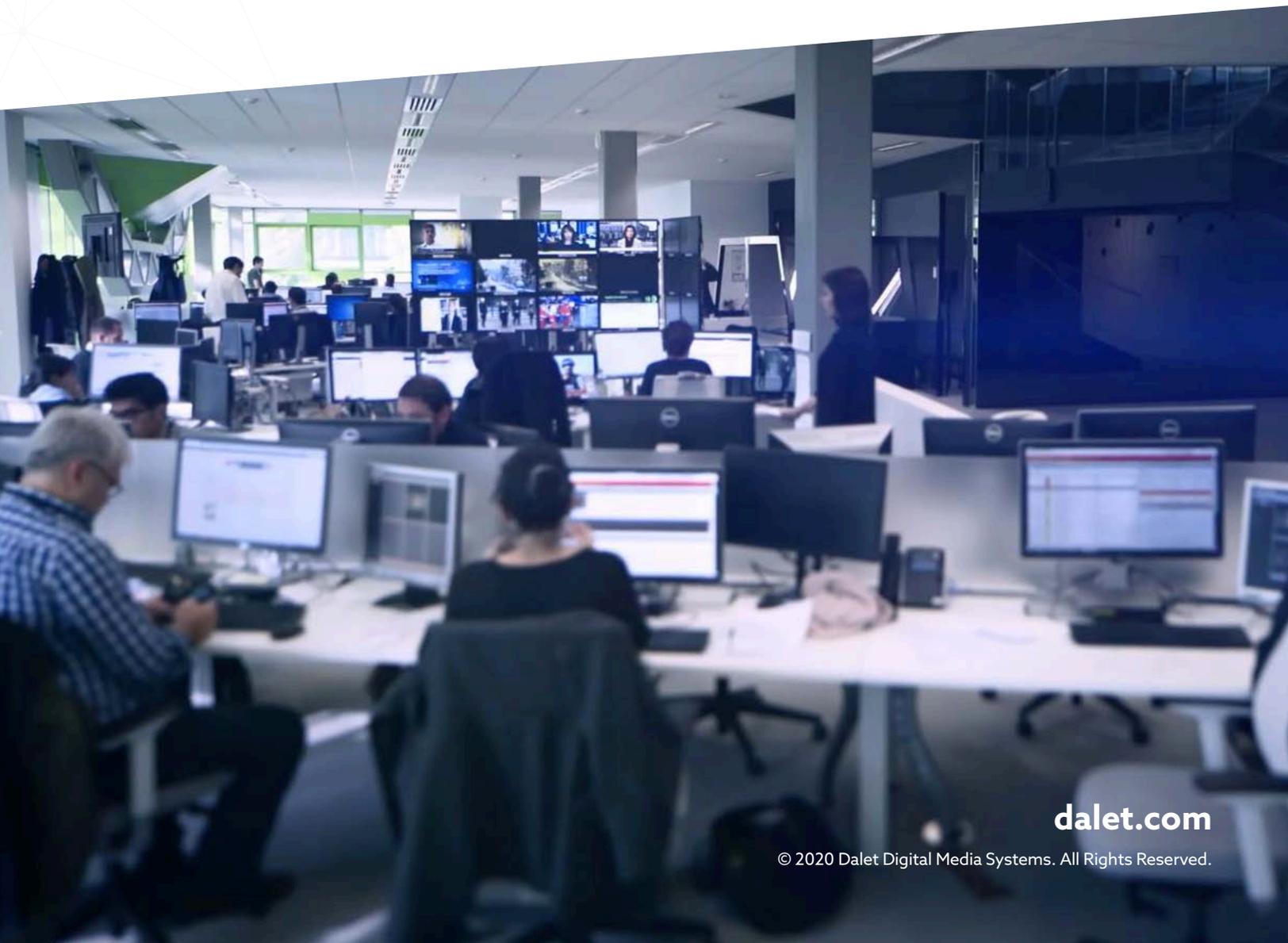DALET

A Brave New Approach to Planning News

An Extreme Case Study for Cloud and AI Enablement

Raoul Cospen
Director of News Strategy
Dalet

dalet.com

Abstract – Fierce competition for eyeballs along with a changing multiplatform world has put newsrooms under immense pressure. With a national election always happening somewhere in the world, the pressure is amplified to deliver even more comprehensive stories and visuals.

Coupled with the logistics of news collection through a multitude of social sources, distribution to numerous viewing platforms, and the public's desire to have the very latest information on the story they are following, many broadcasters must reinvent their workflow to thrive.

Broadcasters need to carefully consider the technology choices that will help their journalists to deliver better content, faster and to more platforms.

With the recent introduction of cloud computing, Artificial Intelligence and web technologies, there is a new generation of tools - if implemented properly - that enable journalists to efficiently access a much larger pool of news sources with capabilities to scale distribution up and out. However, news technology suppliers must adapt their solutions to support these emerging technologies and leverage their vast experience to provide sound guidance for making the leap forward.

Part one of this paper summarizes the new technologies available to newsrooms. It addresses how these new capabilities can be implemented to increase quality, productivity and distribution potential on multiple platforms. It also considers why leveraging emerging technologies is critical to supporting news workflows during major election events.

In addition to these tangible objectives, there is a much more difficult challenge that needs to be quickly addressed: the growing amount of Deepfakes and Fake News spreading damaging information across the Internet. With over two-thirds of the American adult population (68%) [1] getting their news from social media on a regular basis, the concern for Deepfakes and Fake News driving democracy-threatening narratives worldwide is greater than ever. Therefore, identifying Deepfakes and combating the spread of Fake News is at the core of every newsroom strategy.

Part two of this paper will address how emerging technologies and initiatives can help check the veracity and authenticity of media sources and ultimately help expose Deepfakes and eliminate them as news sources and prevent the proliferation of Fake News.

KEY HIGHLIGHTS

- A new generation of tools to increase quality, productivity and distribution potential on multiple platforms

- Why leveraging emerging technologies is critical to supporting news workflows during major events such as elections

- How recent technology initiatives can help check the veracity and authenticity of media sources and ultimately help expose Deepfakes and eliminate them

Next Generation Tools to Foster Productivity, Quality and Better Reach Your Audiences

Election time means extremely busy newsrooms. News planning is at full capacity with larger than normal amounts of content arriving in the newsroom. Everyone is fully focused on being the first to break news events. While the pace is frenetic, quality and well documented stories are vital to remain viable.

There are multiple phases to the creation of good stories:

- A news coverage plan that captures all of the important events (press conferences, candidates on the campaign trail etc.)

- A newsgathering process that follows trustworthy news sources

- Tools to help journalists efficiently and collaboratively produce different versions of the story – usually starting with digital

- A platform to distribute and monetize the story, across multiple platforms

Newsrooms that operate in «Breaking News» mode rarely work with a sequential rundown. Instead they work with a curated pool of media AKA "image carpentry." Anchors, who are on-air continuously, pick media from the curated pool as needed. The studio control plays - by cue - the clips. Often over time, certain clips are played repeatedly.

However, the business requirements have changed and digital first workflows are now becoming the standard practices. We ask the same journalists to look at more sources, to produce more and better content faster, and distribute to many more platforms. This is an equation that technology needs to solve.

STARTING WITH NEWSGATHERING

Getting out a quality story starts with the consistent monitoring of many sources of media. Here is a non-exhaustive list of media sources:

- Wires / agencies
- User generated content
- Social media sources, primarily Facebook, Instagram, Twitter, YouTube
- Archived content
- Syndicated content
- Content coming from live feeds, backpack or simply from cameras
- Media content uploaded from freelancers
- Contacts / people
- Web searches
- Radios, papers, TV news shows
- Press releases
- Emails

This is a lot of news sources and it is physically impossible for a human being to go through them all on a given topic. An emerging notion of "News Explorer" allows monitoring of disparate sources like the one listed above into a single (but potentially advanced) search query. That News Explorer concept can be specifically made for a given story - and the latest updates fall directly in that story with a notification to the relevant journalists.

SECOND NOTION IS THE "COMPLEX STORY"

One key and emerging concept of the new generation news production platform is this notion of "Complex Story." In the past, a story used to be a TV script, with a link to MOS objects such as graphics and videos and a place in a rundown. Thankfully, this era now belongs to the past. Because broadcasters need to distribute news to the web, on social media platforms, over OTT services and play it out on television, the story has many components. Of course, the story may include a TV script, some media, some graphic objects. But it may also include a CMS script, a Facebook summary, or even some actionables such as tags or to-do lists. Also attached to the story are all the components one needs to elaborate on a story: relevant wires, media from archives, media that is not yet there but soon to be ingested, links to other stories, contacts, etc. All this media is structured into a MediaBin - available with the story.

THIRD NOTION, THE "MEDIABIN"

The notion of MediaBin is central for the news story telling process. MediaBin is a structured way to organize text content, graphics, emails, research material, finished projects, rendered media, etc. It is a logical pivot between the text editing world and the media editing world. It unites both worlds, providing a unified view of all potential story assets in a space ideal for story collaboration and evolution. Everyone can contribute in their own way. The field reporter will deposit pictures in the MediaBin while a video editor will edit incoming streams, with digital editors reusing the story narration to prepare the video content for social media.

These three new concepts, News Explorer, Complex Story and MediaBin will drastically evolve best practices in the newsroom. Leveraging these three concepts, digital curators, journalists, video editors, digital editors, and others can achieve the level of collaboration that is needed in a newsroom. These concepts also cover digital first workflows – which is a real benefit during election time where most of the news hits social media first, then newscasts.

Figure 1: The MediaBin is a collaboration tool that organizes and links all kinds of content with a story.
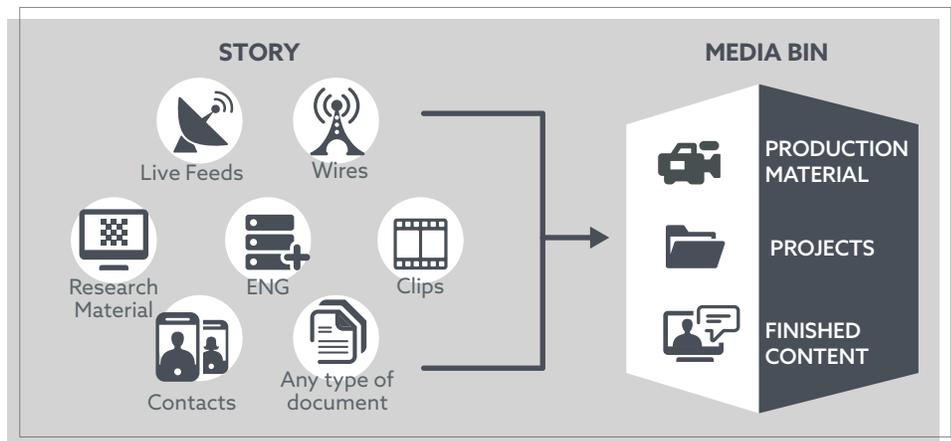
FOURTH NOTION: DISCOVERY SEARCH

Complementary to the News Explorer, "Discovery Search" is an automatic way of searching for content based on the storyline a journalist is developing. For example, let's say a journalist is writing a story about potential corruption in the upcoming Olympics. The Discovery Search will surface stories, videos, pictures, or wires that are already in the production system or in the archives. The more detailed the story is, the more precise and connected the recommendations are. Discovery Search uses standard Natural Language Processing and unsupervised training to analyze and understand the content of story and recommendation algorithms to bring relevant content. There are two significant benefits that apply particularly well during big elections:

• The story is richer with more documentation to elaborate narratives

• Save precious time searching good and relevant content

Complementary to Discovery Search, there is also an interesting productivity approach to ingested content based on AI and in particular speech-to-text. Speech-to-text now gives incredible results – in many languages, with many local accents and dialects. It is not rare that the accuracy reaches over 99%. Coupled with a **transcript viewer**, speech-to-text makes shot selection of raw material easier...and much faster! Journalists can read through the transcript and select the text parts they need, and the underlying technology will **seamlessly build the video EDL**! The productivity gain is significant and depending on the nature of the raw material, current metrics indicate that these AI enabled editorial workflows are 10 times faster than watching the traditional route which entails watching the entire video.

It has a positive impact on productivity as well as the time it takes to distribute the story to air and digital platforms.

There is another AI capability use case worth noting with regards to elections. In many countries, measuring and reporting the speaking and airtime given to candidates is a regulatory requirement. Face recognition technology and voice printing technology can automatically generate reports with accurate candidate data. This information can be displayed in show / daily / weekly / monthly reports but can as well as in the rundown enabling news teams to make better decisions in near real-time.
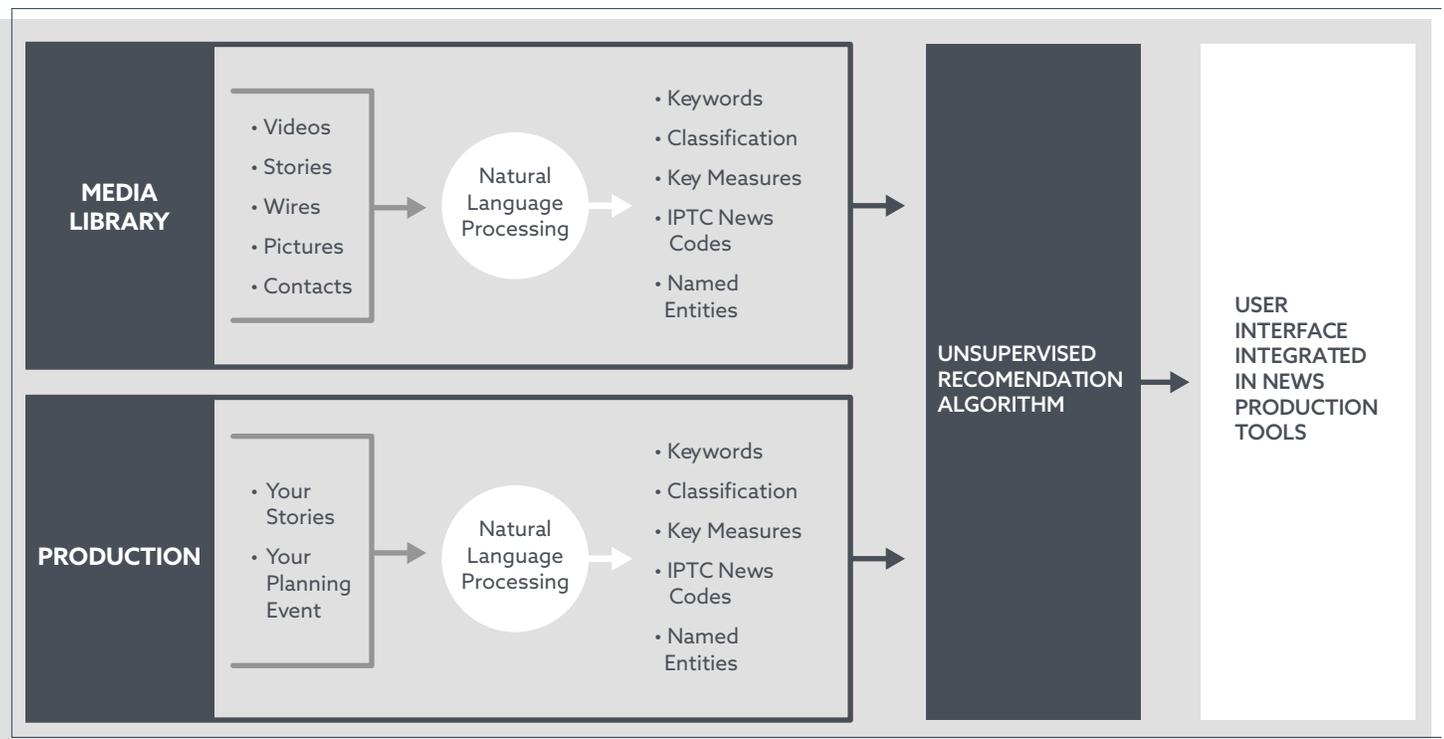

Figure 2: Unsupervised ML recommendation algorithms can bring relevant content to the users based on their work in real time.

Emerging Technologies That Prevent Fakes and Deep-Fakes

The Fake News era erupted in 2016 with the Cambridge Analytica scandal and Russian bots [2]. With companies specializing in mass-spreading false narratives, Fake News is a real threat for the credibility of media outlets and for democracies overall.

Deepfakes, which is a video, graphic or audio piece that makes a person appear to have said or done something they haven't, can have an equally negative impact on society, as Fake News. Deepfake algorithms have become very effective and thus we are seeing a much wider spread of Deepfake audio, video and imagery. It is really incredible just how easy it has become to create credible Deepfakes using the free services found on the internet.

- The following YouTube video demonstrates the real potential to create a destructive Deepfake news piece.
- [Pelosi videos manipulated to make her appear drunk are being shared on social media](#)

Similar to Quality Control engines, there are engines based on pattern recognition methods that can detect synthesized and modified media. However, these are threatened by Deepfake generator technology that has been perfected to the
point that soon neither human nor algorithms will be able to detect the manipulated content.

So how can technology potentially help news organizations mitigate Deepfake sources and deliver fully verified news stories? Further, how can news organizations convey to partners and audiences that the news stories viewed on air or online and picked up for syndication are certified and can be trusted?

One concept to help combat fake information is the implementation of a public authentication service. The authentication technology would stamp verified media source(s) with a certificate of authenticity. This stamp would stay connected to the media source as it passed through the various editorial steps - ingest, production and distribution – up to the end users.

In this paper, we will refer to the authentication service technology as FACT -> "Framework on establishing Authenticated Content Trust."

FACT HAS A DOUBLE BENEFIT FOR THE NEWSROOMS:

- News operations can utilize the authentication certificate workflow to endorse the use of media sources across the newsroom, eliminating the need to verify media sources each time they are used.

- The certification workflow also applies to syndicated content received from partners and agencies. Ingested content with an authentication certificate signifies that the media has been cleared and can be trusted as a verified source.

One important element of the FACT framework is AMP. The result of a 2019 initiative from Microsoft, AMP [3] or Authentication of Media via Provenance is a way to secure media based on the joint use of a strong authentication. It is a solid approach to this fundamental issue.
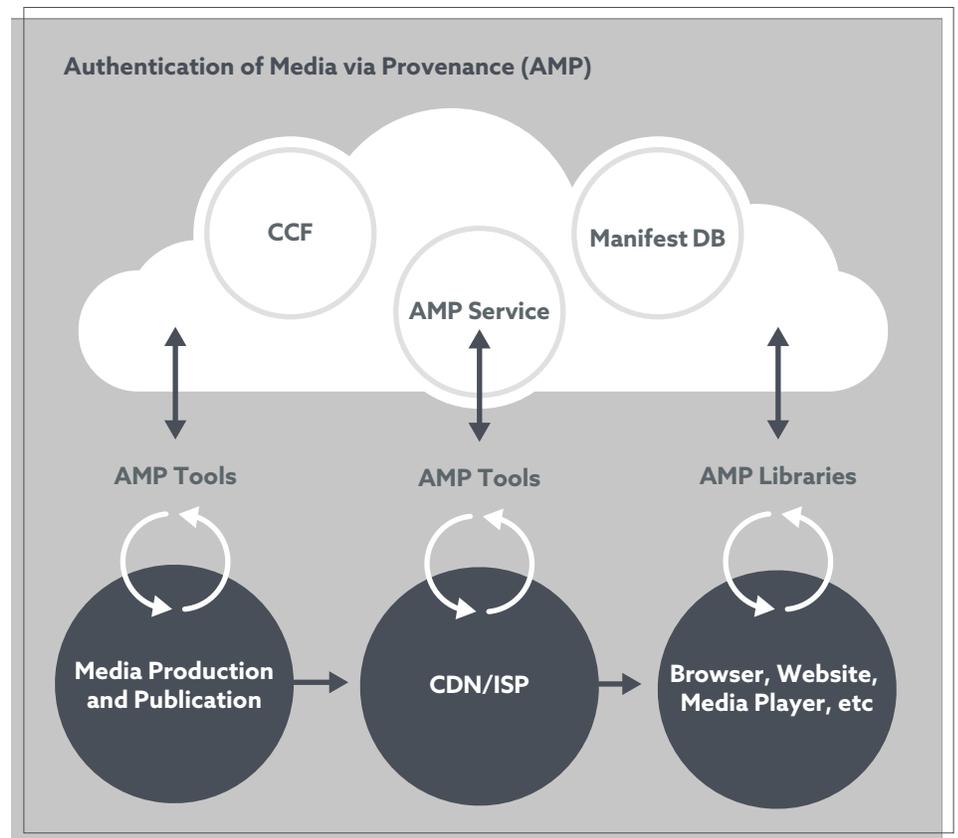

Figure 3: Integration of AMP tools and services into a media production and distribution pipeline.

AMP describes its technical approach as follows: "Approaches to securing media from a validated source to its consumption include (1) strong authentication and (2) fragile watermarking. A complementary approach involves (3) the detection of manipulation or synthesis via pattern recognition employing technical methods, such as machine learned classifiers. Finally, there are opportunities to explore (4) event-certification methods for certifying that media as captured is linked to actual physical events, rooted in activities that are certified via a combination of methods to a time and place. The approach employed in the AMP system to securing media is based on the joint use of (1) and (2), coupled with the certification of the identity and trust of the source of the media."

AMP uses CCF (Confidential Consortium Framework), and a public blockchain to store the AMP Manifests that describe the provenance of the media (media source, credits, link to other media). An AMP Manifest is a data structure that cryptographically authenticates media objects and their associated metadata. AMP Manifests can be used to authenticate the original source material or authenticate transformations from one format to another.

In other terms, the FACT technologies utilize a fragile watermarking technique to link the media (pictures, or frames of a video) to the authentication manifest. When the original media is altered (during a transcode or while burning a logo for instance), the watermarking is equally altered. Note that checking whether a transformation is faithful is not discussed here. However, it is up to the originator of the transformation to then endorse it in an updated manifest that will be added to the blockchain.

There are multiple FACT initiatives similar to AMP. Also created in 2019, "The News Provenance Project" is a partnership between *The New York Times* and IBM. Adobe also in 2019 introduced the "Content Authenticity Initiative."

How FACT can be Implemented in Your Newsroom?

The news production system will need to connect to a FACT framework at different stages throughout the workflow: during ingest to confirm the veracity of the media, during production to certify the endorsement of the media by the news organization and during distribution to register and authenticate the produced media into the AMP service.

1. VERIFYING THE SOURCE OF INGESTED MEDIA

There are several distinct use cases. The most straight-forward applies to content that is owned, say from archive or news gathering source such as a field camera. The news production system will need to authenticate the owned content using the AMP service. This is a simple task: when the operator enters the usual metadata, the news production system will create the manifest with the relevant metadata, including the organization's signature as origin. An invisible, "fragile" watermark will be generated by a back-office process, linking the source media with the manifest, automatically registering it into the FACT public ledger.

A second use case revolves around ingesting syndicated content. The news production system connects to FACT to verify the authentication metadata, which will then be imported into the news production system accordingly. The metadata will include the name of the organization that originally authenticated the content.

When it comes to social media or UGC (user generated content), the situation is different. Most of the time, the media will not be certified. It is then up to the journalist to editorially verify this content (possibly with the help of a fact checking technology, as described later in this paper). The journalist, on behalf of his/her news organization, will need to certify the origin and the veracity of this content and register its authenticity.

2. CERTIFYING PRODUCED CONTENT

During the video editing portion of the news production workflow, the EDL and associated media are translated into a manifest, indicating media sources that are registered by FACT. An important point to note, this process can only be done if the video editing application used to assemble the news story is able track media genealogy.

The video production tools will create new certifications, adding up manifests to the FACT block-chain. To ensure proper tracking of media, video production tools must understand media genealogy hierarchy. This links media sources authenticated previously from newly created manifests to their existing manifests - keeping intact modifications made to the media and source verification details. In essence, FACT holds the authentication of produced media and its relationship with source media.

By certifying the finished content, it means that the journalist has endorsed modifications made to the original media. These modifications could be the addition of a lower third, logo, speed of a sequence, blurring of a character, or simply a format transcoding. This will alter the original fragile watermarking - creating an indicator that modifications have been made. Authenticating the finished piece is indeed mandatory.

3. DISTRIBUTING CERTIFIED CONTENT

The finished material is now authenticated and ready for distribution. When this media is syndicated to other news organizations, they will be able to find the manifest certifying its origin.

As for end-users who consume media through digital platforms, their chosen video player should be able to access the authentication too. This means that distribution platforms, social platforms and CDNs will need to cross-certify the same content during their respective content modification to create streaming renditions.

CONCLUSION ON FACT

To be effective, FACT will need to be widely adopted across the industry by both news production technology and distribution platforms. The FACT technology has a huge potential but first requires deep collaboration between news operations, news technology providers and distribution platforms. If this effort materializes, there is true potential to prevent Deepfakes' damage and spread of Fake News.

Automated Fact Checking

Recent developments in TL (Transfer Learning with BERT or GPT-2), a modern neural network technique, can automatically write an article in a way that is hard to detect that it was machine generated. Try it yourself! For instance, this link uses GPT-2. Assign a Fake News headline, and marvel at the results.

There are a wide variety of new technologies designed to "detect Fake News." Some detect text generated by a machine - like GLTR (a tool issued from a collaboration between MIT IBC Watson AI Lab and Harvard NLP). Others focus on checking the validity of the source. Thwarting fake news, especially during elections, has become critical for journalists. They must check if a claim is true or false, proven or unproven. Every reputable journalist goes through this exercise. And this particular exercise is a very time consuming activity.

CAN TECHNOLOGY HELP FACT-CHECKING ACTIVITY?

We discussed before how FACT can help establish trust based on media provenance. There are also AI algorithms available that can help detect machine-based text.

There is another initiative based on neural network technology that can potentially help with fact checking. Its model trained on a large dataset called FEVER, which is derived from Wikipedia. It is designed to identify sentences that provide evidence of content sources within a large document collection, assigning a score to the potential veracity of a claim. What is interesting about this approach is that it provides a clear view of who supports and who refutes a particular statement. This information can be incredibly helpful to a journalist. More information about this initiative can be found here.

The downside to this approach is the results. Only 58% of the claims that were checked had correct predictions, whereas only 59% of the returned evidence was relevant. This number is far too low to be a trustworthy mechanism.

Transfer Learning is another initiative [4] used to detect Fake News. A fairly new machine learning method where a model developed for a task is reused as the starting point for a model on a second task. The engine was trained to detect Fake News using a very small dataset from the Philippines, with only 3,206 news articles labeled as real or fake. Such method reached an accuracy of over 75%, very promising for such a small dataset, but still too far from reliable. What's interesting about this initiative is the small size of the dataset required to train the engine, considering that collecting a dataset is the most expensive and time consuming part of developing a trustworthy AI based algorithm. This initiative demonstrates that new Natural Languages techniques can lower the size of the dataset required.

CONCLUSION ON FAKE NEWS AND FACT CHECKING

We understand that technology alone is not sufficient to accurately check facts and detect Fake News. No technology can replace good journalism practice - a fact where the news community is in agreement.

However, we understand that technology can potentially be of great assistance to good journalism practices: tools that assist with discovery search and content creation, tools that are designed as user companions from the ground up.

Implementing these technologies within news production tools is not a technical challenge. The challenges revolve around the user experience as well as the potential editorial value that such technologies can bring to news creators. It is the responsibility of technology providers to create intelligent tools that empower journalists to work quickly and confidently through the masses of information available to them, allowing them to create truthful and meaningful stories their audiences can trust.

References

[1] Shearer, Elisa and Eva Matsa, Katerina. "News Use Across Social Media Platforms 2018," Pew Research Center, September 10, 2018.

[2] Shane, Scott. "The Fake Americans Russia Created to Influence the Election," The New York Times, September 7, 2017.

[3] England, Paul, Malavar, Henrique S., Horvitz, Eric, Stokes, Jack W., Fournet, Cédric, Chamayou, Amaury, Clebsch, Sylvan, Costa, Manuel, Erfani, Shabnam, Kane, Kevin, Shamis, Alex. "Authentication of Media via Provenance," arXiv, January 30, 2020.

[4] Blaise Cruz, Jan Christian, Agatha Tan, Julianne, Cheng, Charibeth. "Localization of Fake News Detection via Multitask Transfer Learning," arXiv, October 30, 2019.

About Dalet

Dalet solutions and services enable media organisations to create, manage and distribute content faster and more efficiently, fully maximising the value of assets. Based on an agile foundation, Dalet offers rich collaborative tools empowering end-to-end workflows for news, sports, program preparation, post-production, archives, radio, education, governments and institutions. Dalet platforms are scalable and modular. They offer targeted applications with key capabilities to address critical functions of small to large media operations — such as planning, workflow orchestration, ingest, cataloguing, editing, chat and notifications, transcoding, play out automation, multi-platform distribution and analytics.

Dalet Solutions

| NEWS | SPORTS | RADIO | PROGRAMS | POST-PROD | ARCHIVES | ORCHESTRATION | EDUCATION | GOVERNMENT | BRANDS |

Dalet Technologies

DALET Galaxy five   OOYALA FLEX MEDIA PLATFORM

DALET AmberFin   DALET Brio   DALET Cube   DALET OnePlay   DALET Media Cortex   DALET StoreFront

STAY TUNED

Receive e-mails from Dalet featuring new content that matches your interests.
Visit dalet.com to subscribe to Dalet Newsletters.

Find out more: Contact us | dalet.com

DALET