review.

Ongoing developments will soon allow content review functions – including technical quality control and editorial compliance review – to also be performed with the system. Browser based tools will allow users to mark up content where edits are required, and generate Edit Decision Lists (EDLs) to be imported into a non-linear editing package (NLE) for the final edit.

CHANGING LANES

Despite the early success, the UKTV team quickly found that the solution they'd implemented didn't quite deliver the level of agility, transparency, and control they were looking for. The MAM vendor was able to offer a cloud hosted solution, but it hadn't been built from the ground up for the cloud. It was a virtualised deployment running on reserved computing instances, and architects at UKTV were keen to move to a highly scalable solution built mainly around serverless functions.

So only one year after launch, the Freeway team is already implementing more modular, cloud native components based on the AWS Media2Cloud reference architecture. This new version is called Mediaway.

Only one year after launch, UKTV is implementing more modular, cloud native enhancements

Media2Cloud is not a single software, but rather a set of services that combine to provide many functions that might be offered in a traditional MAM.

This move has put the agility of UKTV's overall solution – and their partners – to the test. The great promise of modular software, integrated via APIs, is that components can be swapped out and adapted over time. But most Solutions Architects will tell you it's not as simple as it sounds.

Many of AccurateVideo's components had to be updated to integrate with different back-end services, especially the new serverless components. But for the Codemill team, that effort is paying off as other clients are already implementing a similar architecture. And for UKTV, success is demonstrated by deploying a second generation of their architecture without causing disruption to users or the business.

WHY REBUILD SO SOON?

Many observers might consider it remarkable to reengineer such an important system within its first year of deployment. But software components have not been swapped without good reason; it's all been in service of the business goals.



It's really important to have a clear strategy for what you're trying to do. It's like a lego set, and you need to know what you're building.

ANTONY JOYCE, UKTV

It is critical for UKTV's technical operations that they have a high degree of transparency from their systems. The more monolithic software tools were seen as behaving too much like 'black boxes', without the ability to understand and interrogate individual components, extract logs, aggregate them across the whole system, and so on.

Monolithic software tools behaved too much like 'black boxes'

Swapping to cloud native components also enabled enhanced system security, because every component operates using identity and access management (IAM) system provided by AWS.

And by implementing a more modular architecture, the flexibility to adapt and extend workflows is further enhanced.

CHOOSING COLLABORATORS

UKTV has a considerable in-house development capability, but partners such as Codemill have still been critical to success. They were required to update and adapt their product line to meet UKTV's needs, breaking it apart into more granular components, and adding new features along the way.

A shared vision was key to success. From the start, it was recognised that a combination of vendors, products, and even open source solutions would be needed to deliver the end to end solution.

“ The entire platform is designed around integrations with other systems.

JONAS SANDBERG, CODEMILL

The AWS Media2Cloud solution itself is intended only as a starting point, and most implementers adapt it, add their own components, integrate different tools, and so on. UKTV is no exception.

CLOUD RELIANCE

Despite all this flexibility, there is no utopia of true vendor agnosticism. By not only building on the AWS cloud platform but also using their reference architecture and software services, there is a significant reliance on Amazon. As discussed in [The Cloud for Media](#), choosing one or more cloud platforms is a business decision that comes with risks, but also with benefits.

“ You will tie yourself in knots and it will cost you a fortune to be cloud agnostic. So we’ve gone with one provider. We’ve drunk the Kool Aid, for sure. But we’re building with principles of interoperability wherever we can.

ANTONY JOYCE, UKTV

The benefits of working closely with a cloud provider include access not only to a scalable infrastructure, but also to a team of experts, and an ecosystem of software partners. And by focussing on a single platform, the UKTV team has been able to develop DevOps and software skills in-house based on a common set of core tools and competencies.

DIRECTION OF TRAVEL

The development of UKTV’s cloud based media supply chain so far has transformed the business, bringing key functions in-house and giving the organisation greater control of its media and its workflows. As you would expect, a number of lessons have been learned along the way.

**UKTV’s cloud supply chain
has transformed the business,
bringing key functions in-house**

Multi-vendor architectures are not without complexity. They open up the possibility of duplication: should the audio waveform rendering function be performed by the AWS toolset or the AccurateVideo one, for example? And there are inevitably gaps that need to be filled, sometimes requiring the development of UKTV’s own bespoke software services.

Time must also be spent considering the fine details of the software architecture and operation; a structured approach pays dividends.

“ How you label the services, how you define them, the formats you use for system logs, and how you aggregate them. These are really important factors. They're not sexy, but they're crucial for operations, for incident management, and for cost management.

ANTONY JOYCE, UKTV

With the foundations in place, the next phase of the project is now underway, taking automation to the next level by significantly increasing the use of machine learning.

AUTOMATING THE AUTOMATION

For UKTV, as for an increasing number of media companies, contemporary software deployment and delivery processes are central to the use of cloud technology.

Cloud infrastructure should be managed using the principles of infrastructure as code, enabling automated and repeatable deployment. The same concepts are applied to software components by using containerisation, serverless functions, and software as a service tools, to ensure that each one can be deployed or accessed programmatically.

The automation of software updates, deployments, scaling, and management, will result in reduced manual effort. But it also has a more important impact. Unlike an automated media processing task, which operates on a different piece of content each time, the automated deployment of infrastructure and software should have the same inputs each time it is run. This leads to consistent results every time, dramatically reducing operational risk.

Automated deployment delivers consistency and reduces operational risk

While automated deployment isn't strictly essential for delivering automated media workflows, strong technical foundations are highly beneficial for building and maintaining supply chains that can continue to meet business needs even as they change.

Defining Automation

Automation, and indeed Artificial Intelligence, are broad terms that incorporate a range of different technologies. It is worth reviewing the definitions used in 2018's [AI for Real](#), as they are just as relevant today.

Artificial Intelligence is a broad term that incorporates a range of different technologies

Rules based automation involves capturing something that is intrinsically contained within humans.

- Robotic Process Automation (RPA) involves a system performing a defined and reproducible set of steps, which previously were undertaken by a human. For example, entering the same data into multiple systems.
- Another rules based approach codifies expert knowledge as rules that can be executed in software. A simple but common example would be processing a piece of content differently based upon its technical parameters or accompanying metadata.

Machine learning (ML) approaches enable the computer to learn from data, rather than only executing predefined rules. ML can also be subdivided into two types.

- The first employs an algorithm trained from a static data set, from which patterns are learned. The machine then applies this learning to other data, for classification, tagging, pattern recognition, and so on.
- The second approach is one in which the machine learns not just once from a training data set, but also continually from feedback as it is used. This is often referred to as **deep learning**. It may include the use of neural networks – an integrated group of individual AI processing units, which mimics the network of neurons in the human brain. Neural networks are often used for voice recognition, for example.

The Progress of Media AI

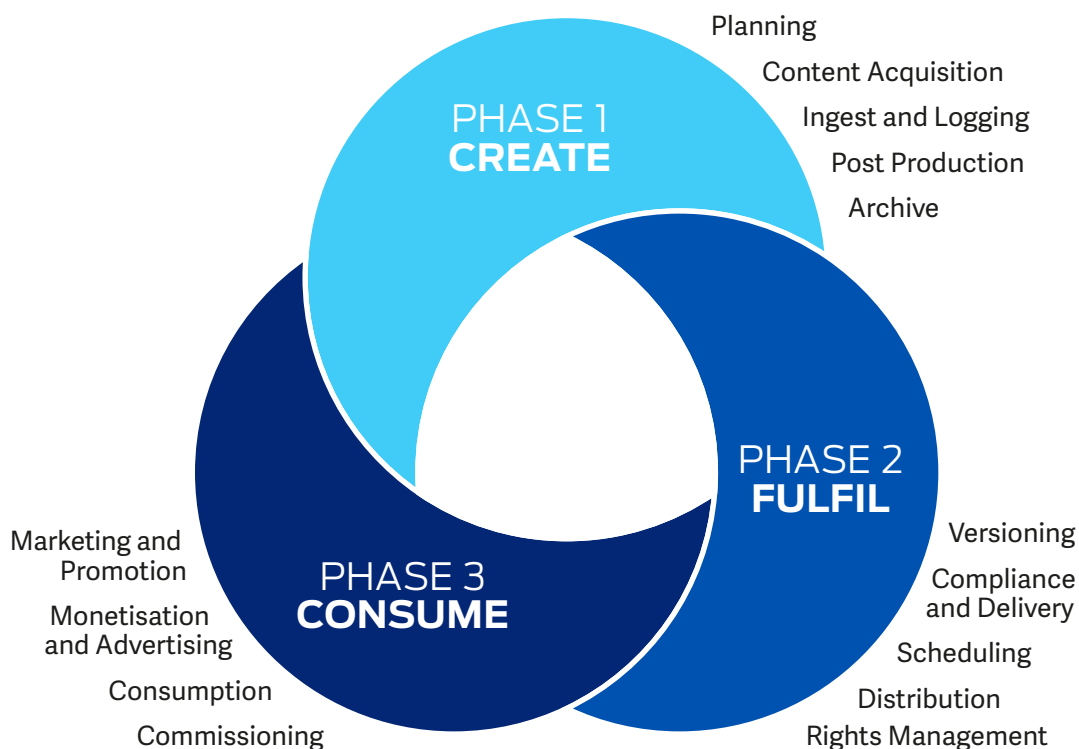
KEY INSIGHTS

- ▶ There is potential for the application of AI right across the content supply chain.
- ▶ Areas with the greatest potential include versioning and rights management, with post production, archive, monetisation, marketing, and consumption close behind.
- ▶ In 2018, excessive hype led to disillusionment around automated logging, but as technology has improved and proven itself, expectations have risen.
- ▶ Steps in the versioning process are being automated, from identifying text elements to automated captioning.
- ▶ The use of AI around content consumption remains largely focussed on surfacing the right content – and crucially the right advertisements – to viewers.

Almost three years ago, 21 experts from the DPP community examined the potential of AI to deliver improvements in efficiency, creative or business opportunity, or consumer experience. They also assessed the actual benefit being delivered at that point in time.

In the summer of 2021, a different set of 21 experts joined a DPP workshop to discuss whether the potential benefit had changed, and where there had been growth in the actual benefit seen in real deployments.

In each case, a three phase model of the content lifecycle served as a framework for the analysis. In this chapter, we will review what our 2018 experts said about the use of AI in each of these phases, and see how our 2021 experts felt the situation has changed.



PHASE 1: **CREATE**

■ THE POTENTIAL

In 2018, our experts saw broad potential for AI to deliver benefits across all of the workflow stages in the creation phase.

Production planning could benefit from enhanced resource planning. And for live and fast turnaround content, greater business intelligence on viewer behaviour and reactions could inform the content as it's produced.

The possibilities in acquisition, ingest and logging were more extensive. Once again, live and fast turnaround content was seen as a key beneficiary, with the structured and predictable environment of a sports field or stadium offering an ideal environment for robotic cameras, AI shot selection, play analytics, automatic highlight generation, and more. Other use cases included metadata generation on ingest (face recognition, shot classification, transcription, and so on), and quality control.

In post production and archiving, the potential benefit was seen as higher still. Selection of best takes, content enhancement, digital effects, automatic assembly and editing, intelligent content reframing, and even automatic product placement featured on the list of possibilities raised by participants. So did archive content selection and recommendation, metadata tagging, predictive hierarchical storage management, and more.

Our 2021 expert group felt that the efficiency gains to be had through automation in the ingest and logging phase are especially significant, and had been underestimated in 2018.

They also called out growing excitement about virtual production capabilities, and other areas including personalised content generation, and synthetic voice generation for dubbing and additional dialogue replacement.

■ MATURITY: THEN AND NOW

Back in 2018, the maturity of real-world implementation was seen as low for content planning, and middling across the rest of the creation phase. A general disillusionment about the efficacy of generic models for content tagging and analysis – which had been much over-hyped at previous years' tradeshows – meant that real business benefit was only being delivered in limited areas.

Three years later, progress has been significant, if still somewhat short of a complete revolution.

Progress in automating ingest and logging has been significant, but somewhat short of a complete revolution

The disillusionment around automated metadata generation is starting to subside. Service providers including BASE and Codemill reported that this is the most common AI function to be requested by their customers.

The use of automatic transcription services to enable content search was highlighted, with some media companies using the transcripts as a basis for subtitling (albeit most often with human review and refinement), and others applying automatic translation.

When it comes to archives, the promise has long been that AI analysis of content catalogues could open up new revenue opportunities by identifying content of value. The challenge, however, is that the cost of running ML processes on an archive can be large, while the return on investment is difficult to anticipate.

But this year, the conversation has moved towards more practical solutions. Those with topical content libraries, such as Rogers Sports & Media and VICE Media, focus efforts on recent content, as it is dramatically more likely to be reused, and therefore has higher value. Others are using machine learning to deduplicate content before running other analyses, reducing wasted processing and saving on storage.

And once useful content has been identified, machine learning models are being employed to identify the correct version: the one with the appropriate language track, for example.



The promise of AI is to open up new opportunities for archive reuse, but the ROI is difficult to predict

Among our expert contributors, it was clear that each was thinking more deeply about how their specific business needs could be met with automation. The BBC, for example, is looking to automated analysis of its archive to assist in delivering against its diversity and inclusion aims. Experiments are underway with the hope that machine learning can help uncover untold or overlooked stories, exposing a richer diversity of content from the broadcaster's back catalogue.

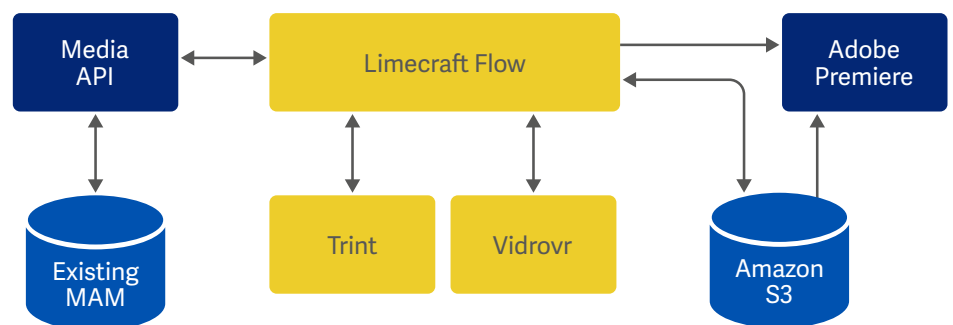
CASE STUDY

AP

BREAKING NEWS WITH AI

Some organisations are starting to apply semantic analysis to content, enabling a range of automated processes.

Associated Press (AP) has been looking for a solution to help create shot lists and rough cuts from content such as press conferences or live news events. It has worked with a collection of vendors including Vidrov, Trint, and Limecraft, to create a proof of concept system, integrated with AP's own platform.



The system uses automated transcription, shot detection, and visual analysis to develop an accurate description of every shot. Following this, automatic shot lists or rough cut EDLs are created, incorporating some basic editorial rules. For example, a press conference might generally start with a wide shot, then follow the speaker close up, cutting away to audience wide shots when there is significant reaction noise.

“ We didn't try to boil the ocean. It's a complex challenge, so we took it step by step. In a first iteration, we used a system that was capable of recognising 250 international politicians and this delivers great results for approximately 20% of our coverage. In the next iterations, we will refine the model and increase the value of the output accordingly.

SANDY MACINTYRE, AP

SPECIFIC USE CASE

AP is unusual, in that it is a wholesaler of content. AP staff commonly produce a relatively long edit of a news story (for example, three minutes), which is distributed to media companies that subscribe to AP's feeds. In turn, those customers will edit the story down in their own style, for example to create a one minute output.

This means that, where other organisations might require finely polished timings in the edited output, a rough cut generated by the system might be sufficient for AP's needs. Nonetheless, EDLs are generated for human editors to refine; a completely automated system is still some way off.

But while the specifics of the system are both ambitious and bespoke, the individual steps are broadly applicable in many media organisations: frame accurate shot detection, human-readable shot descriptions, and editorial construction.

PHASE 2: **FULFIL**

■ THE POTENTIAL

Across the five workflow stages in the fulfilment phase, our 2018 experts saw greater variation in the potential of AI to deliver benefits. Versioning and Rights Management were felt to have the greatest potential of all the stages in the media supply chain, while distribution had the lowest potential.

**Versioning and rights management
have the greatest potential benefit
from the use of AI**

Versioning was expected to benefit from enhanced efficiency, especially from the opportunity to deliver more services to more users. Dubbing might be transformed by the combination of speech to text, language detection, translation, and speech synthesis; while avatar based on-screen signing was also seen as a possibility.

Despite continuing advances in voice recognition and synthesis, this year the conversation turned to seemingly more modest and practical goals, such as the detection of on-screen text, enabling the visual elements of content to be translated from one language to another.

Technical quality control has been using different forms of AI for some time now, but editorial and legal compliance is also an area of great promise. Machines could potentially be employed to detect swearing, nudity, violence, and so on. Yet in 2021, just as in 2018, discussion was tempered by concern over the high stakes of any mistakes. It seems that advances in machine learning have not yet allayed those concerns.

In 2021, just as in 2018, users feared the consequences of any errors made by AI driven compliance checks

Another area of great potential is scheduling, where there was interest in the commercial opportunity generated by automated construction of commercial breaks, including the placement of contextually relevant advertising. Intelligent scheduling of personalised OTT channels was also of interest to some in 2018, though the boundary between this use case and personalised 'Watch Next' recommendations on VOD services is blurred. That may explain why there was much less excitement in this year's group around content scheduling.

The areas of potential in distribution are more limited. Our 2018 contributors discussed predictive capacity planning for Content Delivery Networks (CDNs), and AI enhanced encoding. Since then discussion around shot based and scene based encoding has intensified, indicating a rise in the potential benefit and maturity in this area.

The exceptionally high potential seen in rights management is largely due to the anticipated application of content recognition algorithms to enable identification of copyright content and to manage rights payments. This could lead to more sophisticated rights brokering and pricing optimisation too.

■ MATURITY: THEN AND NOW

The explosion of high profile international VOD platforms has resulted in a growing need for content reversioning. So it might be hoped that automation would be central to this. In reality, content providers in 2021 told us that their developments in this space were nascent, although many are actively investing in more automation going forwards.

Specific steps in the process are starting to be automated, as illustrated by Discovery's use of machine learning to assist with the generation of textless content (see [Discovering Versioning Solutions](#)). And as predicted, there has been some growth in the use of automation for accessibility services, with captioning being the most mature.

Automated quality control algorithms continue to grow in sophistication, but the progress towards automated editorial compliance is being made in baby steps, not strides.

In [AI for Real](#), assessment of the maturity of implementation for rights management had to be separated between linear platforms – where there was very little real-world use of automation – and online platforms. Then, as now, content identification algorithms are used extensively on social and other platforms to enforce rights. This technology is now available to content providers and platforms, extending its reach beyond the largest platforms such as YouTube and Facebook.

Content identification algorithms are used on social platforms to enforce rights

CASE STUDY



COST EFFECTIVE CAPTIONING

For local TV stations, captioning can be a major cost, and automation can deliver considerable savings. And when breaking news comes into the equation, it can be almost impossible to find human captioners and get them up and running at short notice. For both of these reasons, PBS39 (WLVT) in Pennsylvania turned to AI driven captioning.

The ability to switch on automated captions at the click of a button is a huge boon to the station, enabling them to meet their accessibility commitments across all output. Of course this is useful for live coverage of breaking stories, but in the fast-paced world of news, even prerecorded packages may be created very rapidly. As a result, the station often uses live captioning when playing out fast-turnaround or late-delivered content too.



In general, news is a strong use case for automated captioning, because it is usually characterised by a single person speaking at any given time, most often with good audio quality and well constructed dialogue.

Nonetheless, automated speech to text isn't perfect, and live TV can be challenging and unpredictable. To help bridge the gap, provider Red Bee Media offers two services, as part of their delivery to PBS39:

- ▶ An AI driven fully automated option, which has been seeded with basic common vocabulary, the station's presenter names, local place names, and so on.
- ▶ An enhanced option in which human captioners prime the system in advance of a specific programme or event. They provide specific vocabulary, names of relevant people, and words which might be more likely to be misrecognised.

The latter service reintroduces the need to pre-plan and schedule humans, but delivers better results with fewer errors.

Because of this, the human-assisted service is used for pre-scheduled live events. Jim Kercher, Master Control Operator at PBS39, explained that this process delivers outcomes that are generally as good as their previous fully manual captioning, at a much lower cost.

When building their solution, Red Bee Media set out to be agnostic of the speech recognition engine, allowing it to use the best engine for a given job. This required the team to investigate a number of leading speech engines, to identify those that are best suited to captioning. They found that the best results come from engines with access to huge corpuses of data, rather than bespoke engines trained on genre specific information; but that results are dramatically enhanced through the addition of experienced human input to add specific and relevant vocabulary.

PHASE 3: CONSUME

■ THE POTENTIAL

Content consumption experiences have – in the VOD space at least – been some of the earliest beneficiaries of machine learning, as recommendation algorithms are central to most services.

But our 2018 group of experts saw potential for further personalisation based on factors such as consumption context (device, location, etc), and for the promotion of content through trailers and advertising.

Monetisation is of course an important area, where potential was mainly seen around optimising advertising, although some related capabilities were identified including dynamic product placement and personalised pricing.

Meanwhile, there is potential for enhanced automation to deliver innovation, including dynamic versioning and content repurposing, device or environment specific audio enhancement, on-device upscaling, voice search, and sentiment analysis of viewers' responses to content.

Commissioning was placed into the consumption phase, because viewer analytics and consumption data can be used to inform the commissioning of new content. It was not ranked highly in terms of the areas with most potential for automation, although some excitement was felt in 2018 around the possibility for AI to predict the success of a piece of content based on its script and other information.

Notably, this summer's group of experts largely agreed with the possibilities outlined in 2018; there was little in the way of new applications suggested. There was, however, some progress towards those already outlined.

■ MATURITY: THEN AND NOW

There are well documented use cases for personalisation of the consumption experience, including artwork displayed in the user interface, pioneered [by Netflix](#).

Monetisation is, perhaps unsurprisingly, an area where continued investment has delivered progress. Dynamic advertising insertion into VOD was already common in 2018, and the equivalent in linear broadcasts was becoming a reality; one that has continued to improve in the years since. There have been high profile rollouts of set-top-box technology that can replace advertisements on broadcast channels with personalised selections for the specific viewer. Meanwhile, we have seen the rapid rise of FAST platforms (Free, Advertising-supported Streaming Television), which allow for dynamic ad insertion into linear channels. The latter is discussed in more detail in [Cloud Payout](#) and [Streaming at Scale](#).

In the content itself, virtual ad insertion in sports environments has been commonplace for a number of years. And as content is distributed around the internet in more formats and on more platforms than ever, some companies (including the media teams at sports leagues) reported using logo recognition technology to identify the exposure given to their sponsors, in order to maximise sponsorship revenue.



Logo recognition is used to identify the exposure given to sponsors, helping to maximise sponsorship revenue

However, one area that doesn't appear to have grown significantly is dynamic or personalised pricing; perhaps more for reasons of social and consumer acceptance than reasons of technology.

Some additional personalisation of streaming interfaces and user interactions has been noted. Considerable work has been undertaken in recent years on the use of object based media, but it has largely been confined to research labs, rather than reaching mainstream adoption.

Another area of nascent development is in using consumption and consumer behaviour data to inform new content. Some script analysis technology exists, but its deployment remains niche. Other niche deployments include the BBC trialling machine learning to identify music that might be successful in their new artist showcase series across multiple radio stations.

Automating the Media Factory

KEY INSIGHTS

- ▶ The factory processes of media receipt, processing, versioning, and distribution, are the logical epicentre of automation efforts.
- ▶ Discovery has largely automated the process of creating textless masters from incoming texted content, using machine learning models.
- ▶ WarnerMedia has applied automation to its supply chain, including developing AI solutions for audio conform.
- ▶ Training machine learning models takes time, expertise, and data; it requires collaboration between technologists and business experts.

Examination in 2018 revealed a great deal of potential for automation across the content value chain, especially in areas such as versioning and rights management. And although a few additional use cases have been conceived since, the broad landscape looks similar. Media companies and their technology partners have been working to execute against some of these untapped potential areas.

While there is of course a lot of work taking place across the industry, in many different areas, one particular focus area stands out. The most factory-like elements of our industry are those involved in content receipt, processing, versioning, preparation and distribution. And so this is where the most automation is happening.

There is particular focus on automating the factory-like elements of media: receipt, processing, versioning, and preparation

CASE STUDY



DISCOVERING VERSIONING SOLUTIONS

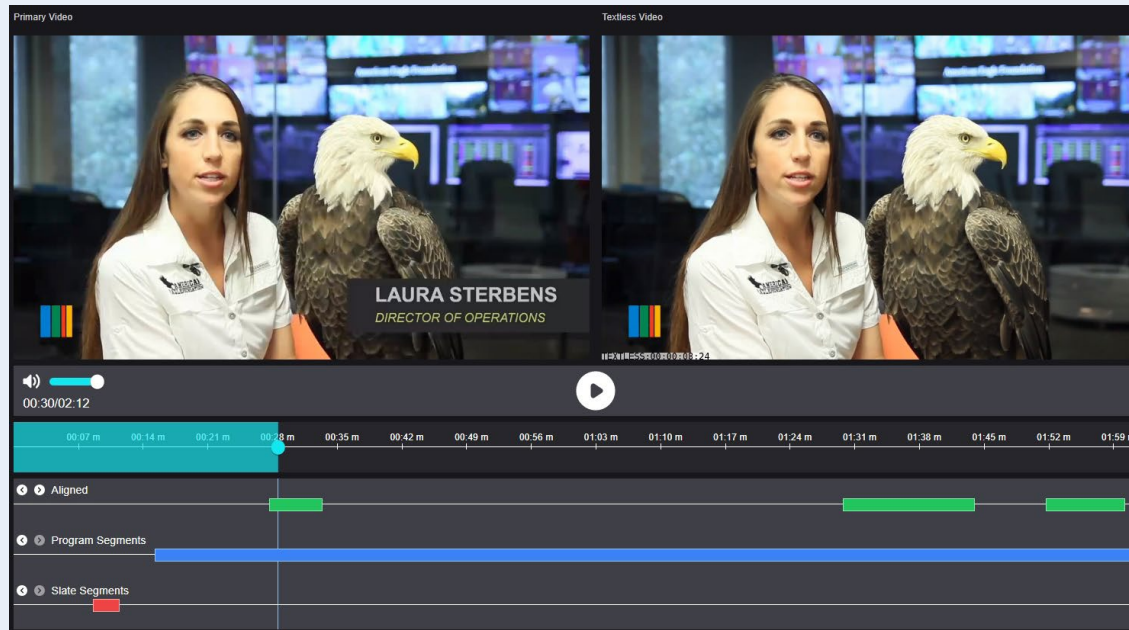
International distribution of content is a key part of Discovery's business. Discovery distributes the content it creates and acquires across over 200 countries and in 52 languages.

Localisation and language customisation are therefore a substantial part of Discovery's content supply chain. It is often necessary to replace audio tracks with new language versions, or add subtitles; but viewers in many markets also expect on-screen text and graphics (such as lower thirds and open captions) to be localised too.

Graphical language customisation requires textless elements, and Discovery's content producers and suppliers provide 'textless clean covers shots' for any texted material. But matching them to the correct place in the master content, and identifying missing elements, is traditionally a highly time consuming manual process. To make matters worse, mismatches would often be found between the durations of the supplied textless elements and the text in the programme.

AUTOMATED ANALYSIS

Piksel was able to develop a machine-learning based solution to this problem. It is provided via a SaaS model in the cloud, and accessed through an API.



The Texted Video Replacement service analyses programmes to find on-screen text within a program master. It then matches these elements to textless equivalents (supplied either as separate media files, or at the end of the main programme) and creates an EDL for the required edits. If the service can't find textless replacements for the texted shots, it escalates the problem to a human reviewer.

The service automatically analyses the programme to find on-screen text to be removed or localised

The analysis begins with a process of shot detection, then identification of which shots contain texted elements. The system is able to identify organic text elements – such as text on signs, computer screens, clothing, and so on – as these need not be edited; only overlaid text is replaced.

The results can be viewed in a web user interface, and where necessary, the EDL can be exported to a non-linear editing system. Reports of missing textless elements are sent to the content supplier to be remedied, generally without requiring human review and intervention.

CLOSE PARTNERSHIP

The story of how the system was developed is as interesting as the results. Pikel worked closely with Discovery's R&D team, collaborating not just on the product requirements, but also the solution itself.

It began with a text recognition and analysis tool Pikel had developed for a different use case. The Discovery team saw potential to solve their business problem by employing the technology slightly differently. Pikel was able to adapt the solution, training the machine learning models with both synthetic (i.e. purpose-generated) test material and real programme content from Discovery. The Discovery team also contributed some of their own work, such as using the orientation of text on screen to help identify whether it was organic or overlaid.

Throughout the process the Discovery and Pikel teams used a shared Slack channel to keep communication flowing during development and evaluation. The commercial incentives were aligned with a high degree of mutual trust: Discovery gave Pikel a conditional Purchase Order, committing to a minimum spend with the service provided the developed solution could successfully meet the requirements.

CONFIDENCE AND ROI

To get to the point where confidence in the system was high enough to use it in production, around three months were spent with operational teams testing and refining the training of the ML models.

The service was then deployed into active use in early 2021, and since then Discovery has run over 14,000 analysis jobs across more than 7,000 hours of content.

The automated workflow saves considerable time in human inspection of the incoming content, and in post production time due to the automatic EDL generation. Cost savings are expected, but the primary benefit is in time to market.

Cost savings are expected, but the primary benefit is in time to market

Global direct to consumer content platforms have created an expectation for simultaneous release of content around the world. Revenue can therefore be maximised by reducing the time from receiving a primary version to releasing the content globally.

MANAGING VERSION OVERLOAD

Continuing their collaboration, Pikel and Discovery have now moved forward to work on additional business use cases.

Discovery's versioning process was traditionally handled as a function of local distribution. But as it moved to launch discovery+ as a global streaming service, it was faced with having tens or hundreds of versions of many pieces of content. These could be from different countries for example, or different brands (Science Channel, Discovery Channel, and so on).

Identifying the difference between these versions is complex and time consuming. It cannot be automated by simple comparison of media files, because even editorially identical files may have been transcoded, and so will not be mathematically identical. Therefore, a perceptual model is required to identify content that is similar or the same.

The problem space is highly nuanced. Discovery needs to understand not just whether two versions differ, but which shots within the programme differ. The business is not interested in the difference in encoding artefacts between MPEG2 and MPEG4 renditions, but it is interested in different languages, channel branding, or other edits. Therefore, extensive training of the machine learning system was required.

The hope was that the resultant saving in storage cost would pay for the solution, while also saving a great deal of manual effort and complexity. Discovery is in the final phases of testing, validating the version analysis systems across thousands of assets in a final large-scale test, and expects to fully deploy the service in Q4 2021.

UNIFYING THE SUPPLY CHAIN

For WarnerMedia, automation of the content supply chain is a long journey which was started many years ago. Under the banner of Digital End to End (DETE), a largely automated supply chain was built to deliver content to broadcast channels, VOD outlets, and license holders.

WarnerMedia's partner Vubiquity acquired the service in 2015, and continued to develop it as part of their own offering. It now supports 4K, and is multi-tenanted, serving multiple customers. Vubiquity's service now delivers WarnerMedia content to over 1,000 recipients.

The system has the ability to store and interrelate video, audio and texted assets in complex combinations. Humans are involved in the creation of new components, but the library, with its Interoperable Master Format (IMF) capability, allows for all of the distribution to be fully automated.

But despite the success of the system, the launch of HBO Max created an unprecedented set of demands. Once again, time to market was the driving factor, as the business demanded the platform be readied for launch quickly. Vubiquity had to deliver tens of thousands of assets to enable the launch.

The launch of HBO Max created an unprecedented set of demands, as the business demanded the platform be readied for launch very quickly

And for a service with global ambitions, the initial launch was only the beginning. Following the domestic rollout of HBO Max, the next set of launches was in Latin America, taking the platform from one country to forty overnight.

The launch in Latin America took the platform from serving one country to forty, overnight

CONFORMING CHALLENGES

The Latin American launch of HBO Max surfaced some specific challenges to which ML solutions were applied.

When compliance edits or other changes are made to a piece of content, international language tracks can fall out of sync or suffer other problems. For example, some sentences are longer in one language than another, and so if the audio tracks were simply cut at the same timecodes as the video, speech would be truncated.

For archives of content, these problems can compound over time. In WarnerMedia's case, a high volume of children's content was found to have been edited at various times for different distribution deals, but with a lack of consistency in whether and how audio tracks were edited.



When large amounts of content were required in multiple languages for the launch, finding and fixing all of these problems manually would have been prohibitive

When large amounts of content were required in multiple languages for the LatAm launch, finding and fixing all of these problems manually would have been prohibitive from both a cost and time perspective. So Stewart Curtis and his colleagues worked with WIREWAX Media Services on a solution to locate audio conformance problems, and create EDLs to implement the required fixes.

Around 40% of the issues could be fixed automatically, estimates Curtis, while the time taken for a human editor to implement the remaining fixes was reduced by an average of around 60%.

The project was no small undertaking. Training and iterating the solution took months of collaborative work, with WarnerMedia providing large sets of sample content to the WIREWAX team.

As is often the case with technology projects, the main challenge was change management. There was a skepticism in some quarters that the solution would work well enough to be used, and so there was a resistance to invest resources into change management early on. But in the end, the system performed well enough to prove its value.

Learnings About Machine Learning

KEY INSIGHTS

- ▶ AI is vital when speed is of the essence; whether producing fast turnaround content, or processing content libraries to deliver new commercial opportunities.
- ▶ AI development requires technical expertise and business understanding; it is therefore most often achieved through close collaboration between a vendor and customer.
- ▶ Ethical concerns around the development, training, and use of AI have not subsided, but there is greater understanding of the need to address them head on.
- ▶ The value of ML models comes from both the technology and the training; there are now good precedents for commercial agreements that recognise both.
- ▶ Automation works best when applied to specific problem spaces; we have moved beyond generic solutions and begun to deliver true business value.

AI for Real ended with the following statement:

To some degree the challenge of understanding how AI in media might evolve feels overwhelming. But there is a self-awareness, and a remarkably common language, among both suppliers and customers about technology innovation in the media sector today.

The potential offered by AI is both scary and exciting. But the preparedness across the supply chain for an open dialogue about how to make the scary less so, and how to make the exciting practical, has perhaps never been greater.

That common language and shared understanding seem to have been borne out in the discussions and case studies that have informed this report. But how have we fared on the journey towards realising that exciting potential?

AI for Real offered five key takeaways about the state of AI in the media industry. So three years on, how have things changed? To find out, we will examine each of the five statements from 2018.

1

AI LOVES FAST TURNAROUND

■ THE 2018 VIEW:

In 2018, the DPP identified that the fast paced worlds of news and sports made them ideal content types for the use of AI solutions. Sport is especially well suited, due to its controlled, predictable, rule-bound, and data rich environments.

Long form non-live content was felt to be less well suited, with long production cycles and project based delivery. We said:

When it comes to AI innovation in long form production, it is likely such innovation will be applied far more to the finished product than to the production processes themselves.

■ THE 2021 VIEW:

All of these statements remain true. However, from the vantage point of 2021, a new trend emerges: AI loves fast time to market.

From the vantage point of 2021, a new trend emerges: AI loves fast time to market

Throughout many of the case studies in this report, AI has been used to deliver faster. Yes, that includes live and fast turnaround content getting to air. But AI also has a new heartland in processing large amounts of content to fulfil distribution deals, populate linear channels, and launch new D2C products.

2

WE CAN'T SEE FOR THE BUBBLES

■ THE 2018 VIEW:

One of the biggest barriers to progress of media industry AI in 2018 was the siloed nature of its development.

Production studios, traditional broadcasters, and streaming platforms each came to the problem space from very different perspectives and starting points.

A huge mismatch of skills and knowledge existed between the companies with technology and data science expertise, and those who understood the content and business challenges.

And very few companies could quantify the potential return on investment from developing automation solutions.

■ THE 2021 VIEW:

It is still, of course, difficult to know in advance whether a solution will meet the required effectiveness threshold to deliver business benefit. A speculative investment is still often required – whether in monetary terms or in time.

And despite many media companies investing in data and ML skills in-house, it is still of course true that technology companies possess the strongest expertise. Media companies remain best equipped to define their business problems.

But by considering these challenges together, and being willing to truly partner with others, some organisations have found solutions. Once again our case studies shed some light on potential paths forward. Customer and vendor teams have worked in close collaboration, accepting and celebrating each other's strengths. And crucially, it's been done within a commercial framework of shared risk and reward.

Customers and vendors are collaborating within frameworks of shared risk and reward

3

THE LACK OF AN ETHICAL FRAMEWORK

THE 2018 VIEW:

Another inhibitor that was much discussed was the ethics of AI.

In many ways, these concerns have been brought into sharper focus in the intervening years. High profile problems with biased algorithms such as face recognition have shown that machine learning is as biased as the data on which it is trained.

THE 2021 VIEW:

The discussion among our experts in 2021 recognised that these issues require our collective effort and attention. However, it did not suggest that we should shy away from AI as a result.

Ethical concerns have been thrown into sharper focus, but this does not mean we should shy away from AI altogether

User privacy remains crucial when processing consumers' data (such as for content recommendations). But this is an area where regulation is relatively mature.

A great deal of work has been done across the media industry and the wider technology community on issues such as biased algorithms, AI-generated fake content ('deep fakes'), and similar issues. Much more work must still be done.

Companies' diversity and inclusion protocols should now encompass consideration of the use of automated systems that make any assessment or analysis of factors ranging from identity (race, sex, age, and more), to political viewpoints. All those involved in training ML systems must remain vigilant about bias in their training data.

THE 2018 VIEW:

In 2018, there was considerable anxiety around the intellectual property status of AI models created using training data sets.

Many content companies felt that they were owed some ownership of ML models trained on their data. Technology companies, meanwhile, generally expected ownership of software that they'd created, no matter how it was trained.

Ultimately, of course, these points of view risk creating a stalemate that halts development of useful solutions.

There was a risk of stalemate that would halt the development of useful AI for our industry

THE 2021 VIEW:

It is encouraging to now see so many examples of positive cooperation between vendors and customers. It's increasingly common for media companies to allow their content to be used for training, in return for some commercially advantageous access to the resulting product.

There is a recognition, perhaps, that although the training data provides value to the ML solution, no value is taken away from the content.

Vendors, meanwhile, are more willing to undertake development with a customer, if the customer commits to using the successful solution. If both parties have skin in the game, there's a shared desire to succeed.



As a rightsholder, there is tremendous advantage to us to share data and have others help us solve the problem.

DAVID KLEE, A+E NETWORKS

Such attitudes depend on the commercial perspective, of course. For service providers, many of the AI solutions being developed give them competitive advantages, so if they train algorithms they are much more likely to want to keep the results to themselves.

Another model appears to have risen in prominence in recent years, although its existence predates [AI for Real](#). Major cloud and technology providers offer pre-trained base models, which customers can then extend using their own training data. This works well for relatively standard use cases, such as face recognition. And it provides a happy medium in which the extended model's IP remains with the content owner, yet they do not have to develop a model from scratch.

5

PROBLEMS AND SOLUTIONS IN SEARCH OF EACH OTHER

■ THE 2018 VIEW:

Perhaps the overriding problem facing AI in the media industry in 2018 was the disillusionment that follows overblown hype.

There were many AI solutions available, but they'd been sold as if media companies could simply push their content through an API and get valuable data back, with no further effort. In reality, most of these solutions were highly generic, designed to address the needs of many use cases and industries, and failed to deliver sufficient value.

■ THE 2021 VIEW:

The hype seems to have died away, and the conversation moved to a much more practical one, focussed on solving actual business problems. More industry specific models and products have appeared, while customers have begun to understand the self-training options discussed previously.

Expectations have shifted from the wildly unrealistic towards a more nuanced view of the need to balance AI and human intervention. This could be using AI as the 'supervisor', checking the work of humans; or more commonly AI as the 'junior', performing the lower level work and escalating to a human where required.

**The hype and disillusionment of 2018
have died away, and efforts have shifted
to solving real business problems
to deliver true value**

Most of all, we can now see a great deal of collaboration between customers and vendors to identify and execute against specific business opportunities. This has been seen in the adaptation of Codemill's products to evolve with UKTV, the development of services by Piksel to solve Discovery's textless challenge, those developed by WIREWAX to service WarnerMedia's needs, and more throughout this report.

These collaborations often benefit both parties: the vendor creates a product that can be sold to others, and the customer delivers value to their business.

There is no universal formula for delivering perfect automation, but defining a specific business challenge or opportunity is certainly a good place to start.

CONCLUSION

Delivering Value Faster

In 2018, we described the potential of AI in the media industry as both “scary and exciting”. In 2021, the excitement remains undimmed. But the fear is subsiding, as we look beyond generalised solutions and focus instead on manageable business problems to which AI can be practically applied.

We also said there was a “preparedness for an open dialogue” between media and technology companies about how to deliver real business benefit using AI. That dialogue is now in full flow, with collaborative partnership at the heart of many successful solutions.

The quest for automation has been greatly aided by the move to the cloud

The quest for automation has been greatly aided by the move to the cloud. As media companies centralise their content and data, the power of automation at scale becomes evident. And with the resources of the cloud at hand, new possibilities become tangible reality.

From mining archives to managing supply chains; from creating content to launching new consumer services; automation is now at the heart of modern media companies. They achieve success by focusing investment where it can deliver the greatest return.

In some cases that means driving efficiencies in existing processes. But for many, it means executing faster to realise new business opportunities. The true value of automation is not just saving costs, it’s enabling revenue.

The true value of automation is not just reducing costs, it’s enabling revenue

The Cloud for Media series was researched and authored by **Rowan de Pomerai**.

The project was managed by **Joseph Adedipe**, and the workshops were organised by **Abdul Hakim** and **Anh Mao**. The report was designed by **Vlad Cohen**.

About the DPP

The DPP is the media industry's business network. It is a not-for-profit company with an international membership that spans the whole media supply chain, covering global technology companies, production companies, digital agencies, suppliers, service providers, post production facilities, online platforms, broadcasters, distributors and not-for-profit organisations. The DPP harnesses the collective intelligence of its membership to generate insight, enable change and create market opportunities. For more information, or to enquire about membership visit

thedpp.com

About Vubiquity

VUBIQUITY, part of the Amdocs Media division of Amdocs (NASDAQ: DOX), delivers premium content to viewers on any screen, device or platform. With a wide-ranging global customer network, the world's leading content owners (Major Studios, Indie Studios, Broadcasters, TV Networks & Digital First Networks) and video service providers (OTT, Mobile, Cable, Satellite, Telco) have selected VUBIQUITY to deliver entertainment experiences around the world. VUBIQUITY has offices in Los Angeles, Toronto, London and Cyprus.

www.vubiquity.com

Photography credits:

Page 27

Courtesy of PBS

This publication is copyright © DPP Ltd 2021. All rights reserved.