

EBU

OPERATING EUROVISION AND EURORADIO

NEWS REPORT 2019

THE NEXT NEWSROOM

UNLOCKING THE POWER OF AI
FOR PUBLIC SERVICE JOURNALISM



ABOUT THE EBU

The European Broadcasting Union (EBU) is the world's foremost alliance of public service media (PSM). Our mission is to make PSM indispensable.

We represent 116 media organizations in 56 countries in Europe, the Middle East and Africa; and have an additional 34 Associates in Asia, Africa, Australasia and the Americas.

Our Members operate nearly 2,000 television and radio channels alongside numerous online platforms. Together, they reach audiences of more than one billion people around the world, broadcasting in more than 160 languages.

We strive to secure a sustainable future for public service media, provide our Members with world-class content from news to sports and music, and build on our founding ethos of solidarity and co-operation to create a centre for learning and sharing.

Our subsidiary, Eurovision Services, aims to be the first-choice media services provider, offering new, better and different ways to simply, efficiently and seamlessly access and deliver content and services.

We have offices in Brussels, Rome, Dubai, Moscow, New York, Washington DC, Singapore and Beijing. Our headquarters are in Geneva.

Discover more about the EBU at www.ebu.ch

EXECUTIVE SUMMARY

Large-scale use of data and artificial intelligence represents the fourth major wave in the digitalization of journalism (after online, mobile and social media). It is a significant opportunity for public service journalism, but there are also considerable risks. In particular, the new communication paradigm could heighten the dominance of technology platforms in the flow of information between people.

According to a recent study, the introduction of artificial intelligence into newsrooms is considered a matter of 'significant concern' by two-thirds of newsroom managers, second only to funding issues.

In the EBU News Report 2019, we analyse the opportunities as well as the threats, and try to find the best practices usable in public service journalism. For a more comprehensive view, we have conducted 75 interviews with subject-matter experts from the worlds of journalism, technology and academia. We have also studied relevant literature and articles and have been in close contact with leading authors who have published recently on the subject of AI and its relationship with the news industry.

As in the 2018 report, [50 ways to make it better](#), we have formulated our recommendations on the basis of concrete examples. Supported by 30 case studies and 10 checklists (containing links to relevant further information), this report aims primarily to provide strategic and practical starting points for any newsroom looking to move to a data-informed model or invest in artificial intelligence.

OUTCOMES

The next wave in the digital transformation – in which large amounts of data will come into use and machine-learning technologies will be implemented in producing and distributing journalistic output – has fundamental implications for public service newsrooms (as well as for society at large). Even the most digitally advanced public-service news organizations have so far found it hard to adapt fully to the online environment and reach new audiences. Most of their online news audience are still older public-service loyalists, relationships with whom are still mostly maintained through broadcast channels.

To be successful in this next wave, we argue that both cultural and technical challenges must be addressed. Strangely, we have found that unfounded fears of the new technologies and undue expectations regarding their capabilities sit side by side. More fundamentally, fears of loss of journalistic identity and autonomy persist. Wherever public service journalism is unable to overcome these fears, opportunities will be lost, raising the risk that organizations will lag behind.

Our conclusion is that only some parts of journalism will be automated. However, virtually all jobs will undergo some degree of change. Newsroom staff will need a better grip on the technology.

We group and analyse artificial-intelligence solutions according to their benefits: in smoothing operations, in enhancing distribution, in producing content. We also delve into the best uses of data for increasing the public service value of journalism.

1. OPERATIONAL USE OF AI

Adding AI on operational level increases efficiency in news production. However, newsrooms will not be able to stand out from the pack simply by using this kind of approach, as the technology will be widely available. Most of it is offered by third-party vendors. As the potential for differentiation for public service journalism is low in this respect, investing in these technologies – whenever the time is right – will be a matter of straight cost-benefit analysis. Operational AI is an important area with which to keep in step, but it is not where the most valuable benefits for public service lie.

2. STRATEGIC USE OF AI IN DISTRIBUTION

To optimize distribution, we recommend that public service media (PSM) should invest strategically in personalization and the targeting of individual needs. Although there is some scepticism surrounding this, especially a fear of creating filter bubbles, personalization can enhance the value of public service journalism, as it enables useful content to be found by interested audiences. The new networked communication paradigm, in which people are in touch with a multitude of others, also requires this type of intervention. If high-grade content cannot be found, it cannot be consumed. We also stress, however, that ‘public service distinctiveness’ has to be coded into the algorithms to ensure societal value and avoid pitfalls and biases.

3.. STRATEGIC USE OF AI IN CONTENT CREATION

AI can be used to contribute significantly to the public service mission when deployed strategically in content production. For example, new technologies and tools such as advanced data science and artificial intelligence can be adapted for use in the mission-critical area of investigative journalism. They can also be used to scrutinize algorithmic decision-making for bias and manipulation, and offer new ways to verify and check information, to maintain trust in journalism. As the technical capabilities for spreading fake news increase, effective verification will in the future require for technology and people to work together.

We foresee significant opportunities arising for PSM from current progress in automated language technologies, such as automatic translation, voice recognition, and voice and text generation. For organizations operating in countries with more than one official language, or organizations collaborating internationally, these technologies represent a unique prospect.

To seize every opportunity, organizations will have to overhaul work processes, cultures and even themselves. Having a data-savvy culture that understands audience diversity will heighten their ability to respond with relevant content. Public service newsrooms should move quickly from measuring only reach to including other metrics that more accurately estimate the real degree of connection with people.

Successful AI projects are typically ‘owned’ by top managers, run by an in-house team, and have the size and scope that matches their strategic aims. Management must stress the value-based goals but then give sufficient autonomy to staff to attain those goals. In this highly competitive field, creating the right working environment is of paramount importance for attracting and retaining top talent. This implies investing in people: training existing staff, hiring people with diverse skill sets, driving cultural shifts and focusing on journalistic identity.

The deployment of ‘machines’ rather than people is often considered an austerity measure. That is far from the truth. AI is expensive: it requires highly skilled work to be performed by highly paid specialists. In most cases, significant scale is needed to successfully build real-life machine-learning systems.

In summary, public service journalism – to be able to respond to the changes in its environment– needs to collaborate more actively with national stakeholders and especially internationally, among organizations. Only by pooling resources can public service newsrooms counter the competitive threat from the ‘big tech’ platforms.



TABLE OF CONTENTS

CHAPTER 01	
IT'S TIME TO CATCH UP!	7
CHAPTER 02	
AI IS A HANDY NEW TOOL: USE IT!	22
CHAPTER 03	
DON'T THINK YOU KNOW THE AUDIENCE	34
CHAPTER 04	
CONTENT ONLY HAS AN IMPACT WHEN IT'S FOUND	55
CHAPTER 05	
FIND THE NEW PUBLIC SERVICE IN AN ERA OF AI	89
CHAPTER 06	
IMAGINE YOURSELF IN A BETTER FUTURE	117
CHAPTER 07	
CLOSING	139

CHAPTER 01

IT'S TIME TO CATCH UP!



THE FOURTH WAVE OF DIGITAL TRANSFORMATION HAS ALREADY STARTED

The fourth industrial revolution, under the auspices of artificial intelligence, is transforming work, lives, economies and democracies. For the news industry, this also marks the fourth wave of digital transformation after the online, mobile and social media. There's much hype surrounding AI, but for those who see its potential, it could make public service journalism both more valuable to audiences and a more inspiring job to do in the first place. Yet the threat is that poor comprehension of AI and unfounded fears of the technology will cause legacy media to lag behind. Again.

"Create compelling themes for your projects faster than ever before." A web service provides AI assistance for composers that is good enough for professional use. AI helps design furniture and clothing that people will actually like and buy. By analysing movie scripts using an algorithm, box-office revenues can be predicted with an accuracy close to 90%.

AI helps to design whiskey. AI turns plain pictures on a screen into paintings in the style of Picasso or Cézanne. AI creates photos of people who don't exist, providing advertisers with synthetic celebrities without the need to agree on rights or pay royalties.

"THE BIG MESSAGE IS THAT THERE IS A LOT STARTING TO HAPPEN AND THAT WE'RE ON THE CUSP OF THESE TECHNOLOGIES BECOMING SIGNIFICANT PARTS OF NEWS PRODUCTION." (CHARLIE BECKETT, PROFESSOR IN MEDIA AND COMMUNICATIONS, LSE)

You thought creative work belongs just to the human domain? Not anymore. The moment of singularity, when self-aware computers will surpass human abilities, may still be decades ahead (or pure science fiction), but computers can even today act creatively – if that means creating something that is new and valuable.

Computers are able to create a news agenda and assign tasks to humans. News stories are being sourced by algorithms, created by algorithms, distributed by algorithms. And algorithms are determining who gets to see or hear them.

IF AI IS DONE WELL, THE RESULTS ARE ALREADY AMAZING

In communication, we are experiencing the largest transformation since the advent of mass media. This also marks the largest societal upheaval since the Second World War. Algorithms know what we do, where we travel and even how we feel. They sometimes make us think we made the decision ourselves, although we were nudged

in a particular direction by a corporation or sometimes even our government. So, what can AI do for journalism in 2019? To start with, let's look at a couple of examples.

In Sweden, regional publisher Mittmedia has achieved amazing results using a personalization strategy that combines journalistic target-setting with a clear approach to contextualizing data and using machine learning. "I think that we will see total personalization and automation of publishing within the near future," says the company's chief technology officer.

In London, The Times and its personalized email newsletter, James, is another example of advanced thinking in this field, using some of the most sophisticated individual targeting methods available.

Chinese news app Toutiao uses machine learning algorithms to source and surface content that users will find most appealing. Toutiao learns about readers through their usage – taps, swipes, time spent on each article, pauses, comments, interactions with the content and location, time of the day the user reads the material – but doesn't require any explicit input from the user.

The results are both amazing and frightening, as the app has hundreds of millions of users who use it for longer than people use Facebook on average.

Journalism as a profession will fundamentally change

With these (and other) incredible examples in mind, it's easy to understand why the OECD has estimated that almost half of all professions will either disappear or fundamentally change within 15-20 years because of automation and new self-learning technologies.¹

Shifting from the one-to-many to the many-to-many communication era, billions and billions of connections exist between individuals, powered by automation, defining our understanding of the networked world we live in.

Technologies don't exist in isolation from the culture and values of news journalism. Some academics see increased automation as leading to a major redefinition of journalistic values in closer connection with computing.² Others argue that communication can no longer be defined as only a human-to-human phenomenon.³ So we may have to ask even the most fundamental questions, namely what is journalism, what is its purpose and how should it be performed?⁴

First came online. Then came mobile. Then social media and platform dominance. The fourth wave of AI is already here. In a recent study of 4,100 newsroom managers and journalists from across 149 countries, introducing artificial intelligence into the newsroom was considered a major challenge by almost two-thirds of respondents, second only to concerns over revenue streams.⁵

This fourth wave of digitalization spells both huge opportunities for public service journalism and considerable risks. Even the most digitally advanced public-service news organizations have so far found it hard to adapt fully to the online environment and reach new audiences. Almost all of the online news audience are the same ageing public service lovers they already have and the relationship is still mostly based on broadcast channels.⁶

Silently, technology giants have assumed the role of news gatekeeper for large parts of the population. The secret behind their success: well-designed platforms superpowered by artificial intelligence combined with very advanced IT. This shift offers a new opportunity to catch up. Well-led public service organizations with a clear vision can now identify ways of increasing value for citizens and societies and, by doing so, become 'first movers', with all of the advantages conferred by this role.

1 (OECD, 2019)

2 (Coddington, 2015; Milosavljević & Vobič, 2019b)

3 (Lewis, Guzman, & Schmidt, 2019)

4 (Milosavljević & Vobič, 2019a)

5 (Owen, 2019)

6 (Schulz, Levy, & Kleis Nielsen, 2019)

Factbox: What do we mean by 'AI'?

Artificial intelligence, machine intelligence, machine learning, deep learning: people use many words when talking about this topic, each meaning something slightly different. But in its most basic form, artificial intelligence is 'a system that makes autonomous decisions, a branch of computer science in which computers are programmed to do things that normally require human intelligence.' One of the pioneers of artificial intelligence, Marvin Minsky, describes artificial intelligence as a 'suitcase term': there are many concepts packed inside. In our Glossary of AI, we explain key terms and concepts.

NEWSROOMS FEEL AN URGENCY TO GET INVOLVED

An astounding 15% of the total data-centre capacity worldwide is now used for machine-learning tasks. However, legacy media seem slow in adopting these technologies.⁷ And when you look at the media companies that are dominating the market, you see mainly tech giants: Google, Apple, Facebook, Amazon, IBM, Microsoft, Baidu, Alibaba and Tencent.

At Microsoft, for example, there are 8,000 machine-learning specialists in employment, at Amazon's Alexa 5,000 and at Alibaba a massive 25,000 employees are working on AI-related tasks. Google has made the work of its approximately 1,000 researchers publicly available, and its subsidiary DeepMind has doubled its payroll since 2017. Even back then it already had close to a 1,000 people developing deep-neural-network technologies for use in general-purpose AI.

As Patrick Arnecke of Swiss SRF concludes: "These are stunning numbers compared to anything that we as broadcasters might have at our disposal."⁸

Legacy media companies have not been completely absent from the field though: various public service media (PSM) organizations are experimenting with and implementing machine intelligence, as the examples in this report will show. And the EBU has recently set up an '[AI and Data Initiative](#)' (formerly the Big Data Initiative) to collectively address topics like cross-platform audiences, content creation via data-driven editorial production, and new marketable products and services.⁹

An insight shared by most PSM organizations is that artificial intelligence will quickly grow in importance, not just as a tool to be used, but as a topic to report on. Since machine intelligence will be applied increasingly more in all facets of society, it needs to be held accountable.

That newsrooms feel an urgency to get involved in the matter was noted by researchers in a project at the London School of Economics who ran a survey amongst 'AI savvy or technologically savvy people' in more than 70 news organizations worldwide.¹⁰ According to Charlie Beckett, Professor in Media and Communications and the project lead: "The big message, I think, was that there is a lot starting to happen, and that we're on the cusp of these technologies becoming significant parts of news production."

THE CHALLENGE: FILTER THE OPPORTUNITIES FROM THE HYPE

The increased use of artificial intelligence is not leading to a decreased requirement for human intervention. On the contrary, human capital will be needed even more in the future, as the authors of a recent book about the 'simple economics' of predictive AI explain: "All human activities can be described by five high-level components: data, prediction, judgment, action, and outcomes. As machine intelligence improves, the value of human prediction skills will decrease because machine prediction will provide a cheaper and better substitute. However, the value of human judgment skills will increase."¹¹

From a practical point of view, it seems that the challenge is more about aligning processes in such a way that the strength of both machines and humans can be used in the best possible way to fulfil the public service mission. Nicholas Zembashi, of Forensic Architecture, says: "We see AI as a team member, with a very specific skill set, but with its shortcomings."

⁷ (Bremme, 2019)

⁸ (Arnecke, 2018)

⁹ <https://www.ebu.ch/ai-data>

¹⁰ (Beckett, 2019)

¹¹ (Agrawal, Gans, & Goldfarb, 2016, 2018)

But exactly how beneficial or dangerous artificial intelligence can be remains difficult to forecast. As the Future Today Institute Trend Report from 2019 puts it: "The lack of nuance is one part of AI's genesis problem: some dramatically overestimate the applicability of AI in their workplaces, while others argue it will become an unstoppable weapon."¹²

In this report, we will seek to address these issues from both a strategic and practical perspective. We do so recognizing that it is the role of journalism and particularly public service journalism to inform and educate the public about such significant technologies and their implications for communications and informed societies. But we do so acknowledging as well that these developments present both significant challenges and potential opportunities, and spell a certain degree of disruption for the media itself as agents of communication – and particularly for PSM, who have the responsibility to act in the interests of the population.

HOW TO READ THIS REPORT

This 2019 EBU News Report is closely connected to the previous EBU News Report – [50 ways to make it better. Building audience and trust](#). We have given special focus to the first of the five major challenges identified in the 2018 report: [The value of news: On trial](#). This year we ask: Which are the best means to increase the value of news, in a way fit for PSM? What are the newest AI and automation technologies? And how can we perceive the real value while filtering out the hype?

Ultimately, the question is: How can public service news be successful in its mission to serve citizens in the new era? The answer will be crucial to the ongoing existence of the public service concept.

We attempt to answer this answer in our six chapters:

- In Chapter 2, we start by giving an overview of all the ways artificial intelligence can help in the newsroom, as practical tools. This is AI at a purely operational level, aiming primarily to automate repetitive tasks and reduce cost.
- In Chapter 3 we dig deep into the world of journalistic data and metrics, since creating a data-savvy culture and knowing your objectives and ways to measure them are prerequisites to using new technologies in a sensible way.
- In Chapter 4, we zoom out and look at current strategic uses of artificial intelligence, namely content targeting and automated journalism, looking at how to use the technologies in a way that will really make a difference.
- In Chapter 5, we zoom out even further and look into the future: the specific public service opportunities that AI has to offer for PSM to be distinctive and offer value in societies undergoing change.
- In Chapter 6, we address the core challenge of the future: How to produce positive outcomes in newsrooms while maintaining creativity, employee morale and a strong journalistic identity while avoiding the pitfalls that could result from combining these values with technology, however good it may be.
- In all chapters, various case studies will deepen our insights. Many examples represent projects run by EBU Members, but we find inspiration elsewhere as well. The names and contact details of each 'case owner' are listed, if available, and these persons may be contacted for further information.

This report is by no means the ultimate survival guide for journalism in the fourth industrial revolution, so we also cross-reference some worthwhile books, articles and websites throughout the report. Several checklists and factboxes have been drafted containing additional links. Our goal: to guide newsrooms through the AI and data maze.

The analysis we present in this report is based on interviews we conducted among 75 people; from media and communications professor to head of strategy, from director general to head of marketing. In the source list [link](#), we have listed all the names and positions and quotes of the interviewees can be found throughout the report.

12 (Webb et al., 2019, p. 23)



BYTEDANCE, CHINA

TOUTIAO, A MOST INTERESTING NEWS APP

TAKEAWAYS

- China is a major player in AI; and because of its high level of sophistication in this field, and huge home market, disruptive -potentially challenge established tech giants such as Facebook and YouTube.
- Algorithms, when used extensively, are able to detect context and user's interests from mere behaviour and other data available without any user contribution.
- Mixing soft and hard news may be an effective way to collect the data needed to suggest the right news content to users.

The Chinese invented paper in around 100 BC. Two thousand years later, as paper has been widely replaced with digital news sources, one of the most noteworthy technical innovations has again hailed from China. Some might say that having more than 250 million monthly users is not such an amazing feat in a country of 1.4 billion people. But then consider also the average usage time per day, which is more than one hour. That's more time than people worldwide spend on Facebook, on average.

You swipe twice and your profiling has already started: Where are you? What phone are you using? What time is it? What content are you using? The app starts to suggest news to you that you might be interested in. It learns on the fly, and you don't have to do anything else other than using the app. In 24 hours, the system has created a decent profile about you as a user.

That's the key for user experiences with Jinri Toutiao (Today's Headlines, in Chinese), the news app which is built on a whole host of advanced machine-learning algorithms. Its success and original approach have attracted the attention of some of the most technological advanced public-service news organizations in Europe, and several study tours to China have already been arranged to learn more.

Toutiao was launched in 2012. It is run by ByteDance, a company valued at \$75 billion in its latest financing round. The company also runs Chinese social media app Doyin and in west better known TikTok, an app that distributes amusing selfie videos containing special effects. This app has 1,2 billion downloads worldwide, of those more than 100 million in the US.

Altogether, hundreds of millions of Chinese are using the company's various apps monthly. Toutiao is reported to be used, on average, for 74 minutes each day per user. Over half this time is spent watching videos, with an estimated 10 million views per day.

Toutiao is now entering the English-speaking market as well.

Toutiao is like Facebook, YouTube and all your news apps rolled into one. The uniqueness of it lies in the deep-learning algorithms powering it, enabling it to recommend content without any proper user input or purchase history. There's a fundamental difference, however, in its algorithms compared to Facebook: TikTok does not rely on user's connections with others but uses AI to analyse content and users and maps out interests and desires they may not even be able to articulate to themselves.





Toutiao learns from taps, swipes, time spent, time of the day, pauses, comment, interactions with the content and location, and then cross-tabulates these in its algorithms and detects what type of content might be interesting to each user.

The technology is proprietary, of course, but it's said that each user profile has millions of dimensions, based on algorithmic calculations. Toutiao analyses content: Is it trending? How long is it? How up-to-date it is? Is it evergreen or just interesting today?

Most western news apps still recommend content based on the algorithm developed for Google News more than ten years ago. This 'collaborative filtering' assumes that if your behaviour is similar to others, it will continue to be so. Toutiao is much more context-aware and so it has claimed able to spot the needs of the user more specifically. Even so, it's still easy to use.

To be successful in machine learning, you need data. To get data, you have to be successful, preferably in a large market. Like China.

The ML engine profiles the user quickly and typically reaches quite a high level of accuracy in just one day. As a result, people come back and stay.

Now, let's dig into an even more fundamental process in Toutiao: It has also started to produce and assign content with algorithms. It started its robot journalism operations by producing news during the 2016 Olympics in Brazil.

Toutiao's approach to automatic content creation is more advanced than most western counterparts. They select relevant sentences from live commentary and relevant images by analysing their pertinence based on historical data.

Add to that automatic highlights and selection of best titles for each user. All these technologies exist separately and in labs, but you rarely see them working together in a live system. Recommendation platforms the company is willing to licence to other companies but not the ones on content: they are the focus in the company now and predicted competitive assets of the future.

It also has a programme encouraging people and organizations to contribute content to the platform, with almost 1 million registered accounts, and it shares its advertising income with those who provide the content.

By its very nature, Toutiao is a news aggregator. That gives it a fundamental advantage with diverse content over conventional news organizations that publish only their own stories. It can select the personalized content from more than a million sources and is not bound by the material its own newsroom produces. ByteDance is constantly facing accusations of breaking privacy rules or pumping their figures up by buying both content and traffic that help in its way to the front of the listings.

Toutiao also uses algorithms to filter out content that it doesn't want to show users. This may be used to censor content according to authorities' requirements, but also to combat 'fake news' or optimize brand image of the app towards the desired outcome. However, a great deal of celebrity news is still available, to which both users and algorithms seem to respond positively.

FOR MORE INFORMATION

<https://www.toutiao.com/>



MITTMEDIA, SWEDEN

ALGORITHMS MATCHING EDITORS

Robin Govik
CDO
robin.govik@mittmedia.se

TAKEAWAYS

- Mittmedia had clear goals based on deep contextual understanding: that it is important to know the audience and their news consumption in order to select relevant content.

- Mittmedia's data strategy and data architecture made it possible to use different kinds of data in complex algorithms.

- Journalistic values were carefully taken into consideration.

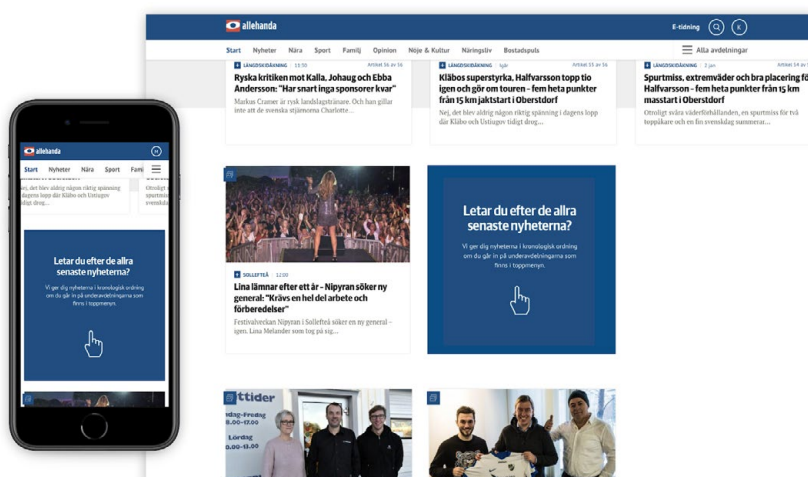
Swedish regional newspaper publisher Mittmedia has become an international superstar by using AI to personalize its online offerings and produce robot journalism. The need for automated content was detected by analysing data. And the company has a clear understanding about how data should be used.

It was a frightening insight for Swedish newspaper group Mittmedia. New subscribers, those who finally had decided to pay to read their newspaper online, after an initial period of intense reading activity, reverted to their previous consumption level after just one month.

If people read more, they remain happy subscribers. If they don't, they'll cancel. In commercial terms, this is called 'churn'.

"I THINK THAT WE WILL SEE TOTAL PERSONALIZATION AND AUTOMATION OF PUBLISHING IN THE NEAR FUTURE." (ROBIN GOVIK, CDO, MITTMEDIA)

"The business problem we had was that people stopped using our products. So we tried to enhance retention, and personalization was one of the methods we started to look at," tells Robin Govik, Mittmedia's CDO.

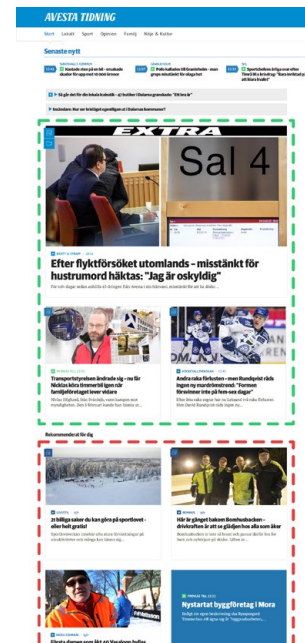


The data-science team set themselves a clear business target: double the number of stories subscribers clicked on from the homepage during each visit. It started analysing users based on data from three dimensions: users, content and interaction between the two. In other words: who reads what, and when, as an attempt to ascertain the why.

The team discovered that the usage patterns of different user segments differ radically. One segment may consume 70% of its content during the morning, while another hardly ever visits the site in the morning, instead spending half of their time on the site in the afternoon.

This pattern stays strikingly the same day after day, indicating that people have hard-wired routines for viewing news.

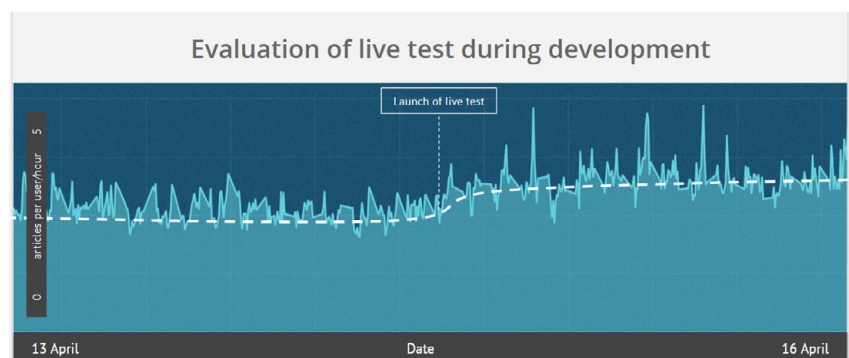
Next questions: What kind of content is each of these segments using? What is their typical age? What is the probability of churning?



This analysis yielded many immediate practical applications. The company designed a daily publishing clock, similar to models developed by Yle or NRK, or earlier by tabloids such as Aftonbladet. The importance of weekends was taken into account along with recognition of which content is consumed at different times of the day, and by whom.

In addition, one's interests stay relevant even in the presence of breaking news. The usual interests are simply deferred: people read the stories that interest them once the breaking situation has died down.

Mittmedia's team concluded that the main problem for users was that they had trouble finding the right content that they felt was valuable for them. At the same



time, the company had to save money and reduce the effort deployed in publishing stories on different websites, which in Mittmedia's case are numerous, as the company is a regional publisher.

Meanwhile the company had made a strategic investment in people. One example of new hires was Xiu Li, who carried out complex data analysis of content consumption from the small Swedish town of Sundsvall. Mittmedia's data science team are all in-house staff. No outsiders are involved.

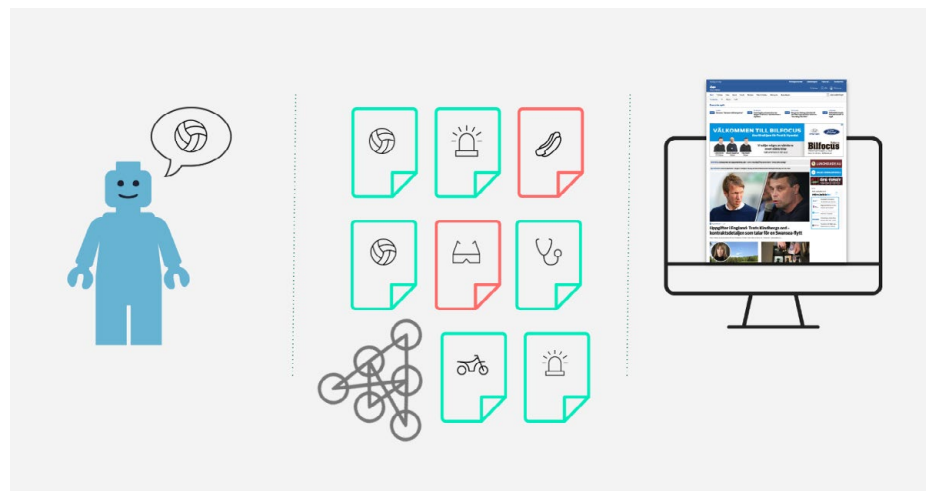
Mittmedia started to explore personalizing their content offering using machine learning. They set algorithms to cross-tabulate interest clusters with geographical clusters. One elegant example from the solution is that geographical interest was not only detected based on the location you were in but also on the geographical interest patterns yielded by the content viewed, as supplied by machine-learning algorithms. The first trials were run in November 2017. When measured, algorithms could already then match editors in selecting stories for front pages. "We were quite convinced that an algorithm would do this as good - or even better - than people."

After around one year of developing the algorithm, at the same time making sure that all journalistic values were all taken into consideration, the real challenge was taken on in October 2018.

The results were amazing. The automated pages, mixing top stories selected editorially and rest of the pages with algorithmic selection, increased metered usage significantly.

Time consumed on the sites increased immediately by roughly 10%. Click-through rates from the front pages rose 50%. But even more importantly for creating loyalty, the frequency of unique visits per day more than doubled and the churn rate declined around one- third in the six months after the launch of new personalized front pages. And yes, the number of articles read during each visit doubled.

For public service companies, here is one more key finding: consumption of quality niche content increased significantly once the content found its audience. For example, the culture segment roughly doubled its visits in the new system.



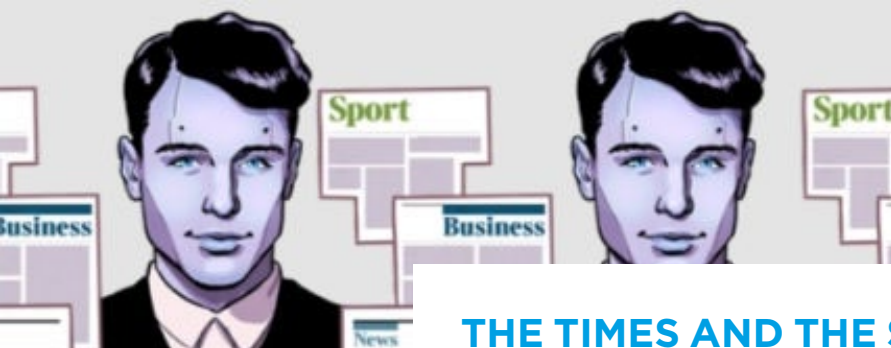
Robin Govik believes that having the data readily available and properly organized is key: "... in the same place and organized in a way that is accessible. You can call it a data lake. We have descriptive data on articles, we have the addresses, gender, all the demographic data and we can match that with behavioural data based on interactions."

The results yielded for Mittmedia, which is facing a constant need to reduce costs, foreshadows the company's next move:
"I think that we will see total personalization and automation of publishing within the near future."

MORE INFORMATION

<https://www.mittmedia.se/>

<https://medium.com/mittmedia>



THE TIMES AND THE SUNDAY TIMES, THE UK JAMES, YOUR DIGITAL BUTLER

Michael Migliore
Head of Customer Value, News UK
michael.migliore@news.co.uk

TAKEAWAYS

- There's a clear distinction between segmenting the audience and targeting their needs individually. All news organisations can learn from subscription-based newspapers about how to create value that users are prepared to pay for.

- The Times' digital strategy is bold and distinctive. It does not follow others by publishing news once it's ready but waits until the next edition. This is one way of creating habits: readers know when to expect an update.

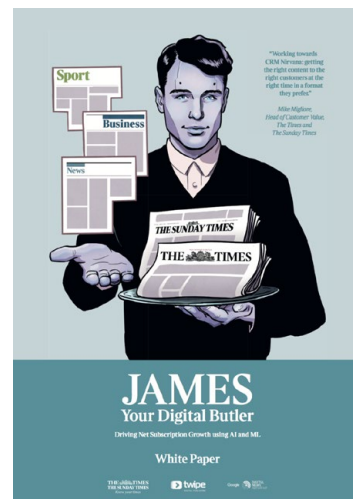
- A newsletter as an email can be a powerful tool to distribute news, especially when tailored individually. Tailoring a newsletter is cheaper than tailoring digital offers across the board.

The Times of London created an individualized news service, in which its digital butler James carefully tailors email newsletters to each subscriber. Consequently, subscription cancellations have fallen by 49%. The Times was among the first newspapers to establish a paywall. It is a no-brainer that returning customers are more likely to stay on as subscribers, so the key to long-term success is habit creation, and individual tailoring has proven its value in this respect.

"How, as a business, are we ever going to create relationships with people we're never going to meet?" says Michael Migliore, formulating the key question of digital mass-marketing – or communications – in our time.

"Irrelevant". That's how he describes the results of the frequently used marketing tactic of segmenting people into groups and using personas to describe them. He says that most traditional mass-marketing campaigns targeted at these groups result in people feeling that the three seconds they spent opening the marketing emails was a huge loss, especially when the technology exists to better consumer expectations.

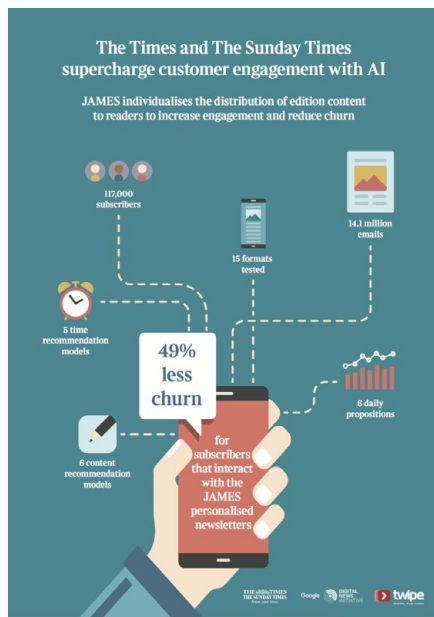
Instead of personalization, Michael Migliore likes to talk about 'individualization'. Here the objective is to understand, predict and meet individual consumer needs. "Good individualization does all of those things for each person," he says. "It's getting the right content to the right person at the right time, in the right format."



"THE COOLEST THING ABOUT THIS, WHICH DRIVES OPERATIONAL EFFICIENCY, IS ADAPTING AUTOMATICALLY TO WHAT THE USER DOES OR DOESN'T DO." (MICHAEL MIGLIORE)

Michael defines efficient marketing and stakeholder relations using key questions: "Who am I talking to? What am I going to say? And how am I going to say it?"

In designing *James, Your Digital Butler*, who emails an individualized news summary to readers, he only had to add two more questions to the original three: 'When?' and 'How frequently?'



If public-service companies are hesitant in adding personalization to their offering, because they fear for their heritage, they can find inspiration in The Times, which was established in 1785, five years before the French Revolution. In the last ten years, the newspaper has faced some of the toughest times in its existence. These days, it is one of those few news outlets in the UK that is growing. Over half of the readers are on digital.

Establishing the digital paywall in 2010 was not an easy task. About half of those who subscribed soon cancelled. The solution was found by looking deep into the data on the subscribers and their behaviour. On the surface, it seemed that the longer you had been a subscriber, the less likely you were to cancel. But looking more carefully, analysts at The Times could see a deeper layer: "It was about how quickly we can build a habit with you."

By that time, the segmentation strategy had reached its limits. Grouping audiences did not create a loyal readership. "That's why we eventually turned to AI and James."

The Times received support from Google Digital News Innovation Fund and started to collaborate with Twipe, a Belgian analytics and AI company. "The vision for *James* was to meet this crucial challenge to deliver better-individualized customer experiences, stronger growth and unprecedented levels of operational efficiency, all through machine learning. So more relevant content, for the right people at the right time, more cheaply, to demonstrate value."

When *James* was built, the content was the answer to the question of *what you are going to say* and format to the question of *how*. But how were we to answer the question *when* and *how often*? "Basically, we created algorithms across all of these dimensions," says Michael Migliore. They copied the testing model from Netflix, interleaving, in addition to implementing 'multi-armed bandits', to collecting and adapting all the time according to the behaviour of each customer.



For example, the most effective email layout for some of the audience was one without any photos, which was against the previous received ideas in the industry. But this is not necessarily the best layout for all types of reader, so the individualized user preference must be sourced from the data. "Our multi-armed bandit chooses which format to use for each user. And then, of course, the coolest thing about this, which drives operational efficiency, is adapting automatically to what the user does or doesn't do." This removes the need for an over-reliance on creative production and post-campaign analysis, driving costs down.

Success factors were, according to Michael Migliore, the development of cross-continental collaboration between Bangalore, Barcelona, Belgium and London, and working cross-functionally, with flat structure and autonomy that gave everyone a voice. "Talking to senior leadership was: leave us alone and do our thing. And we'll let you know."

James sent 14 million personalized news-summary emails over a 12-month period. Consequently, 70% of subscribers are actively engaged with this service. On the other hand, 20% are not, and according to Michael Migliore, "they're doing fine." They have already their other firmly established news-related habits. The biggest impact has been among low-to-medium-engaged subscribers.

While the traditional recommendation model can reduce the churn by 14%, *James* can decrease it by 49%. "That says pretty loud and clear to me that individualization provides a powerful return on investment."

FOR MORE INFORMATION

<https://www.thetimes.co.uk/>

<https://www.twipemobile.com/media-innovation/james-your-digital-butler/>

<https://www.blog.google/outreach-initiatives/google-news-initiative/news-uk-mike-migliore/>

https://newsinitiative.withgoogle.com/assets/static/docs/the-sunday-times.pdf?_ga=2.34118241.1177072267.1572261416-1302913871.1559743076

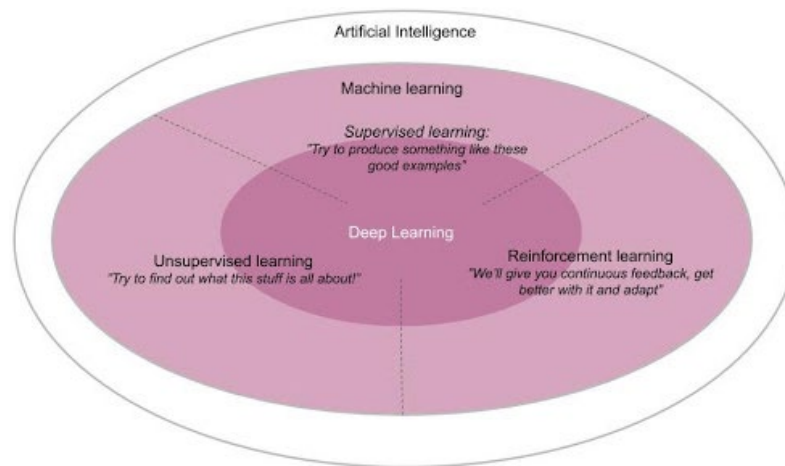
HOW TO

GLOSSARY OF

ARTIFICIAL INTELLIGENCE

The term 'artificial intelligence' was coined by John McCarthy, who began research into AI in the 1950s. He assumed that human learning and intelligence could be simulated by a machine.

The term 'artificial intelligence' is not clearly defined, however. One common definition is that artificial intelligence describes machine processes that would require intelligence if performed by humans.



SUBCATEGORIES OF ARTIFICIAL INTELLIGENCE

Most applications of artificial intelligence today use technologies that fall into the domain of machine learning. With **machine learning**, computers learn from data without being explicitly programmed.

- In traditional programming, data is run on the computer, which produces the output. Programmers therefore have to know the rules if the desired output is to be achieved.
- In machine learning, data and examples of the desired output are first run on the computer, which then learns from them to create its own rules. These are then used to produce the output.

A very powerful and resource-consuming kind of machine learning is called deep learning. Deep learning uses algorithms called 'artificial neural networks', which are modelled on the way neural networks operate in the human brain. Advances in deep learning have made language technologies and image recognition much more sophisticated. Deep learning name actually means that there is one or more extra layers in the neural network which usually has only three layers. The term "deep" doesn't come from deeper intelligence but deeper structure (i.e. more layers) in the neural network.

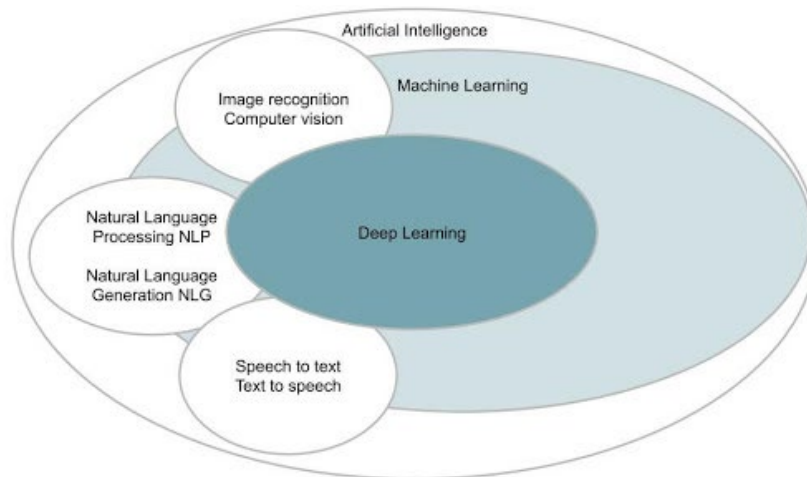
The key benefit of deep learning is that it is able to absorb huge amounts of data. This has allowed machine learning to accomplish tasks it never could have done before. On the other hand, it requires huge amount of training data and expensive computational resources, so using it on a large scale has only recently become possible and is still limited. Often deep learning doesn't work as well as some other type of algorithm.

LEARNING APPROACHES WITHIN MACHINE LEARNING

The different types of machine learning are based on how the machines use data to learn rules.

- **Supervised learning** is the oldest kind of machine learning. It uses a set of desired outcomes to train the computer. The algorithm then comes up with rules that will allow the computer to produce results similar to like the training dataset. An example is comment moderating, where a system searches for comments similar to those that have been considered inappropriate by human moderators.
- In **unsupervised learning**, the computer is used to group a huge dataset in a meaningful way. Because the desired output is different, this approach does not require a set of training data. One application is in clustering: news consumers or news articles can be clustered with this model.
- In **reinforcement learning**, the system learns on the fly from feedback it receives from its environment. This field is developing fast and is quite useful for news organizations, as these systems can quickly adapt to new situations and learn from user behaviour that can suddenly change for any reason. The most advanced recommendation systems use reinforcement learning as a way to censor the behaviour of users.

All three learning models can be combined in a complex system. An advanced system typically has multiple algorithms and can combine different approaches to achieve the desired outcome for a specific situation. For this kind of system, it's crucial to understand the problem, the context and the information contained in the data.



KEY APPLICATION AREAS OF ARTIFICIAL INTELLIGENCE

What's confusing about the different areas of artificial intelligence is that people often mix up the dimensions of AI.

For example, machine learning often appears alongside image recognition and natural language processing in lists of key AI application areas. But machine learning is present in almost all modern applications of artificial intelligence, including medical diagnosis, self-driving cars, prediction systems and automatic classification.

GLOSSARY OF TERMS

Algorithm

A step-by-step mathematical instruction to solve a problem. In machine learning, algorithms perform various tasks, such as clustering, recommending or classifying data.

Artificial general intelligence (AGI)

Also called 'strong AI', this is a type of AI that can solve complex problems in any context. Its opposite is 'weak' or 'narrow' AI. This kind of AI has not yet been achieved.

Artificial intelligence

A subset of computer science that aims to build machines capable of performing human-like tasks.

Artificial narrow intelligence (ANI)

Also known as weak AI. ANI performs a single task on a specific problem. Artificial neural network (ANN) algorithm with layers of 'neurons' modelled on the biological neural networks in the human brain. One algorithm family of machine learning.

Clustering

A task that involves grouping data into previously unspecified clusters. For example, it can be used in customer segmentation.

Convolutional neural network (CNN)

A neural network designed to detect, analyse and classify images.

Computer vision

A field of AI concerned with teaching machines how to interpret the visual world.

Generative adversarial networks (GAN)

A neural network that can, for example, create images based on the input it's given. Used in creative solutions.

Image recognition; Computer vision

A task in which systems try to identify objects, people, places, and actions in photos or videos.

Machine learning

A field of AI in which systems learn from data sets and perform a task with a specific target without being programmed to do so. This is the main field in AI today.

Natural language processing (NLP)

A system that helps computers simulate the ability of humans to understand language.

Reinforcement learning

A type of machine learning in which computers learn through reward and punishment and improve by trial and error.

Supervised learning

A type of machine learning in which an algorithm learns from labelled sample data, usually provided by humans.

Transfer learning

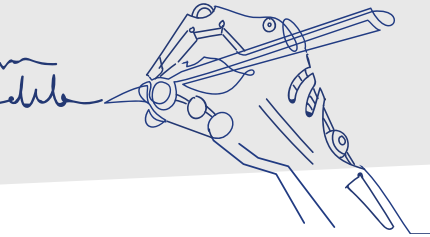
An approach that tries to use systems previously trained in one task to accomplish new tasks.

Unsupervised learning

A type of machine learning in which a computer tries to make sense of unlabelled data.

CHAPTER 02

AI IS A HANDY NEW TOOL: USE IT!



MANY BASIC AND AFFORDABLE AI RESOURCES ARE ALREADY AVAILABLE

Technology and journalism have always been intertwined. Smart machines already perform everyday tasks in newsrooms, sometimes even without journalists knowing it. AI helps to find new stories, automates tedious production tasks and helps with distribution. No need to be fearful: this will change journalistic work, but it won't replace it. Since common applications are – and will become even more – available for a decent price (sometimes even for free), the key is to use them smartly. This will make journalism not only more effective but also more interesting and fun.

One of the deep fears of artificial intelligence is that human work will be largely replaced with automation. But is this true? Will journalists ever be replaced with robots? “Not any time soon,” says Erik Brynjolfsson, MIT professor, who has studied the impact of AI extensively. He reminds us that artificial general intelligence (AGI) able to perform the broad set of tasks that people can do is still very far away. Instead of entire jobs, machines will replace routine repetitive tasks.

“IT’S VERY RARE THAT YOU SEE AN ENTIRE OCCUPATION BEING AUTOMATED. THE SAME IS TRUE FOR JOURNALISM.” (ERIK BRYNJOLFSSON, PROFESSOR, MIT)

“Machines have gotten extremely good at these,” he says. “But it’s very rare that you see an entire occupation being automated. The same is true for journalism, where some parts of the job will become automated, but not all.” Finding and asking the right questions, interacting with people, contextualizing the results and showing creativity will most likely still be the responsibility of human reporters in the future. Machines are editing audio and video, transcribing audio and finding relevant footage from archives in almost real time. AI is used in smart search engines and automated assistants. It generates silent videos and highlights reels.

It helps television weather forecasters publish in multiple channels. This is all very useful. And why then not use the tools available, as this will leave more time for more interesting and meaningful work?

¹ (Broussard et al., 2019; Mullin, 2016; Pallanich, 2019; Marconi and Siegmund 2017)

Many broadcasters rely on in-house software systems, and tech departments are often detached from the daily journalistic routine. Legacy newsrooms are not prone to use external tools and since the pace of technological advances is lightning fast, it's hard for them to keep up. This means that the main challenge may be to create a newsroom culture that is willing to explore, accept and keep up to date with new technologies.

We have summarized all the useful tasks that machines can do in the following paragraphs. A longer list can be found in our checklist [AI tools for newsrooms](#). We have grouped the tools according to journalistic process as described in the diagram below – ranging from idea to research, production, publication, feedback and archiving.



MACHINES FIND STORIES YOU WOULD NOT OTHERWISE SEE

In the idea-creation phase, machines can act as useful scanners for interesting and newsworthy information. Some media organizations are monitoring data sources to find new story ideas and discover topics that have not yet been covered.

Sometimes monitoring can be done with simpler rule-based tools, which can do the work with no advanced AI. Bayerischer Rundfunk (BR) in Bavaria, for example, has a [social listening desk](#) that uses a social monitoring tool coded with keywords that relate to their areas, and two people are working full-time to analyse the findings and inform the newsroom. By doing this they discovered, for example, the importance of [blue-haired video blogger](#) Rezo before he became an influential figure in the German election campaign.

The Finnish News Agency STT developed an award-winning simple-to-use tool for this: a system called [STT Little Bird](#) alerts them when a field on a website has changed in a pre-coded way; a politician has blogged, or a statistical figure moves beyond a threshold - for example.

Some newsrooms use more sophisticated [anomaly detection software](#) that analyse large datasets to [find newsworthy events](#) anywhere on the web. [Reuters](#) has developed the tool Tracer to help them find breaking news stories on Twitter, sometimes even several hours earlier than would have otherwise happened. It is now investing in an even more sophisticated news-detection and pre-processing tool, Lynx Insight, as it aims to become a 'cybernetic newsroom'.

CASE 1: FINDING STORIES WITH AI, REUTERS, INTERNATIONAL

MACHINES PRODUCE STORIES YOU WOULD NOT OTHERWISE MAKE

Machines can also be productive in later phases of the production pipeline. They can, for example, suggest archive material or information that fits the story you are writing, as the [Editor](#) tool is doing for the New York Times.

Sometimes even a whole programme can be assembled automatically from archive material, such as the BBC's [AI met the archive](#), which was co-presented by mathematician Dr Hannah Fry and a virtual host.

You will often hear that good metadata is needed to succeed in such a task and content needs to be tagged correctly in archives, otherwise the machine will miss it. But this is not an absolute requirement anymore as metadata can be [generated automatically](#), and machines are becoming so sophisticated that they are able to detect relevant content even when the metadata is missing.

For example, the French company [Newsbridge](#) has worked with France Télévisions to associate facial recognition, object detection and live-feed audio transcription and match with semantic context to find relevant footage much faster than before. And they are not the only company active in this field.

[Irish RTÉ joined forces with Al Jazeera](#) to adopt a solution from a technology provider that can detect political leanings in election programmes. Al Jazeera even uses this solution to track potential lapses in editorial quality in the footage and not just to identify objects and individuals. Alternatively, [RTS in Switzerland](#) has developed advanced archive functionalities in house.

Similar technology has been used in real-time news production, with cameras feeding their footage wirelessly to the cloud, which changes the news process dramatically. Footage can then be monitored, relevant pieces found and published in just minutes – if not seconds. Launches from leading tech players are to be expected in this field.

MACHINES DO PART OF THE WORK YOU NO LONGER WANT TO DO

There are numerous examples of tools that are designed to automate the more tedious parts of the journalistic work process. AI is for example moderating discussion boards using sophisticated learning models.²

Typically, these systems automatically select posts that can be published and those that should be removed. Then the system asks for human advice, and once the moderators have decided what to do with the posts, the system uses those decisions as learning material to make smarter selections in the future.

The tools can be adapted to reflect specific editorial standards as with the case of the [Washington Post and its ModPot engine](#), which allows more criticism in the comments if the subject is a public figure.³

In marketing and commercial social-media operations, automation is used regularly for the distribution of stories, and some news organizations are following suit and [automating their social media publishing](#), like Echobox is doing. The calculation is that computers do a better job than humans in mastering large numbers of distribution channels and targeting specific groups.

The technology's potential is shown in the innovative way which Dutch media company [Talpa](#) has developed its new social-platform video service for football. Tens of thousands of amateur football games are transmitted without any human involvement, using high-definition cameras and a machine-learning-based virtual director who can detect hotspots in the game and zoom in to where the ball is. In Germany, the football association has created a system in partnership with technology firms to automatically mix voice based on the position of the players in the field. These technologies could open new paths for inexpensively producing multi-camera-like programmes and live streams.⁴

RECOMMENDATIONS FROM THIS CHAPTER

1. The technologies highlighted in this chapter are widely used and developing fast. There will be lots of companies providing these services at a decent price. It's up to individual broadcasters, in their budget policy and cost/benefit analyses, to decide when to use the technologies.
2. Many tools exist to automate production and will change the work processes significantly.
3. The above technologies typically reduce costs or make the work in the newsroom more effective. However, they typically don't make a strategic difference for public service media.
4. The data gathered or created by some of these resources may be needed later for more strategic purposes. Therefore, it's important to plan how the data is stored for future use.

2 (Sallinen, 2019)

3 (Broussard et al., 2019)

4 (The Broadcast Bridge, 2017)



REUTERS

FINDING STORIES WITH AI

Rogelio Escalona

Director, Business Development and Strategy, Reuters

Rogelio.Escalona@thomsonreuters.com

TAKEAWAYS

- Automated systems are worth investing in for cases that scale. Even Reuters, with its global coverage, has found it wise to limit the efforts in language processing and let humans do the part of the job at which they are better.

- AI is best used with topics that pop up often, and in which the datasets are standardized. The more manual work needed to make a system work, the more logical it is to have the machines do it.

- News agencies are looking for opportunities to collaborate with their clients in personalization.

The global news agency Reuters has a distinctive vision of the future of human-machine collaboration with its Lynx Insight project. With substantial benefits for its financial and sports news operations, Reuters is developing a 'cybernetic newsroom'. Part of that vision are robots that are proposing stories to reporters, and the real value is created by reporters who are then able to evaluate the newsworthiness of the findings, and if needed, put them into context. The key idea of the concept: Let both machines and humans do what they are best at.

Imagine this: Manchester United beats Liverpool. The first goal was scored in the 50th minute. Right after that, there was a red card. These events are recorded as metadata of the match. At Reuters, with both a vast archive of pictures but also fresh images of the match, statistics can automatically be combined. Final product: A video-like mock-up of pictures with automatically created text explaining the key events of the match.

This is an example of what Reuters is doing with the technical capabilities already in action in its 'cybernetic newsroom'. The concept began 5 years ago with Reuters Tracer, an attempt to monitor the news that breaks on Twitter. It sounds great and the

"YOU SHOULD LET THE MACHINES DO WHAT THE MACHINES ARE GOOD AT, WHICH IS ANALYSING THE DATA AND TURNING OUT SOME FAIRLY BASIC SENTENCES." (REGINALD CHUA)

approach is sophisticated: Tracer gives tweets newsworthiness scores. The algorithm evaluates both the newsworthiness of the tweet and the credibility of the source. "The truth is that it's not in widespread use, it's still somewhat experimental," says Reginald Chua, Global Managing Editor of Operations, Reuters.

S&P 500 risers in a time period (1st, 2nd, 3rd) 1:16 PM	
Biggest monthly fallers - L Brands Inc	
L Brands Inc's <L.B.N> monthly calendar loss of -23.85% was its largest monthly fall in the last 10 years.	
Year to date performance of L Brands Inc shares: -23.4%	
Sector - Specialty Retailers	
Timestamp - 01:16 p.m. EDT (NY time)	

But there's another much more ambitious project going on at Reuters: Lynx Insight. This system combines three key functionalities: input of a dataset, an analytical layer that detects patterns in the dataset and, finally,

a natural language generator that writes news texts. Worldwide there are dozens of robots writing stories, but the Reuters' approach is different. "What we decided was that it was more important for the system to find interesting facts than it was to write a whole story," Reginald Chua says.

Lynx Insight stock trigger rising % 12:15 PM	
Kraft Heinz Co shares on 6-session falling streak	
Kraft Heinz Co <KHC.OQ>, last down 1.21% to \$ 25.21, falls for sixth trading day.	
Kraft Heinz Co share performance since end of last week: -4.9%. Kraft Heinz Co next reports earnings on October 30.	
Sector - Food & Tobacco	

At present, Lynx Insight is mostly based on rules coded by the Reuters team in Toronto and analyses archived news stories to determine the features that made them headline news. What's the typical threshold of change in the share price to make it news? For how many days does the stock have to decline for the news to become worth covering?

Reuters is shifting to a more 'AI-based' anomaly detection system that learns from the feedback it receives from journalists (concerning the recommendations that didn't work) to improve performance and become better at recommending.

Based on this, the computer might be asking questions that a reporter didn't even know they should ask. Easier said than done, says Chua, but that's where they are heading.

Reuters' thinking goes to the very heart of sensible investment decisions in AI. "Obviously, breadth is useful. And obviously, efficiency is useful. But the really interesting innovation is here: Let's not try to write the whole story if we don't have to since the amount of effort needed to make a story completely publishable is very high compared to your return on it. You should let the machines do what the machines are good at, which is analysing the data and turning out some fairly basic sentences. And you should let humans do what humans are good at, which is calling somebody up, getting a little bit more insight or asking the right questions."

Reuters knows by experience that building AI systems is expensive and hard work, so it has had to be selective. "There are off-the-shelf resources but you're going to have to customize it yourself. That's a big investment for any news organization. The second issue is that it takes effort to ingest a new dataset. There's always some work involved."

The conclusion is that AI is a sensible investment when investments can be scaled. And lots of data doesn't have that scale, in reality. For example, crime or public finance data is different in various countries, and some of those stories are news only once a year. Not good cases to invest in.

Naturally, news agencies have an advantage as they are doing the job for many news organizations. However, they don't have access to user data, as that often belongs to their clients. Reginald Chua has a vision for a project that would need collaboration with the media clients that have the information of single users. "What if we personalized stock market news around your portfolio? What if we could – and we can with Lynx Insight – create a Premier League video based on which team you support, of a length that we know suits you?"

"I think we have to understand that news is not going to be a one-to-many but more of a one-to-one process. Building a personal relationship with readers is, I think, going to be one of the competitive advantages that news organizations will have in the future. My goal is that we can help them do that."

FOR MORE INFORMATION

<https://www.reuters.com/article/rpb-cyber/the-cybernetic-newsroom-horses-and-cars-idUSKCNIG00Z0>



TALPA NETWORKS, THE NETHERLANDS

VOETBALTV, BROADCASTING WITHOUT THE HUMAN TOUCH



Martijn van Balken,
Director and COO, VoetbalTV
info@voetbaltv.nl

TAKEAWAYS

- Newest high-definition broadcast technology combined with machine learning and digital distribution may open up opportunities for disruptive innovations, like producing live 'multi-camera', fully automated broadcasts from events other than sport.

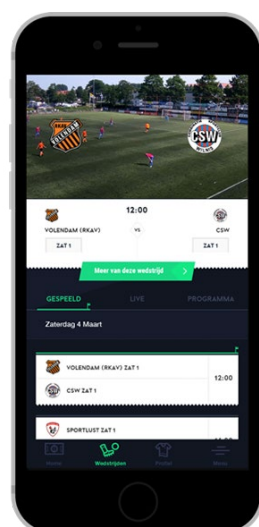
- Through these innovations, the operating costs may become so low that private-sector companies are able to provide the societal services that public broadcasters cannot afford through ageing technologies.

- The road from great idea to tangible achievement is where the battle for success is either won or lost.

Broadcasting live all the tens of thousands of amateur football matches played in the Netherlands, all fully automated and with a machine-learning algorithm running the broadcasts: that's VoetbalTV. It is today one of the most stand-out examples of what can be achieved by automating the entire production pipeline.

Eighty-thousand amateur football matches broadcasted live yearly. No people involved. Can it be true? Almost. Dutch football association KNVB and Talpa Networks have partnered in the Netherlands with this aim. The service has only been in production for around a year and is currently in ramp-up phase.

"WE WERE A LITTLE SURPRISED OURSELVES THAT IT WORKED SO WELL ON SUCH A LARGE SCALE" (HYLKE SPRANGERS)



The technical concept itself is already up and running in several stadiums. It's based on two 4 k high-definition cameras, one on the left, one on the right side of the field. These signals are transferred via fibre to a cloud server, which combines the two signals and creates one optimized wide-angle picture of the field.

The software detects the ball and the players in the picture and zooms in to where the action is. And since the cameras are extremely sharp, the 'virtual director', a computer relying on machine-learning algorithms, is able to produce a full match broadcast featuring multiple cameras angles – without a human touch.

"We were a little surprised ourselves that it worked so well on such a large scale," says Hylke Sprangers, Talpa CTO.

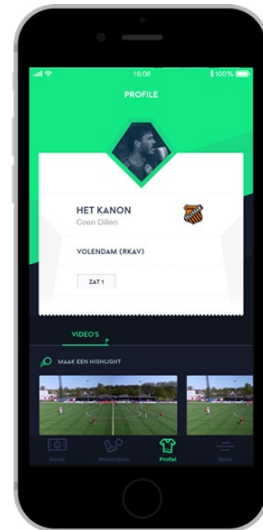
Some 200 clubs of a total of 2500 have joined the venture. About 400,000 people have already watched the games.

Talpa estimates that, by the end of 2019, 25,000-30,000 matches will have been broadcast.

“The objective was always to increase the joy of football and to increase the skills of the trade by making analysis easier,” says Martijn van Balken, one of the founding fathers of the system and director of VoetbalTV. “But, of course, this is a very nice project, and there’s a commercial side as well.”

“So far, we’re satisfied. But our ambition is always to go higher, so never sit back and say we’re satisfied!” says van Balken. The Netherlands, with 17 million inhabitants, is a country full of football enthusiasts.

Making the system even easier is that no human involvement is needed once set up – including starting or ending broadcasts. Data on games and kick-off times are imported from the national football association’s database, and this information is used to manage when the signal airs. New features powered by machine learning are already on the drawing board: producing match highlights, summaries of key events and full player statistics. Tests have shown that computers can detect these details from match footage. For sport news organisations, that kind of functionalities would open interesting opportunities.



The broadcasts are available live on mobile devices using an app called ‘social’ – because of its features allowing users to create their own highlights and share moments with teammates and friends. In addition to catering to football fans, the system also serves teams and coaches by enabling them to review their own on-field performances. One of the main reasons why the KNBV is supporting the venture is that it wants to strengthen the football culture and teams’ ability to analyse their games.

The business model of VoetbalTV is based on long-term sponsorships. While the app is free to use, the clubs pay a yearly fee to use the system including analysis product, but the revenue from this is so small that sponsorships are in effect being extended to the clubs to complete the business model.

The technical concept took less than one year to develop before general release. Three different technical and ML teams from two cities were involved, and the entire AI engine was created by Talpa’s internal data-science and machine-learning experts, in accordance with company policy.

But importantly, while the production of footage is fully automated, 20 people are involved to make it happen. After all, cameras have to be installed in new stadiums, the system has to be monitored and updated, new features developed and connections to customers and clubs maintained. And then, something unexpected may always happen, which is when people need to intervene.

According to Talpa CTO Hylke Sprangers, the key to success in the project has been a combination of a great idea and lofty ambitions + capable people. Talpa’s common technical platform was a good base to build on. It was important to keep in mind that technology should serve both users: the audience and the teams, adds VoetbalTV COO Martijn van Balken

The aim is to expand the concept internationally. This is a credible goal for VoetbalTV as Talpa is a commercially successful media company with substantial cash resources, owned by John de Mol, founder of the Big Brother and The Voice television formats.

“We have a long-term history of developing popular concepts and products here in the Netherlands and then scaling them out to other countries,” says Hylke Sprangers.

FOR MORE INFORMATION

<https://www.voetbaltv.nl/>

SIMILAR CASES

[Uefa and EVS paper for IBC: AI for better storytelling in football](#)



ECHOBX, THE UK AUTOMATING SOCIAL-MEDIA PUBLISHING

Joe Rudkin
Head of marketing, Echobox
joe@echobox.com

TAKEAWAYS

- Social media publishing is a typical area in which computers are able to predict and optimize the selection of channels and publication times better than humans since there is such a high number of variables that have to be taken into account.

- Data collection is enhanced if media organizations work together. This may take the form of direct partnerships but may also encompass external partners, who then collect the business data.

- Media companies should also look at products and services developed for the wider corporate market, as they frequently offer the same features and use-cases as required by media companies.

There are many social-media platforms using algorithms to tailor the content shown to users. Such algorithms are regularly updated, and each update has the potential to massively impact the traffic received by those news publishers which use social media to promote their content. Automation may offer solutions capable of mastering this complexity, which is most of the time beyond what humans can manage. Specialized companies drawing data from hundreds of publishers find it easier to mitigate the impact of these updates.

In January 2018, Facebook changed its News Feed algorithm to prioritize content from 'friends, family and groups' – and shocked the world of news publishers. Some publishers saw a 60% drop in their traffic from Facebook. The decision also made many news outlets far more reluctant to rely on Facebook as a strategic distribution partner.

Engagement and traffic have since recovered close to levels seen before the 'switch', perhaps because of further changes made to the algorithm, or potentially because news content is now being shared by these 'family, friends and groups' as well. But having learnt these lessons, more and more news organizations are adopting a more opportunistic strategy in relation to social media, using it, of course, to promote content but with the ultimate target to establish direct relationships with users. Knowing how the algorithms underlying social media work is essential for them to be handled effectively. However, it is practically impossible for a human editor to understand and predict these algorithms on their own. That's where full or partial automation enters as a possible solution.

“ANY MACHINE-LEARNING ALGORITHM IS ONLY EVER AS GOOD AS THE DATA IT DRAWS UPON.” (ANTOINE AMANN)

In a way, there's a war of algorithms going on. Social-media companies such as Facebook do not publish the structure of their algorithms, making it impossible for publishers to optimize their social media presence. Instead, another solution has emerged: using machine-learning techniques to reverse-engineer those algorithms and optimize publishing to social media accordingly.

For this purpose, Echobox harnesses the analytics and social media data of 600 publishers that use its services.



Antoine Amann, the founder and CEO of Echobox, cites “peace of mind” as an important reason to use the service. “Our clients enjoy no longer having to worry about when and how Facebook will change their Newsfeed. We do it for them. We are constantly monitoring Facebook’s News Feed algorithm.”

Echobox can use much wider datasets than any single media company. It models the algorithms of Instagram, LinkedIn and Twitter in a similar way. It uses machine-learning predictions about what content will be successful, on which channel and at which time. The company claims to be able to predict virality, suggest relevant hashtags, optimize the timing of posts and spot content that is worth republishing at other times.

With natural-language processing technology, Echobox generates ‘share’ messages automatically. It also creates mathematical representations of each media customers’ audience using its data.

Social-media automation has been used in corporate communications for some time, and many of the technologies developed in that field could work for news distribution as well. However, only some of them use machine-learning techniques. Several US news organizations and Reuters, for example, are using True Anthem, which was originally created for marketing. There are plenty of technologies available but these are mostly offered to corporate communicators. That reveals where the world is heading: more and more content consumed within social media is coming from sources other than traditional or independent media.

Antoine Amann underlines the role of data acquired from media clients in developing services like his own: “Any machine-learning algorithm is only ever as good as the data it draws upon. Even if you have the most advanced machine-learning algorithms, if you don’t have good data sets, you can’t really do much with them.”

The company analyses the historical data of every client and trains algorithms individually for that client, based on its specified targets and editorial policy. Audience behaviour is, of course, different for each news brand. Sometimes the company uses insights into the news market in general, but typically the work has to be performed on a client-by-client basis to reach the desired objectives. “In the UK, The Times has a very different audience to that of a tabloid. That’s why merging data together typically doesn’t yield good results in most cases,” said Amann.

One exception is breaking news. Publication time can’t be optimized, as breaking news articles have to be published immediately. It is crucial for publishers to be first when publishing that type of news. That’s why using AI technology to detect such items within split seconds is another worthwhile application to drive publishers’ success on social media platforms.

Amann distinguishes strategic decision-making from executional automation in the newsroom. “It’s important for the editorial team or publisher to make the strategic decisions as to how they want to publish on a daily basis. However, the execution itself can be fully automated today. In ten years’ time, I believe strategic decision-making will be fully automated as well.”

FOR MORE INFORMATION

<https://www.echobox.com/>

SIMILAR CASES

[True Anthem, a social media marketing automation tool used by Reuters, Hearst and NBC Universal](#)

[Socialflow, used for social media marketing by many large US media brands](#)

HOW TO

AI FOR NEWSROOMS

There are many tools that journalists can use in the newsroom that harness some kind of machine learning. Some of them can be very helpful in making the journalistic production process quicker and less repetitive. We've drafted a list of AI tools drawn from various sources and grouped them into the six phases of the production process. The diagram below shows the six phases and their possible applications.

PLEASE NOTE:

- The list is by no means complete.
- Most tools on the list are commercial tools that can be purchased for use or are free for use.
- A lot of tools in the list work mainly in English.

IDEA

- Detect trends on social media with Newswhip: <https://www.newswhip.com>
- Get tips for breaking trends with Dataminr: <https://www.dataminr.com>
- Get story angles, visuals and people suggested through an online search with Project Inject: <http://injectproject.eu>
- Browse the archive of investigative platform OCCRP with OCCRP Aleph: <https://aleph.occrp.org>

RESEARCH

- Check for facts with Claimbuster: <https://idir.uta.edu/claimbuster/>
- Check for fake photo and video with:
 - Truepic: <https://truepic.com>
 - Serelay: <https://www.serelay.com>
 - WeVerify: <https://weverify.eu>
 - InVID: <http://invid-project.eu>
- Do a reverse photo search with:
 - TinEye: <https://www.tineye.com/>
 - Google reverse image search: <https://images.google.com/>

- Check for deepfakes with:
 - Deeptrace: <https://www.deeptracelabs.com>
 - Deepnews.ai: <https://www.deepnews.ai>
 - Truly Media: <http://www.truly.media>
- Find email addresses connected to websites with SNOVio: <http://www.snov.io>
- Implement facial recognition in your data with Clarifai: <https://www.clarifai.com>
- Discover bots on social media with:
 - Botslayer: <https://osome.iuni.iu.edu/tools/botslayer/>
 - Botometer: <https://botometer.iuni.iu.edu/#/>
- Detect if texts could be real or fake (written by an algorithm) with Giant Language Model Test Room: <http://gltr.io/dist/index.html>
- Recognise objects in photo's (open source) with Tensorflow Object Detection: <http://www.tensorflow.org>

PRODUCTION

- Get your audio transcribed through:
 - Trint: <https://trint.com/>
 - Otter: <https://otter.ai/>
 - Happy Scribe: <https://www.happyscribe.co/>
 - Descript: <https://www.descript.com/> (has special features for audio, video and podcasting)
 - Sonix: <https://sonix.ai/> (also video to text)
 - Temi: <http://www.temi.com>
 - Google Live Transcribe: <https://play.google.com/store/apps/details?id=com.google.audio.hearing.visualization.accessibility.scribe&hl=nl>
 - AI Media TV (also automatic subtitles): <https://www.ai-media.tv>
 - Meeting to Docs: <https://meetingtodocs.com/>
 - Google recorder: <https://www.techradar.com/news/google-recorder-is-a-real-time-transcription-app-that-even-works-when-youre-offline>
- Make stories out of data with Quill: <https://narrativescience.com>
- Automate data visualisations with:
 - Graphic: <https://www.graphiq.design>
 - Tableau Public: <https://public.tableau.com/s/>
- Have stories automatically written (natural language processing) by:
 - Wordsmith: <https://automatedinsights.com/wordsmith/>
 - ReporterMate: <https://github.com/nickjevershed/Reportermate-Lib>

- Have texts automatically translated with:
 - Mediawen: <https://mediawen.com>
 - Zoi Meet: <http://www.zoimeet.com> (transcription as well as translation)
- Get automatic text corrections with:
 - Grammarly: <http://www.grammarly.com>
 - Hemingway: <http://www.hemingwayapp.com>
- Get AI assistance in video editing with:
 - Adobe: <https://www.adobe.com/nl/products/premiere-elements.html>
 - Wibbitz: <http://www.wibbitz.com>
- Filter noise out of audio: <http://www.krisp.ai>
- Automatic edit audio with Descript: <http://descript.com>
- Automate workflow yourself with If This Then That: <https://ifttt.com>

PUBLICATION

Automatic personalisation with:

Amazon: <https://aws.amazon.com/personalize/>

Gingalabs: <https://gingalab.com> (videos)

Vionlabs: <https://vionlabs.com>

FEEDBACK

- Improve the comments on your content with Perspective API: <https://www.perspectiveapi.com/#/home>
- Leverage predictive analytics to drive revenue and increase customer satisfaction with Prevision: <https://prevision.io/home/en>
- Make your own chatbot with Flow: <http://www.flow.ai>
- Bring AI to Slack, the collaboration software: <http://www.botz.ar.com>

ARCHIVE

- Clean up messy data with Open Refine: <http://openrefine.org/> (works e.g. in English, Spanish, French, Russian, Portuguese, German, Italian, Hungarian)

CHAPTER 03

DON'T THINK YOU KNOW THE AUDIENCE



BECOMING DATA-INFORMED IS THE FIRST STEP TO RESPONDING TO THE NEW REALITY

In a world where humans are looking for a meaningful way to cooperate with machines, data is becoming an increasingly important way to give guidance about what should be done and what should not. And even then, data is only valuable if it's understood and gives a clear indicator of how behaviour can be changed.

Setting the goal of becoming a data-informed newsroom is the first step towards breeding a culture of self-reflection based on deeper understanding of audiences. This is vital to succeeding in a digital environment. For public service, there's a clear need to move from volume metrics towards measuring relevance, engagement and impact.

On the surface, it seems that being data-informed has become the standard for news organizations. You can hardly enter a newsroom and not find a big screen that gives real-time information on how online stories are performing.

But are these metrics on screens giving us real information on our audiences? Mika Rahkonen, Head of Strategy at Yle, isn't so sure: "On digital channels like the web, new holes from the metrics are discovered pretty much every month. 'This didn't work. Oh yeah, that wasn't correct.'" In short, he argues that the metrics we use consistently overestimate consumption and relevance. "The media think that people love the media much more than they actually do. And this, I think, is a pretty big deal."

"PERHAPS ONE OF THE IRONIES OF THE DIGITAL ERA MAY BE THAT THE MOST PERSISTENT AND VIVID READER PERCEPTIONS ARE STILL BASED ON ACTUAL PERSONAL CONTACT, A FUNDAMENTAL HUMAN CONNECTION THAT VIRTUAL COMMUNICATION — BE IT THROUGH NUMBERS, GRAPHS OR EVEN EMAIL — STRUGGLES TO REPLICATE." (JAMES ROBINSON, PROFESSOR AND RESEARCH FELLOW AT COLUMBIA UNIVERSITY'S SCHOOL OF JOURNALISM AND FORMER DIRECTOR OF NEWS ANALYTICS AT THE NEW YORK TIMES)

Most journalists have entered the profession because they want to be relevant and make a difference, and the value of their work lies in the impact they can make with their audience. But often the metrics shown on newsroom monitors do not reflect this impact. And they are hard to act on. Bård Romstad, Head of Product Development for Data and Insight of NRK, says: “Often metrics only indicate past results, not the possibilities for future success.” As a result, metrics are quite often not believed or accepted by journalists in newsrooms.

So the metrics need to change if they are to be truly valuable. Previously we were dependent on unreliable, analogue-based measuring systems to keep track of the behaviour and needs of our audiences. Nowadays, new systems enable us to track every move a browser is commanded to make. By formulating the right goals and connecting them to new ways of measuring success, we can obtain more answers.

But this does require something of newsrooms, of journalists. As we concluded in the [2018 News Report](#), journalists are not always prone to self-reflection and find it sometimes hard to accept that they need to learn continuously from feedback. Journalists must become more comfortable with the idea of challenging their previous assumptions so that they can fully harness the feedback they obtain from metrics. Adopting a more curious, fact-based and data-savvy culture is a foundation for any serious attempt to serve the audience better and succeed in the next digital wave.

FROM VOLUME TO RELEVANCE/ENGAGEMENT

Newsrooms often have a vague or even wrong image of our audience. But why? First of all, public service news has in general been focusing on vague volume and reach metrics for far too long. Public service has considered ‘reach’ as the main indicator of success. This is partly because charters and laws emphasize the universality of PSM services. The target group is ‘everyone’, and therefore you should reach huge numbers of people. This has in some cases even led to an unhealthy relationship with this metric, to the extent that the public value of the content is not fully considered. “Public service seems to be addicted to large numbers, and this is odd,” says Ben Fields, Lead Data Scientist at BBC News. “Reach does not equal love, caring, or even liking,” adds Mika Rahkonen of Yle.

Even in privately-owned media, the focus on clicks (and advertising) has shifted towards creating a deeper relationship with the audience (and subscriptions).¹ In this respect, these companies have acknowledged that the success of digital stories can’t be measured only by the number of people that have clicked on an article. Rather, the correlation between time spent, recirculation, retention and loyalty is key. Should we then combine different kinds of metrics? Jill Nicholson of Chartbeat thinks so: “If you don’t balance reach with quality, you run the risk of grabbing people only when there’s something viral going on. The job is not to chase traffic. In the business of news, random indiscriminate traffic is not what a business is built upon.” Ideally, a combination of metrics should be used, each connecting to a different goal – a different part of the overall success story. Metric company Parse.ly points to three areas on which newsrooms ought to focus: distribution (how are we growing?), connection (how are readers responding?) and impact (do readers want more of what we are creating?). For each area, different metrics play a role: page views, unique visitors, shares, referrals.

Conversely, Dejan Nikolic of Content Insights is critical of most of these metrics: “The problem is that the metrics that we are used to, like page views, unique visitors, new visitors, return visitors; that’s all browser behaviour, not reader behaviour. It tells you nothing about how your content is being consumed.” So you may well measure them, but you wouldn’t know if you have been of value to the user.

In our checklist A way to design metrics, we list the various approaches used for metrics, listing multiple sets of indicators that help ensure a successful outcome.

ARE THE METRICS BEING USED AT ALL?

Another complex factor with metrics lies in applying insights. Do metrics actually give journalists more insights? And do journalists act on them? Even if they understand the message, are they willing to change the content they are producing?

While there are conflicting views on the issue and things are already changing,² research by the Tow Center for Digital Journalism argues the rise of available metrics has not really altered journalists' behaviour: "While journalists are open to engaging with readers, the ways in which they form audience perceptions remain largely unchanged despite the rise of audience metrics and analytics."³ There are several possible reasons for this:

1. The data never really reaches the journalists

Most of the analysis is taking place elsewhere, in audience research or in departments handling distribution data, so the insights are not reaching journalists in newsrooms. A recent study by The Lenfest Institute and Harvard's Shorenstein Center found that "... the advertising people look at the ad metrics, the subscription people look at subscription metrics, and journalists look at outdated metrics such as how often their articles made the front page of the newspaper. Nobody was looking at these metrics together to get a clear view of the business as a whole."⁴ Production processes are often so tightly packed that there is hardly any time to look at metrics – new stories need to be made. However, a recent study suggests that things are changing fast: in 2019, 44% of journalists regularly consulted analytics, compared to 29% in 2017.⁵

2. Journalists don't want the data in the first place

Even when the audience data reaches the right people in newsrooms, it's not to say that every journalist welcomes these insights. One out of five journalists never consults analytics.⁶ Professor Media Studies at the University of Amsterdam, Mark Deuze: "There is a classic tri-division between 1) a small group of enthusiasts for measuring audience data, 2) a fairly large group of people who don't like it at all and say something like 'this has nothing to do with real journalism, I just want to make beautiful pieces, or reports' and 3) a very large group that is not involved in this at all; not so much because they don't believe in it or something, but because they are already busy." And even when they're not busy, they sometimes simply don't want the data. As research by Jacob L. Nelson and Edson Tandoc concludes: "Journalists face two goals they perceive as mutually exclusive: the pursuit of a mass audience and the aspiration to provide mission-driven reporting."⁷

3. Journalists don't understand the data/metrics

Then, when data does reach the right people, it's not always properly interpreted. Journalists are no data scientists. In general, they don't like maths and don't get on well with numbers. Receiving audience data might be more confusing than enlightening. David Bunker of the BBC: "I sometimes wonder if we don't measure too much. It's confusing for the people making the content." Some people therefore suggest that metrics should be presented to journalists much more simply if they are to be useful. As Eric van Heeswijk, CEO of Dutch analytics company Smart Octo puts it: "If you don't bring the data in a very useful, simple, friendly form to the newsroom, all of this intelligence will be lost."

4. Journalists don't act on the data

Journalism is often considered to be a creative profession. To be focusing on data and optimization might be in direct contrast to this and push the wrong buttons.⁸ Whereas for organizations it might be logical to focus on optimization, for individual journalists the opposite might be true. It might be more important for them to spend time understanding the real story and getting their version right.

2 (Ferrer-Conill & Tandoc, 2018; Hanusch & Tandoc, 2019; Lamot & Paulussen, 2019; Owen, 2019; Petre, 2018)

3 (Robinson, 2019)

4 (Mele, Skibinski, & Spector, n.d.)

5 (Owen, 2019)

6 (Owen, 2019)

7 (Nelson & Tandoc, 2019)

8 (Neheli, 2019; Zamith, 2018a, 2018b, 2018c)

Focusing on data and optimization can cause a considerable amount of stress for journalists. Journalists feel they are being forced to focus on metrics by their superiors. “Research shows that journalists are becoming increasingly aware that their choices have consequences for the brand and the metrics. And they involuntarily adjust themselves to them,” says Professor Mark Deuze.

“We interviewed lots of journalists who said that this was one of the reasons why they left journalistic organizations. That it was all about the clicks, or they felt that they were being forced into a sort of straitjacket where you have to do journalism in a certain way.” So ‘changing behaviour because the metrics say so’ may run counter to the role journalists want to have. That may then cause tensions or even conflicts in newsrooms.⁹

There is a way out of this, and that is by tying the metrics to concrete actions journalists can take, which are in line with their gut feeling on ‘quality journalism’. At Norwegian NRK, they are planning to do just that, explains Bård Romstad: “It’s important that the metrics journalists see are quite simple and reflect their instinct. ‘Quality’ is hard to define, but when you feel you have done a good job, and the numbers indicate the same, you trust the metric currency. And with that, you are willing to change your behaviour.”

DATA IS NOT JUST THE END BUT ALSO THE BEGINNING

One way to foster a data-informed culture in the newsroom is not to look simply at the impact of stories afterwards but rather to focus on the audience at the beginning of the journalistic production process and try to understand who is being addressed with what message and why: “As much as journalists like to talk about the five W’s of a news story—who, what, where, when and why – the practice of journalism rests on three other, equally important questions: ‘Who am I writing for? Why is it important for them to read it? And what will they find interesting?’” noted James Robinson, professor and research fellow at Columbia University’s School of Journalism and former director of news analytics at The New York Times.¹⁰

This is underlined by Head of Data at BR, Ulrike Köppen, who as a Nieman fellow at Harvard wrote about effective interdisciplinary news teams and described the ideal workflow as: “prototyping, testing, revising – repeat. It means trying out ideas by breaking them up into small portions of work, taking first steps, evaluating the results, and deciding as a team to go on or dismiss the idea. The key is responding to change.”¹¹

So true audience measurement is only accomplished in iterative processes, from real contact and openness and willingness to learn. James Robinson: “Audience knowledge is best revealed by iteration, not just experimentation. It is not enough to try new things; having the ability to observe and evaluate what worked is critical – what resonated with one’s readers, and why.”¹² A good example of how you can design this kind of journalistic iteration is the Dutch Newsmap. This project uses artificial intelligence to alert editors about the areas of the province they are covering the most, and where the audiences that consume those stories are geographically located. In this way, they can see imbalances in their coverage and discover areas that need more reporting.

The BBC even went one step further, wagering that algorithms could help predict content gaps. Magda Piatkowska, Head of Data Science at BBC News says: “My dream state, the ambition, is to have algorithmically driven commissioning. I think that with the pressure to address more audiences, with the pressure of efficiency, we just can’t afford to waste resources on content that is not relevant to audiences. And those two drivers are going to lead us to adopt more machine-driven commissioning.”

Systems with more and more predictive features are becoming available and being used. The aim is to use resources more efficiently and adjust the content if it seems that it won’t be consumed by enough people.

9 (Chadha & Wells, 2016)

10 (Robinson, 2019)

11 (Köppen, 2019)

12 (Robinson, 2019)

STEP AWAY FROM THE COMPUTER

Despite all this data, it's important to acknowledge that in order to receive relevant contextualized information, journalists still need to engage directly with the public. They need to step away from their computers and meet their audiences. James Robinson: "Perhaps one of the ironies of the digital era may be that the most persistent and vivid reader perceptions are still based on actual personal contact, a fundamental human connection that virtual communication – be it through numbers, graphs or even email – struggles to replicate."¹³

Professor Mark Deuze adds: "If you really want to know what the public finds important, then you shouldn't look at clicks, you have to talk to them. You have to build and enter into a relationship. Just look at how commercial brands work. This is not so much about just dumping stuff on the market and hoping that they will be sold. No, it is really about entering into a kind of relationship, sharing knowledge, exchanging information, collaborating at certain times and then you build something up, so that you can put products on the market. And that is something that you should be able to do very well in journalism."

In last year's News Report, we signalled that getting to know the audience offline is [crucial for building trust](#). This time round, we are adding the viewpoint that this is actually crucial for truly understanding audiences and, ultimately, for making better content. We know that, inspired by the case studies from 2018, such as [Fika by SVT](#) and [Tiomiljoner](#), other EBU newsrooms have set up projects to meet audiences offline.

RECOGNIZE BIAS IN BOTH YOUR DATA AND YOUR NEWSROOM

A pitfall in relying on offline physical contact, however, is that diversity comes into play. You need to make sure that the people you get in contact with are a proper reflection of your audience – in all its shapes and sizes. When editors rely only on their own peers and sources, i.e. the people they know, there is a risk of seeing only the obvious.

It's also true that data may have its own biased and pitfalls as well. However, real people have a much stronger impact on how editors perceive audiences. As James Robinson puts it: "Analytical tools describe numbers, not people; and artificial constructs such as personas are often ignored. Given how deeply one's peers and sources inform one's perceived readership, increased newsroom diversity might be the most effective way to ensure that the readers in one's mind's eye accurately reflect the audiences for their work."¹⁴

Diversity is another issue covered in the [EBU News Report 2018, 50 ways to make it better](#).¹⁵ We concluded that widespread trust can only be built if multiple groups recognize themselves and their lives in the content on offer. Therein lies a huge risk, namely relying on the demography of the current newsroom alone to be the mirror for the world. All initiatives that attempt to broaden perspectives are therefore to be encouraged.

This needn't always be HR projects, or projects focusing on bringing in new people. ORF has for example recently set up Reflective Practice Groups, where editors from different departments learn from each other on how to reflect on their own biases and assumptions.

FROM RELEVANCE TO IMPACT

In a further step in measuring how well end-goals are being reached, public service media is increasingly focusing on impact. The claim goes like this: it's the effect the story or investigation has on people and society that proves its true value, not only the quantity of people it reaches.

As EBU's Head of Strategy Roberto Suárez Candel puts it: "As news professionals, we now have those great tools to measure what works, what doesn't, how we distribute and how we communicate. But these are tools. These are not the goals. The goals are to raise awareness, to make people understand better, to let people undertake action as citizens." In other words, the ultimate goal is impact.

¹³ (Robinson, 2019)

¹⁴ (Robinson, 2019)

¹⁵ (Jääskeläinen & Olij, 2018)

Impact is yet another word that needs further clarification. Impact may be both positive and negative. The audience research department at the BBC dissected the term in a recent project aimed at improving the impact of their programmes. David Bunker led the project: “We needed to define what is impact, before we could include it in our measurements.”

They finally developed a typology of impact consisting of six elements: emotional release, learning, inspiration, believing, helping and interacting. “And this is all just about the impact a production has on an individual; we have left the societal impact out of this typology.”

The BBC’s typology reveals one aspect of the complexity of moving to impact as a key indicator of success: it’s really difficult to measure, and the measurement can’t easily be automated. Questionnaires or real-time interviews seem to be the only way. And even then: “It’s still difficult because a lot is perception. And if a brand has clever marketing, one’s perception might be that there is a great impact, but this might not be true,” says David Bunker.

Public service media is historically focused on providing value to society. The EBU in its Contribution to Society initiative says: “the legitimacy of PSM might well be at stake if citizens cannot perceive the full value of what public broadcasters offer them... At its core, this initiative calls on public broadcasters to identify and explain the value that citizens receive from public service media.” There is great urgency to do so, as commercial media companies are positioning themselves within the framework of impact. Impact, defined as “the ability to contribute to positive change with respect to a societal issue”, can be felt on an individual level, but also on a group level and by society as a whole.¹⁶

However, effective ways to measure this kind of impact are still rare. Roberto Suárez argues it can sometimes be simple measures: “We can go into very complex mathematical models for measurement. But we can start with very simple things like asking people whether you are really helping them.”

In Austria, ORF has created a comprehensive way to analyse the fundamental question of what really is quality and public value.

Impact might include how the media organization itself can lead change and better reflect society itself. In the UK, the Diamond project has brought together public and commercial broadcasters to measure all the characteristics of all people shown on screen or working in production of a programme, in an earnest attempt to address diversity head on.

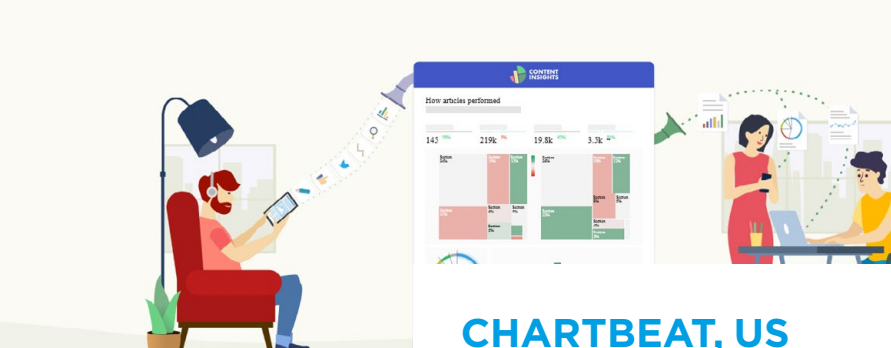
While the impact actually always should be measured from the standpoint of audiences, not based on the content itself, the project that Frederic Filloux is conducting with Deepnews.ai is interesting to regard. Here they try to measure the quality of a news story – meaning depth of reporting and investigation, ethics and resources deployed – with deep-learning models called convolutional neural networks, attaining an accuracy rate of 80-85%.¹⁷

¹⁶ (European Broadcasting Union EBU, 2019)

¹⁷ (“Deepnews.ai FAQ”)

RECOMMENDATIONS FROM THIS CHAPTER

1. Public service journalism should aim for relevance and genuine relationships with audiences. It should therefore use metrics that measure engagement, loyalty and impact – not only reach. There is a multitude of metrics available that provide far more valuable insights than reach alone.
2. Data is only valuable if it influences decision-making in a healthy way, creating a culture in which the organization tries to create a proper understanding of audiences, their expectations and behaviour. An iterative model, which consults audiences regularly and then translates insights into concrete actions, is most effective.
3. In addition to the numbers, journalists need other people for true sense-making. Stepping out of the office and getting to know real people in all their diversity will help create firmer foundation for a real relationships.
4. In the fast-developing area of measuring impact in public service, it's important to define what the desired impact is and how public service journalism contributes to the outcome.



**CHARTBEAT, US
SMART OCTO, THE NETHERLANDS
CONTENT INSIGHTS, SERBIA
UPSCORE, GERMANY**

DIFFERENT WAYS TO LOOK AT THE SAME DATA

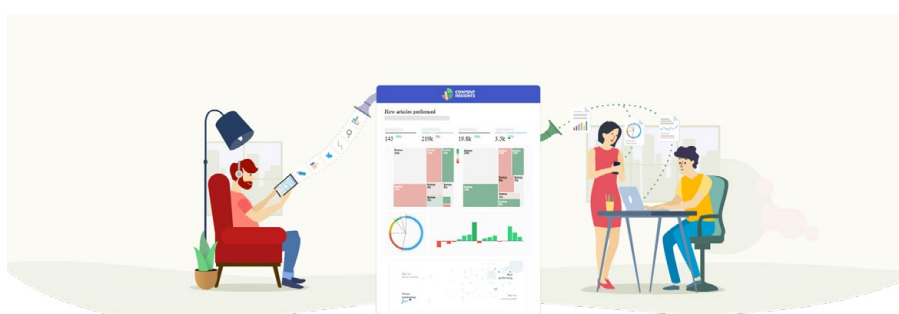
TAKEAWAYS

- To define success metrics for your organization, you have to know what you are aiming for. Metrics should reflect that decision and be understandable and actionable: How do I change my story so that the metric rises?

- Third-party tools have different approaches to what is important. Choice of tool may give you opportunities but also limit your criteria.

- PSM organizations should carefully consider moving from quantitative to qualitative indicators because the latter more accurately reflect the actual relevance of a service to the audience.

To make use of audience data, it has to be made meaningful and capable of being used to improve performance. Third-party data analytics companies have different views on which data is relevant, what should be monitored and how actionable insights should be communicated to newsrooms. The common message from all these companies is that 'reach' is overvalued as a metric.



"Are the people you are attracting to your site ones likely to become more loyal to the publication, so you're kind of focusing on building the right audience instead of just going for overall reach?" asks Jill Nicholson, Director of Customer Education from Chartbeat – thereby formulating the key question for privately owned media. "It's not only are they opening the page, but they are reading and interacting with the content."

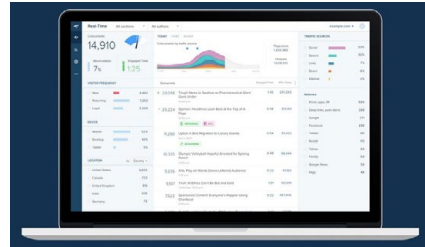
"THERE'S NO SYSTEM IN THE MARKET THAT HANDLES PUBLIC VALUE VERY WELL." (ERIC VAN HEESWIJK).

For public service media, selecting only audience niches is not appropriate, since at the heart of public service mission is the idea of bringing people together. There's no right or wrong audience.

However, the question of right metrics and the best tracking tools is crucial for public service too, if turned back to front, namely: How can a loyal relationship be established with those visiting your site? How can analytics tools be used to understand this relationship better and act upon the findings?

Many PSM organizations, such as NRK of Norway and Yle of Finland, have built internal dashboards to track their chosen metrics, inspired by examples like The Guardian's Ophan or The Financial Times' Lantern. Most companies are, however, using standard third-party tools.

Chartbeat, the dominant player, is facing a challenge in Europe from small companies trying to differentiate themselves through innovation. These companies have differing views on how to measure success and impact newsroom culture.



Mario Maracic, Chief of Product and Technology at UpScore, a Hamburg-based analytics company, believes in combining metrics into a single score, a so-called 'content score'. UpScore wants the tool to be intuitive and easy to understand. The basic premise is that success is not measured the same way at every organization. Each newsroom can and should define its own metrics of success.

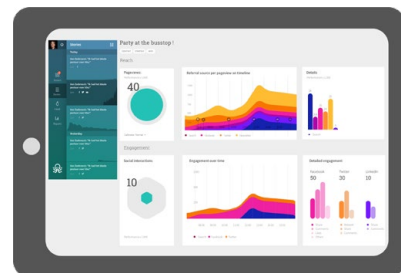
Smart Octo, a Dutch analytics company, has a completely different approach. "Basically, I believe that there is no single success metric," says Eric van Heeswijk, CEO. "You have to have articles that give you reach, articles that give you conversion, and articles that give you engagements. And you have to have different success metrics for all these different kinds of articles," he says.

"And I think basically, every newsroom that has only one set of metrics to define success is going to die. I think it's actually damaging to journalism to go to a zero-sum game with one metric."

The distinctive factor in Smart Octo's method is the use of a visualization tool that analyses all the distribution channels and which transmits this to the author of the story. Smart Octo has built tens of algorithms into the system, combining data from different sources, and giving clear tips for actions for editors, like: "Make the teaser better" (seems to be a good article on your own page but no action in Facebook).

Chartbeat gives priority to engaged time and recirculation as metrics. Engagement indicates loyalty, says Chartbeat, and recirculation helps create a long-lasting relationship with the reader.

"The problem with only focusing on reach is, you kind of have to start from scratch and acquire people over and over again. Whereas if you're focused on engagement, and loyalty, you're building a deep enough relationship with those readers, that they're consistently coming back to the site," says Jill Nicholson. She says that loyalty gives credit to quality journalism, with investigative pieces and properly reported and well-told articles rising to the top of the rankings.



Dejan Nikolic, the editor of Serbian website Njuz.net, developed the core engine that became later the asset of another analytics company, Content Insights. It was originally designed to calculate freelancer fees based on the actual value that their stories brought to his site.

He claims that the key to measuring behaviour is not found in simple metrics, but in calculations and inter-relationships, because that's where behaviour is being revealed. He also states that the success metrics have to be tailored to the business objectives of each site, and their single score is called the Content Performance Indicator.

He says the insights have to be actionable, and the road to action passes through understanding. He says most people in newsrooms read the reports served to them by analysts or analytical tools but still make their decisions based on gut feeling.

Chartbeat reports in its case study that just looking at scroll-depth data resulted in its client Swiss newspaper NZZ throwing long-term assumptions out of the window. Everything from sentence length to info-box placement was questioned. Best practices on image content and article views on mobile were disrupted.

Let's return to the question of how best to measure success in public service, and how to use external tools in that pursuit.

For example, if you can adapt UpScore's Content Score to each news organization's objectives, it can be deemed suitable up to a certain point. You can define your own Key Performance Indicators and calculate your own performance based on those indicators. The system will then compare the readings with your own historical performance.

More and more, analytics tools are adaptable for each customer. Choosing the right metrics and making them actionable is a challenging task for public service media, whose goals are not market driven.

What about the special objectives tied to public service journalism? "There's no system in the market that handles public value very well. There's more of a qualitative measurement system behind it," says SmartOcto's Eric van Heeswijk.

LINKS TO WEBSITES

<https://smartocto.com/>

<https://contentinsights.com/>

<https://www.upscore.com/en/>

<https://chartbeat.com/>

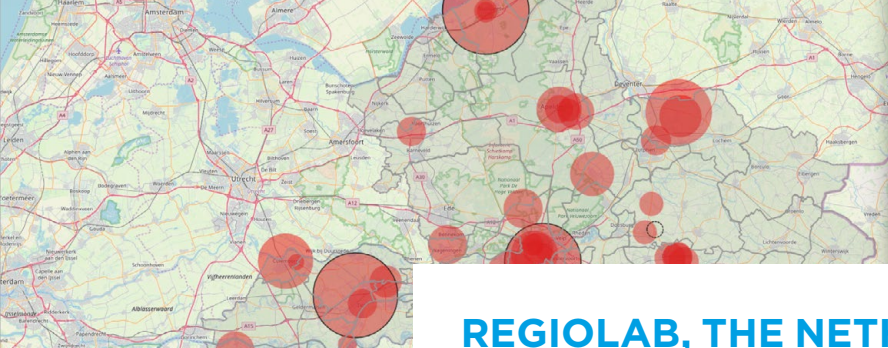
OTHER TOOLS TO MONITOR DATA:

<https://www.parse.ly/>

<https://www.newswhip.com>

<https://www.adobe.com/analytics/adobe-analytics.html>

<https://www.native.ai>



REGIOLAB, THE NETHERLANDS NEWSMAP



Just Vervaart
Projectmanager Regiolab at Omroep Gelderland
jvervaart@gld.nl, 0031-651492247

TAKEAWAYS

- Public service goals can be also attached to content production (where is this news located?), not just content consumption.

- Collaboration between various news outlets makes the product more valuable, as people's interests are not restricted to provincial boundaries.

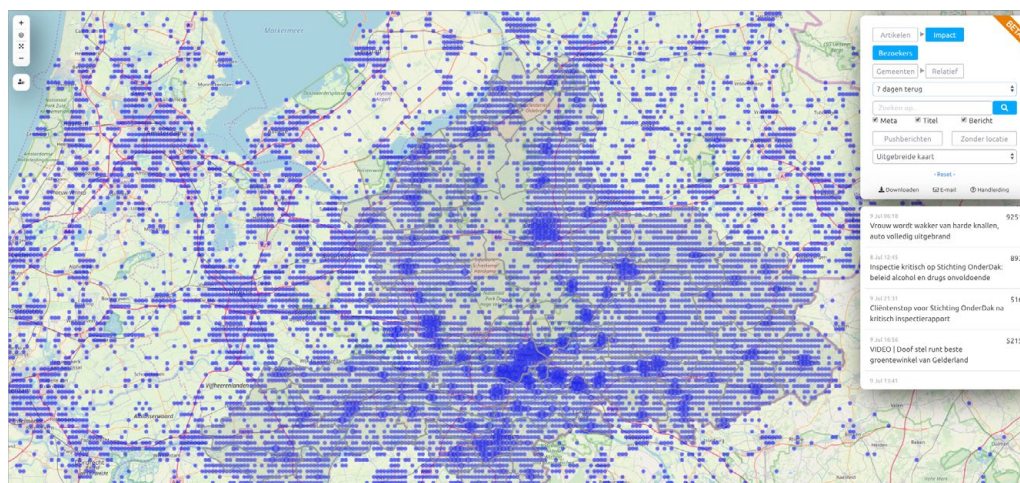
- Application of the tool by newsrooms will eventually prove how successful it is. Will editors indeed focus on the blank spaces that the Newsmap reveals?

The Newsmap uses big data to produce heatmaps of local news coverage. Which municipalities have been over- or under-represented in the news? And where are those consuming the news located?

The publications of regional broadcaster RTV Oost seemed to be more popular in the East of the Dutch province Gelderland than in the West; there were more visitors coming from that area. The assumption in the newsroom was that the brand was somehow favouring the East in its stories. "Everybody felt that the cities in the East were more prominent in the news than those in the West. But we now have evidence that proves the contrary," says Just Vervaart, Manager at Regiolab: a collaboration of nine regional Dutch broadcasters working on (journalistic) innovation.

The proof comes from Newsmap, an interactive analytical tool that compares geolocation data of articles with that of the news consumers. In this particular case, the Newsmap showed that there was a fair balance in the stories from both parts of the province. "The explanation for the difference in reach lies in promotion: there is more competition in the East." So instead of focussing on more stories from the West, Omroep Gelderland now knows it's probably wiser to spend more energy on marketing efforts.

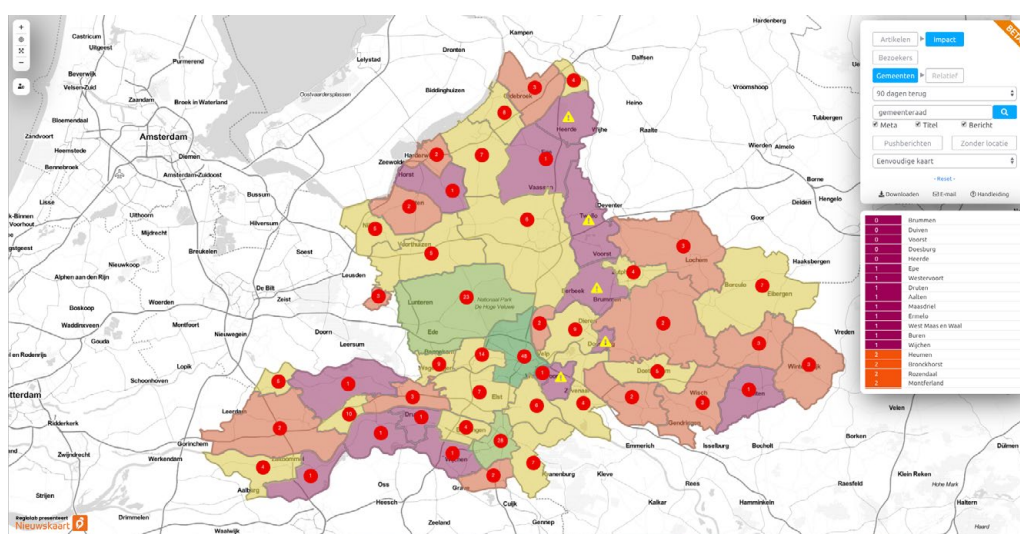
The Newsmap shows a map of the region in which news articles appear as red circles and consumers are represented as blue dots. Users can analyse the geographical impact of single news stories, by clicking the red circle and checking where the readers of that article come from. But the real interesting analyses can be made with the absolute numbers of articles per area, for example on a topic ('local council') and/or a specific time frame. This will show how well each area of the province has been covered, marking uncovered areas with an alarm colour. Editors can use this information to invest extra attention to this area, or at least check why there have not been articles so far.



“THIS IS SLOW ANALYTICS, WITH NO REAL-TIME QUERIES. IT’S NOT ABOUT OPTIMIZING REAL-TIME REACH BUT ACHIEVING GEOGRAPHIC SPREAD IN TERMS OF ARTICLES AND READERSHIP.” (JUST VERVAART)

“Our map combines data on news coverage with data on news consumption. On the production side we tag all of our reports with a geolocation. This is carried out using the regular CMS that the regional newsrooms are all working with. On the consumption side we have a representative sample of about 20% of our audience base that shares their location data with us. This gives us insights into where stories are consumed. When you combine the two, you can make all sorts of relevant analysis,” explains Vervaart. In order to highlight the most important insights, users receive a daily email showing scores in relation to newsroom-specific targets.

The tool has proven successful in challenging newsroom assumptions and biases. One example is that people’s interests are not restricted to the boundaries of provinces.



“The Dutch regional broadcasters are connected to the constitutional definition of provinces, but people’s interests are more likely to focus on a radius of a few kilometres around their town,” explains Vervaart. With Newsmap, Regiolab now has proof of that. Newsmap has also triggered more cross-border cooperation and exchanges with local news sources, as their content can be more easily distributed to relevant audiences.

“It all started with this image me and my colleagues saw of Sweden; a heatmap on which they had plotted news stories based on their geolocation data. We figured we had to develop something similar for the local news coverage in the Netherlands, to track how well we are doing in our mission ‘to serve everyone’ with our regional news. As an added value, our news brands will become more relevant for people, which will improve our reach and ultimately, our impact.”

The map cannot be used in real time as the data is processed overnight. And while this started out as a technical limitation (the rendering simply takes too much time), there is also an editorial rationale behind it. “This is more like slow analytics, with no real-time queries. It’s not about optimising real-time reach, it’s about achieving geographic spread in terms of articles and readership. And this can be done better using data that has been collected over a period of time.”

In the future, Regiolab hopes to combine the findings of Newsmap with its projects on personal recommendations.



ORF, AUSTRIA

REFLECTIVE PRACTICE GROUPS

Christian Schueller
christian.schueller@orf.at
 +4369917115145

TAKEAWAYS

- You need to truly understand your own position and assumptions before you can detect biases in AI or instruct artificial intelligence correctly.
- Bringing in psychologists can help journalists overcome their reluctance to reflect on themselves.
- Interdisciplinary and diverse teams that acknowledge each member's professionalism can operate as self-regulatory systems on ethics.

With the Reflective Practice Groups, ORF invites editors to openly and critically reflect on their own judgements. By making their thoughts explicit they learn to become more sensitive to the people around them and more aware of their (routine) work processes. As we need to truly understand our own position and assumptions before we can instruct artificial intelligence correctly, Reflective Practice Groups could be a good starting point for an organization that wants to focus on AI.

If you want to implement artificial intelligence in your newsroom, you need to have a clear idea of who you are, what you are doing and what effect you are hoping to achieve. You need to make assumptions, gut feelings and biases explicit. But this is not always easy for journalists, noticed self-proclaimed 'dinosaur journalist' Christian Schueller at Austria's ORF: "My experience is that journalists are very good at talking about the outside world, but they often find it difficult to listen to each other carefully".

Almost three years ago he decided to set up Reflective Practice Groups. The idea was to provide a secure space where journalists could share specific work experiences with others who are actively listening. The group reflects on what is said, what is left out and asks questions. As the different participants contribute, various perspectives come into play, providing a multi-dimensional picture that portrays unspoken assumptions and routines.

"AS JOURNALISTS WE ARE USED TO FOCUS ON CONFLICTS OUTSIDE OF US, THUS OFTEN AVOIDING REFLECTION ON OURSELVES. BUT IF WE WANT TO BE MORE OPEN TO THE WORLD OUTSIDE, WE FIRST NEED TO BE MORE OPEN TO OURSELVES." (CHRISTIAN SCHUELLER)

"We assume everybody sees the world as we do. But whether we find, for example, an interview quote interesting or boring, moving or exaggerated, authentic or artificial, depends on how we relate to the person saying the quote. An interviewee will evoke different emotions within a reporter," says Christian Schueller.

"And the – often unspoken – expectations of colleagues and superiors will also

influence the way the reporter perceives what she or he hears. Our worldview is shaped by all these factors – rational and irrational, personal and organizational.”

While journalistic work routine often leaves this subjective filtering process unnoticed, Reflective Practice Groups make the reasoning behind routine choices visible. And as human bias should be avoided as much as possible in any ethical AI project, these reflection groups can be very useful. “They can help in detecting some unexpected biases – a good starting point for any dialogue on artificial intelligence.”



“The more technical journalism gets, the more its content is influenced by constraints we are not directly aware of,” explains Christian. For one, the interaction with other people in the production process is decreasing. This results in less interaction and discussion about the best possible choices. “In one of our Reflective Practice Groups, a magazine journalist spoke about her experience as camera-journalist (camjo) – a role that some of them clearly enjoy. However, others felt that their stories are getting more ‘one-dimensional’, as the verbal exchange with the cameraman is no longer there.”

The ORF currently has three Reflective Practice Groups, consisting of eight employees of various job grades and ages (and hence different levels of self-confidence). They come together for 90 minutes once a month to reflect on real-life cases. What happened with the production of a particular story, and how can other group members relate to this on the basis of their own experiences?

The key is the diversity of the perceptions and perspectives that are brought together in each group. A senior television magazine editor will be talking to a junior radio news editor and someone from the online department. “The whole idea of the Reflective Practice Groups is that we take the group members seriously. They are all professionals. They don’t need to be told something by someone in a workshop, but they need space to be able to reflect together.” It does require something of each member, as sincere and non-judgemental listening is needed. “Each group member needs to fully accept the diversity of the group and has to overcome his or her imprinted perceptions of hierarchy.”

The aim of the groups is not to find issues that can be solved by the HR or news departments. “The result is more subtle and personal. The people that are in the groups will improve their self-confidence and sensitivity to others, making it more a ripple effect than something that needs to be implemented top down.”

Christian is the longest-serving journalist at ORF (42 years!), which provides him with the much-needed credibility and trust of his peers. Not only because he has worked in almost every department and position, but also because he is a psychotherapist by training.



ORF, AUSTRIA TAKING QUALITY METRICS SERIOUSLY



Klaus Unterberger
Head of Public Value
klaus.unterberger@orf.at

TAKEAWAYS

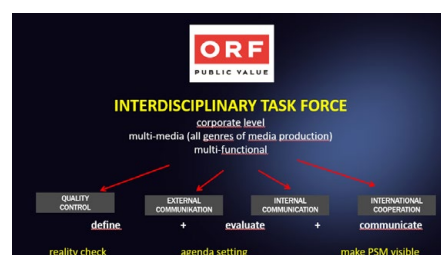
- The challenge to define what quality actually means is a huge intellectual task. However, working with the issue for years has provided ORF with substantial capabilities that provide a foundation for discussing the issue.

- Categories like national and international value have been chosen because they have been considered important components of public value in Austria. The qualities defining the value of the service may have to be decided by each country, thereby limiting the possibilities for expanding the system directly to others. However, it can be used as inspiration.

- Efficient Quality Control is not a police inspection. It empowers journalists to create distinctive quality and delivers data for credible legitimization of PSM funding.

In trying to measure or even automate actions in order to increase the public value of journalistic content, specific goals should be formulated in such a way that a computer can understand them. Some guidelines for this work might be found from the company that has perhaps the most structured and comprehensive quality control and measurement system in Europe, Austria's ORF.

In using data and machine learning, it's essential to know what your goals are and how to measure them. But how can we measure public service value and distinctiveness as a complementary benchmark to existing quantitative data? ORF, the Austrian broadcasting corporation, has taken this challenge extremely seriously.



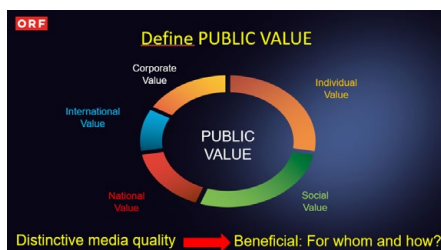
The 29-page paper describing their quality assurance system begins as follows: Trying to define quality is like trying to nail a pudding to the wall. The rest of the document explains how to nail it anyway.

"I think we have one of the most comprehensive quality-control systems", says Klaus Unterberger, Head of Public Value of ORF. Crucial benchmark for its relevance are Public Value Criteria, focusing on the distinctive quality of Public Service Media. The Board of Trustees at ORF approves the quality assurance system. Then, an independent expert assesses the individual measures taken within the company to meet these standards. There's also a Quality Committee, set up by the ORF Public Council, which is regularly consulted on quality issues and which gives recommendations. And every two years, the media regulator reviews the quality system as well.

"I THINK IT'S A NECESSITY THAT WE BRING IN THE IDEA OF THE COMMON GOOD TO ARTIFICIAL INTELLIGENCE." (KLAUS UNTERBERGER)

In the assessment and reporting, quality is defined in an orderly way, containing five broad areas and 18 performance categories:

1. Value in individual context: reliable, credible information, practical value in use, quality entertainment, providing education and knowledge and expressing social responsibility
2. Societal value: considering social and cultural diversity of the society, variety in orientation, integrating of different groups to the society, closeness to citizens and contributing to cultural and artistic life.
3. Nation value: producing original Austrian media content and covering relevant topics on Austrian identity, such as history, culture and social development, as well as contributing to the national and local creative industries.
4. International value: contributing to European and international news through an extensive correspondent network and international cooperation, specifically covering issues on Europe.
5. Company value: innovating and developing ORF as a business, increasing its skills and the transparency of its operations.



ORF has also commissioned several scientific studies looking into the challenges it faces, both present and future. Looks exhausting? It may be, but the company states that: “This orientation towards the common good is one of the decisive distinguishing features of public broadcasters compared to private broadcasters and thus one of the decisive premises for the legitimacy of public broadcasting in general.”

The audience is surveyed about their satisfaction in relation to the above criteria. “We create data by a survey, simply by going out and asking people if they are happy with what we are doing using the public value criteria.” They create quality profiles for all programme genres and use expert assessments as part of the process.

Another crucial issue at ORF is how to make this system applicable in practice, to everyday operations, and how to make tight monitoring of quality fit with the principle that journalists must be independent, creative and responsible. The documentation, though exhaustive, is divided into applicable and concrete sections on a programme level. In news, for example, the longest chapter elaborates on quality concerns and issues of trust.

The quality documents contain the key elements of journalistic ethics, with strict requirements applying specifically to ORF. For example: “In the independent design of broadcasts, such contributions must be included that contradict the personal opinion of the editor.” Or: “Essential for any objective representation is the clear separation between facts and opinions.”



In the latest quality assessment in 2018, one of the independent experts, Dr Florian Oberhuber, gives ORF credit for meeting the targets. In the analysis of 22 quality characteristics, ORF achieved acceptance rates of between 78% and 97%. However, he advised that there were areas where things could be improved: 8% of respondents were critical about the neutrality of reporting, and although this number was low he concluded that neutral behaviour by ORT moderators is given a high priority by the audience.

Dr Oberhuber concluded: “Mostly the audience perceives positively the presence of a high diversity of protagonists. In particular, there are several occasions when they wish that in the future ORF would have a greater diversity of studio guests and experts.”

In addition, for example, a broad variance in terms of interview styles and the signposting of the corresponding goals (e.g. ‘in the crossfire’ versus ‘in conversation’) can increase audience satisfaction.

Applying quality metrics like ORF's in a machine-readable technical format, so that support can be obtained from algorithms, would surely be a challenge, but this example shows that you can approach quality in a structured way, almost from a scientific standpoint.

Klaus Unterberger, would it be possible for a computer to measure quality and code an optimizing algorithm based on the metrics like the ones developed at ORF?

"Only if it's possible, that it's not market-driven, but socially driven. Once we have something that can be checked by the public. Then I think it would make sense."

"I think it's a necessity that we bring in the idea of the common good to artificial intelligence."

LINK TO THE WEBSITE:

<https://zukunft.orf.at/>

https://zukunft.orf.at/show_content.php?sid=144&language=en



CREATIVE DIVERSITY NETWORK, UK DIAMOND

Amy Turton, Project Manager Diamond (amy@creativediversitynetwork.com)

Deborah Williams, Executive Director (deborah@creativediversitynetwork.com)

TAKEAWAYS

- By having all major broadcasters on board in the project, the industry is making a loud-and-clear statement that they truly value diversity.
- Diversity is one of the major public service values. By using an empirical and scientific approach to collect data, PSM organizations can see if they are indeed upholding this value, and where improvements need to be made exactly.
- Sometimes you still need people filling out forms to obtain valuable data sets.
- Having this kind of data on diversity issues leads to concrete changes in the industry.

The Diamond project has for two years been collecting the demographic data (age, gender, ethnic background, sexual orientation) of all on- and off-screen contributors to television programmes of all major broadcasters in the country. This data gives nuanced insights into the question of how diverse the television industry really is. The upcoming report will also include analysis of news productions.

Perhaps contrary to popular belief, there are more people visible on British television that are lesbian, gay or bisexual¹, or who have a minority ethnic background, than there proportionally are in the UK population. But on age (50+) and disability, there is significant under-representation, with only 29% of the on-screen individuals aged over 50 compared to a national average of 36%. Concerning disabled persons, the percentage is even lower: only 8% of on-screen contributors have a disability compared to 18% in society.



These remarkable insights have been gained thanks to the Diamond (Diversity Analysis Monitoring Data) Project, which has been running nationwide since 2016.

'Who is on TV' and 'Who makes TV' are more difficult questions to answer than one would think. There are studies that list the demographic characteristics of people portrayed on television based on what is visible – gender, rough age, disabilities. But these fail to list nuanced or invisible characteristics and do not include information on people behind the camera, who are just as important for creating a diverse climate in the industry.

"WE ARE COUNTING. THAT IS UNIQUE AND USEFUL, ESPECIALLY IN THESE TIMES OF POPULISM." (DEBORAH WILLIAMS)

With the introduction of the Diamond project in 2016, the entire television industry in the UK started to collect data on diversity. "BBC, ITV, Channel 4, Channel 5 and Sky all participate," explains CEO of the Creative Diversity Network, Deborah Williams. She underlines that the real problem is the system, not one particular organization: "Our aim was never to blame a particular broadcaster. The system is the problem, not one particular organization. And the question is how the industry as a whole can solve the problem."

¹ In Diamond's data gathering, sexual orientation (gay, lesbian and bisexual) is measured separately from gender identity (transgender).



“We are counting. That is unique and useful, especially in these times of populism,” explains Deborah. The counting is done through an online questionnaire. Anyone who has contributed to a television production receives an email requesting to share information on their gender, age, religion, ethnic background and other specifics. In line with data-protection regulations, these are anonymised and kept in a database for a maximum of two years (people contributing regularly only need to fill out the form once every two years). The success of the data relies on the willingness of people to fill in these forms, but as awareness of the project is growing, the number of participants has grown from 8,000 people in the first year to 28,000 in the second and probably to around 50,000 this year.

Various genres are compared, such as entertainment, drama and factual, but it’s striking that news and sports are not yet included. “In the first two reports, we have not yet been able to look at these genres, as their production methods and systems significantly differ from those in other genres. But for the third report that will be published in September 2019, we will have data on news and sports from three main broadcasters (Channel 4, ITV and ITN), making it possible to draw some initial conclusions.”

Deborah quotes Ben Frow, Director of Programmes at Channel 5 on the reason why the broadcaster wants to focus on news: “Broadcasting is a sector which holds a lens up to society and nowhere is this more the case than in news and current affairs. It is therefore an important step forward for Diamond that news and sports data can be reported alongside other genres. This will enable us to see where we need to improve.”

The outcomes of the reports are important for the industry so that it can make changes. “The more data we have and the more we share, the more the industry can understand what they can do, instead of focusing on the impossibilities,” says Deborah. And to make the right decisions, it is important to exactly know where the problem lies. “In sports, for example, we can see that there have been major changes on-screen in the last years, with more women and ethnic minorities, but behind the scenes, it’s still very much a white male-dominated genre.”

For the Diamond project, the next step is to look at diversity in various professional roles: who are the decision-makers? And who are the executors, recruiters and financial backers? “This will lead to an even fuller conversation, especially when compared over genres.”

FOR MORE INFORMATION

<https://creativediversitynetwork.com/diamond/>

HOW TO A WAY TO DESIGN METRICS

Newsrooms analyse quantitative data on audience behaviour using both generic and in-house tools. But some types of behaviour are hard to measure, and even deciding which ones to measure can be a challenge. And as the Reuters Institute points out in its [compelling report on editorial analytics](#), even with the right data, newsrooms still need the right internal culture and organizational structure so the data can and will be used¹.

There's no one 'best way' to use metrics in the newsroom. Success depends on many variables, and different organizations will need different things. What's clear, however, is that a single metric rarely yields all the insight a newsroom needs. Most use a combination of metrics.

D.L. Rogers, in his [Digital Transformation Playbook](#), says organizations need to figure out how they want to create value before they choose which metrics to focus on. For public service media, values like engagement, loyalty and impact have become increasingly important, although reach remains dominant. The table below, from the Reuters Institute report, shows how different metrics are used to measure different values. For example, 'reach' might be measured by the number of unique users, while 'engagement' might be better measured by session time.

		RELATIVELY CLEAR DEFINITIONS/MEASUREMENTS ←-----→ LESS CLEAR DEFINITIONS/MEASUREMENTS			
Interest		REACH	ENGAGEMENT	LOYALTY	IMPACT
Definition		The number of people exposed to content in a given period of time	Time someone spends with content during a session or a given period of time	Frequency with which someone seeks out a given brand's site, app, or social media content	Whether the content made a difference in people's life (individual or societal)
Measurement		Unique users (website) User sessions (apps) Impressions (off-site social media use)	Session time (website) Time in app (apps) Engaged time (time spent actively interacting with content) Number of pages/visit	Return visits (website) User sessions (apps)	No agreed upon metrics
Data issues		De-duplication from devices to people App/browser proliferation Access to off-site data	Aggregation of individual users' engagement across devices, apps, and browser use Integrating on-site/app engagement with off-site engagement	Aggregation across devices, apps, and browser use Integrating on-site/app use with off-site use	No agreed upon sources of data

Figure 4.1. A range of metrics mapped in terms of relative clarity of definitions and measures.

1 Cherubini & Nielsen, 2016

INTERESTING READS

Much has been written on how metrics can best be used in newsrooms. Here's a list of useful resources:

- The Reuters Institute's 2016 [report](#) on how news media are developing and using audience data and metrics
- Dan McFunley's blog entry on how [dangerous](#) it can be to rely on real-time analytics
- The American Press Institute's clear and detailed guide to creating a metrics-driven newsroom [in six steps](#)
- [Journalism.co.uk](#)'s interview with metrics experts [David Brauchli](#) (Content Insights), [Rob Hammond](#) (Trinity Mirror) and [Elinor Shields](#) (BBC News) on how to best use analytics in the newsroom
- [This article](#) on Medium, which summarizes a talk given by editor Esra Dogramaci on building digital strategy, sharing key advice like: "Things like views, reach, clicks and impressions may look impressive on aggregate but are very superficial. They aren't actionable metrics – meaning we can't really use them to feed into editorial or content strategy. Things to pay attention to are dwell time, retention rate and watch time. Look at how your content is consumed and shared."
- Digiday's articles, in which [retention](#) is considered a key strategic element for publishers.

COMPANIES PROVIDING METRICS FOR NEWS ORGANIZATIONS

Newsrooms can either use generic analytics tools or build their own product. Chartbeat is the number-one generic tool in use, but there are many other off-the-shelf solutions:

- Metrics for news: <https://www.metricsfornews.com>
- SmartOcto: <https://smartocto.com>
- Content Insights: <https://contentinsights.com>
- UpScore: <https://www.upscore.com/en/>
- Chartbeat: <https://chartbeat.com>
- Parse.ly: <https://www.parse.ly>
- Ezy Insights: <https://ezyinsights.com>
- Newswhip: <https://www.newswhip.com>
- Crowdtangle: <https://www.crowdtangle.com/>

POSSIBLE INDICATORS FOR SUCCESS

This overview of metrics from Reuters is extremely useful (to be found on pages 34 and 35 of [their report](#)):

Bounce rate: Percentage of single-page sessions (e.g. users who land on a site and leave immediately).
Click-through rate: Percentage of visitors to a page who follow a hypertext link to another webpage (e.g., when a user gets from one article to the next through links, 'related articles', and similar means).
Concurrent visits: The total number of people on a site at a given point in time. Unlike metrics like pageviews or unique users that are typically aggregated over a set period of time, concurrent visits give an impression of the number of real-time users and can inform decisions on when to publish as well as decisions around newsroom workflow.
Conversion rate: Percentage of users who take a desired action (e.g. buy a subscription, register, sign up to a newsletter, etc.).
Engaged time: Amount of time that users spend actively interacting with a page or site – reading, writing, scrolling, watching. Sometimes reported as an aggregate of 'attention minutes', sometimes as median attention time for a given article, section, or the whole site.
Entry rate: Percentage of visits starting on a particular page (i.e. the number of entry pageviews divided by total number of visits). Important for thinking about how to use popular parts of a site to draw people to other content. The reverse is 'exit rate', the percentage of visitors who left a site from a specific page.
Exit page: Last page accessed before a user leaves the site.
Pageviews: Any time a user views a page by any method, such as clicking on a link, typing in a URL, or refreshing a page. Pageviews are sometimes called 'clicks', 'hits', or page impressions. Pageviews are sometimes broken down by device type, e.g. PC, smartphone, tablet.
Pageviews per visit: Average number of page views per visit (often expressed over a given time period).

Reach: Number of people who have theoretically been exposed to a given piece of content from a particular media organisation or exposed to content from that brand. Reach is normally calculated as a percentage of a known population, which can be relative to population (for broadcast ratings and print circulation), internet users (for site visitors), or the number of Facebook followers of a site who have had content shown in their NewsFeed. In most cases, reach is measured in platform-specific ways.
Recirculation: Percentage of the audience that has engaged with a particular piece of content (article, video, etc.) who proceeds to engage with another piece of content.
Referred traffic: Traffic that comes from an external source (as opposed to direct traffic). Referred traffic includes social referrals, search referrals, link referrals (from other sites), and 'dark social' where analytics tools cannot determine the referrer (e.g. emailed links).
Scroll depth: How far users scroll down on a page.
Site performance: Various metrics including load time and execution speed that measure how quickly a site or an app serves users.
Social shares: The number of times a piece of content has been shared via social media sites like Facebook, Instagram, Twitter, etc. Other social metrics include interactions (number of 'likes', etc.) and mentions.
Time spent: Amount of time (in minutes or seconds) visitors have spent on one site or on a particular page. Time spent and engaged time is not always the same, as not all analytics tools take into account whether users are actively interacting with a site or simply on it (while potentially doing something else).
Unique users: The number of different people who have visited a website or app in a given period of time. In reality, unique users (often referred to simply as 'uniques' or more precisely as 'unique browsers') are normally counted by associating individual IP addresses with a cookie. This tends to lead to both inflated numbers and uneven data quality – an individual accessing the same site from multiple devices, multiple browsers, via both apps and browsers, or who has deleted cookies from her browser, will be counted several times. Without registration, it is hard to link unique users directly to individuals.
Visitors: Normally measured on the basis of unique users . Can be broken down to understand for example loyalty (new users, returning users, loyal users) or device use (PC, smartphone, tablet). Like uniques, visitors are normally authenticated on the basis of cookies in the browser or through registration/subscription.



CHAPTER 04

CONTENT ONLY HAS AN IMPACT WHEN IT'S FOUND

THE MOST STRATEGIC USE OF AI CURRENTLY: HELPING GOOD CONTENT FIND AN INTERESTED AUDIENCE

One of the fundamental challenges in the era of digital overload is that it is hard to find good content amid all the clutter. So the mission is to connect content with the audience that's interested in it. While it's indisputable that personalization is valuable for individuals, for public service media it's crucial to serve the public interest as a whole. And yes, it is possible to find a balance between these potentially competing interests.

The logic is simple: if content is curated to match a person's individual interests, it will be more relevant for them. However, this is also the disruptive idea that created first Google and then Facebook, Amazon and Netflix. Using algorithmic and machine-learning technologies, these newcomers have been able to capture both the attention of the audience and the advertising money without any content production of their own, initially.

Some of the most fundamental challenges for journalism – and democracies as well – have risen from these 'targeting machines': fake news on a massive scale, the ability to live in one's own filter bubble, an overabundance of information, the unbundling of news products into atomized pieces and a destruction of the old income model that financed a large part of quality journalism in the 20th century.

"THE DATA OVERWHELMINGLY SHOW THAT PERSONALIZATION DOES INCREASE ENGAGEMENT, AUDIENCE SIZE AND LOYALTY. IT DOES ALL OF THE THINGS THAT PERHAPS WE WISH IT WOULDN'T DO, BUT IT DOES." (PHILIP NAPOLI, PROFESSOR OF PUBLIC POLICY AT DUKE UNIVERSITY)

So it's important to ask whether these problems can be solved by avoiding these technologies altogether. We think not: the reality is that these companies have created highly innovative services that serve people's key needs.

Automated targeting, or segmentation using artificially intelligent systems, is a lever for increasing the value of good content because content that does not find an audience has no impact.

However, the solution isn't simply to copy what the tech giants invented years ago. The value for modern public-service journalism has to be found in new combinations that are applied in a distinctive, public-service way. It is in the public's interest to have technologies that combat or counteract the negative or distorting impact of these other mechanisms.

PERSONALIZATION CONNECTS STORIES TO INTERESTED AUDIENCES

There is considerable fear that public service media might lose its role as a common arena for discourse and understanding if the content offered is not the same for everyone. Many (including privately owned media companies) argue that personalization is a better fit for privately owned media. However, meeting individual needs more accurately can be in the public interest too.

Deliberate avoidance of news programmes is at an alarming level, typically expressed in tens of per cent of audiences.¹

Klaus Unterberger, Head of Public Value of Austrian ORF, warns: It's about relevance and access: If more and more people are disconnected from quality media – as they are in times of fragmented audiences, echo chambers filled with populist propaganda, personalized communication driven by commercial interest and an alarming polarization of society – there's a real danger for democracy. Never before in the history of journalism, the situation has been as bad as today."

Jill Nicholson of Chartbeat says: "Even in public media, there really isn't one single tactic that is going to work for everyone. Audiences all have their own personalities and preferences. I think it's going to be a crucial skill for every newsroom to stay up to date with those individual reader behaviours, and really learn from them."

When Executive Product Manager David Caswell, from BBC News Lab, talks about personalization, he speaks about three different layers. The first is content recommendations – something almost all advanced newsrooms already do when they suggest related content to their audiences. Then there is personalizing the presentation of stories – like Netflix is doing with different promotion pictures of the same offering for different users. The third is the personalizing of the story itself: making different versions for different people, based on their individual profile. Combining all three layers may be the most effective approach.

SEE OUR CHECKLIST: UNDERSTANDING PERSONALIZATION

As unpleasant as it may feel, the whole idea of reaching the masses with the same broadcast news needs to be replaced by something else in this new era of many-to-many communication. In public service media, that means linking common agenda-setting to serving individual needs. But succeeding in practice is a much more complex task than simply understanding the concept.

One reason why tech giants have been so successful is that they have vast content inventories: for Google, that's the whole internet, for Facebook, it's close to two billion daily users and their posts. For public service media, personalization strategies have to be coordinated with content strategies; otherwise, there won't be enough suitable content to recommend.

The challenge here is much tougher than just trying to understand audience diversity based on data. In effective personalization, diversity of interests may have to be addressed with thousands of different story options.

Director Nicholas Diakopoulos from Northwestern University's Computational Journalism Lab says that the impact of targeting and optimisation of content distribution is clear: "We just finished up a study where we analyzed several hundred thousand A/B tests from hundreds of different publishers: It really does work. There are double-digit boosts on average."

Philip Napoli, Professor of Public Policy at Duke University, agrees. He says that there is no doubt that personalization works if done well: "The data overwhelmingly show that personalization does increase engagement and increase audience size and loyalty. It does all of the things that perhaps we wish it wouldn't do, but it does." He underlines that it captures some part of journalistic value creation but leaves other parts out, such as the journalistic mission and journalistic values in a democracy as well as community building. For public service media, this may mean combining the two approaches by offering

¹ (Newman, Fletcher, Kalogeropoulos, & Nielsen, 2019)

important, editorially selected news for everyone but then adding a personalization layer.

The EBU's PEACH project, an initiative between several public service companies and the EBU, is aiming to achieve this goal.

DON'T BE AFRAID OF A ROBOT; HE'S YOUR HUMBLE SERVANT

Personalization efforts can benefit from robot journalism, and robot journalism benefits from personalization. When audiences need to be targeted individually, news organizations may need to have various versions of a story to distribute. And robots are very capable of producing these on a large scale.

In robot journalism, investments scale only when the topics and datasets allow it. There are robots producing stories based on data in sports scores, election results and in some cases from public datasets. Among the most famous are Bloomberg's Cyborg, The Washington Post's Heliograf and those of Xinhua of China. In practice, this means robot journalism is normally reserved for stories created from numbers.

The rise of robot journalism has been slower than once predicted. "Five years ago, there were many bold predictions about how automated journalism will develop," says Andreas Graefe, Professor at Macromedia University, in a report published for WAN-IFRA in March 2019. "In reality, not much has changed. Progress is steady but slow." The reason is the complexity of language: the text robots are still mainly based on templates, and there is seldom any real AI in them. They are more like old-school computer programs: rule-based text generators. Therefore: "It's hard to design and contextualize narratives in advance for subjects where the outcome is uncertain, like politics," conclude the authors of the WAN-IFRA report.²

"Algorithms can only describe what happened – not why, making it best for routine stories based solely on facts that have little room for uncertainty and interpretation," writes Andreas Graefe.³ So it's still up to journalists to put the numbers into context and provide in-depth analysis.

The predicted main growth outlet for automated reporting will be breaking down journalistic work into actual informational pieces and processes, and analysing what can be automated and which are inherently human tasks. Computers will step in by providing raw material and insight for human journalists to take to the next stage, as at Reuters.⁴

Norwegian news agency NTB analyses what can be automated and what tasks are essentially human through their innovations on robot journalism.

Carl-Gustav Linden and Laurence Dierickx in a recent LSE publication wrote that the metaphor of 'robot journalism' frames the issue as machines coming to take the jobs of human journalists. That means the users of the term "have managed to destroy, or at least delay, a move towards a future of augmented journalism, where smart machines play a supportive role and allow reporters to do their jobs better."⁵

The balance between human and machine is important. "The work of a journalist is creative, it's about curiosity, it's about storytelling, it's about digging and holding governments accountable, it's critical thinking, it's judgment – and that is where we want our journalists spending their energy," said Lisa Gibbs, the director of news partnerships for AP about why 3,700 automated financial stories are being produced per quarter.⁶

² (Lindén & Tuulonen, 2019)

³ (Graefe, 2017)

⁴ (Graefe, 2017)

⁵ (Linden & Dierickx, 2019)

⁶ (Peiser, 2019)

Do people trust stories written by robots more or less than those written by humans? A study conducted in Germany shows that the credibility scores for computer-written stories were higher than for human-written stories, but human stories were still favoured because they were more readable.⁷ Professor Philip Napoli says that the jury is still out but wonders: “Wouldn’t it be ironic if we realized, boy, that the average person trusts an algorithmically produced news story more than one written by an actual human being.”

WE’VE COMPILED EXAMPLES OF ROBOT JOURNALISM AND COMPANIES PROVIDING TECHNOLOGIES FOR BUILDING SUCH SYSTEMS IN OUR CHECKLIST ON ROBOT JOURNALISM.

IT’S NOT ENOUGH TO PERSONALIZE; THE CONTENT HAS TO CHANGE TOO

So we have personalization on the one hand, and robot journalism on the other. What can they achieve together? And will it be enough? To get a grasp of this, we will look at the Chinese example of Toutiao. This news app has a massive 250 million users who consume more than an hour of news content every day. That’s more than people use Facebook on average. What can we learn from that?

Toutiao, together with other examples included in this report (Mittmedia from Sweden, Yle in Finland and the BBC), face the same bottleneck in news personalization. Even when the personalization engine is well built, it may not find the content people would like to consume. That’s why it’s worth noting that Toutiao, however controversial its success may be, is working on robot journalism and at the core of its strategy has created a contributor program with income-sharing schemes that now has close to one million participants.

For public service media, this means that something about the offering has to change based on the feedback from the algorithm. It may be the topics, the angle or even the tone of voice of the stories.

David Caswell of the BBC says that in the new many-to-many communication environment, even the news artefacts that we’re used to will have to change. Structured forms such as news programmes and newspapers present news in a hierarchy of importance, but in a digital environment, those narratives are disappearing. The overall picture is created in the mind of the consumer.

“The entire profession of journalism arose within the one-to-many communication environment, with its one-size-fits-all artefacts, and that is what makes the transition to the many-to-many communications environment so extremely difficult. The audience, especially the younger audience, is quickly making the transition to the many-to-many communications environment – because it gives them choices and it gives them a voice – but they as yet have no ‘journalistic utilities’ available to help them assemble coherent mental narratives from the vast, global, permanently accessible torrent of one-size-fits-all news and other communication artefacts. To me that is the heart of the challenge facing journalism,” says Caswell.

In practical level, take, for example, the digital audio bites created at SR (Swedish Radio): they have to be understood without intros and outros. So there’s a need to change the forms themselves of radio-making and invent a new way of contextualizing the news.

⁷ (Graefe, Haim, Haarmann, & Brosius, 2018)

Jukka Niva, Head of News Lab at Finnish Yle, and David Caswell from the BBC agree that the next step is to ask fundamental questions in newsrooms: What kind of stories are we producing? What are the artefacts used for storytelling and making meaning? “If you want to create an excellent user experience in the output, the input has to change. Our development has been centred around technology so far, and now the editorial has to react to the feedback we receive,” says Jukka Niva. His colleague Jaakko Lempinen, responsible for Yle’s AI development, adds: “We are reaching a point where ROI of implementing AI into different user interfaces is getting lower and lower. Instead, we need to turn our focus to how our content design, production, acquisition, distribution and marketing can be helped with artificial intelligence.”

TARGET INDIVIDUAL NEEDS IN A PUBLIC-SERVICE STYLE

Not tailoring content to individuals is not an option – the choice is therefore how to do it. When used by commercial media organizations (like Google and Netflix), personal recommendations can reinforce an audience’s pre-existing preferences and undermine shared and collective media experiences. Not what PSM hopes to achieve.

Where platforms and commercial organizations use the techniques to sell more products or satisfy people’s expectations, public service journalism can use the same technologies to match content more accurately with higher purposes and societal values. PSM values, such as impartiality, distinctiveness and diversity, can be used as indicators for personal recommendations.

The BBC connected four possible areas of ‘success’ to this: “To connect audiences to new content, to connect diverse audiences to shared content, to connect audiences to new experiences and formats and to connect audiences with external publicly valuable content and services.” This means designing algorithms to counter filter-bubble effects and explicitly optimizing for social and cultural diversity.⁸

This goes back to the question of how to define diversity in public service broadcasting. Is it enough to offer diverse content, or should the content that is actually consumed be diverse as well? And how much should PSM organizations influence how people consume content? Natali Helberger, Media Professor of the University of Amsterdam argues: “In a situation of content abundance and limited attention, public service media could have an important function in contributing not only to diversity of supply but also to diversity of exposure.”

“This proposition inevitably triggers the question of whether we would like to live in a world where the (public service) media research, determine, and even predict our information needs and preferences and steer us toward more diverse choices,” she recently wrote.

Helberger, together with Dr Mira Burri from University of Luzern and others, is even examining the concept of whether the public service mission should be redefined to prioritize offering diverse content, even if the content is produced somewhere other than in organizations’ own newsrooms.⁹ But before considering such a concept, there must be the recognition that if a PSM company is only curating content without producing it, it has to do the curation extremely well.

The machines are merely executing the codes that humans have developed. Incorporating PSM values into machine-readable algorithms requires clear goals. And it is difficult to find appropriate measures that can work accurately and in real time. Nicholas Diakopoulos sees a tension between what can be optimized for PSM objectives with data and the data that can be easily collected within budgets. “I don’t think anyone really knows how to measure that in a reliable way,” he says. He lists various interesting points to start with: diversity of topics and opinions, transparency, establishing common ground for public debate, etc.

SEE OUR CHECKLIST FOR PUBLIC SERVICE PERSONALIZATION

⁸ (Fields, Jones, & Cowlishaw, 2018)

⁹ Bodó, Helberger, Eskens, & Möller, 2019; Burri, 2013; Helberger, 2015; Helberger & Burri, 2015; van Dijck & Poell, 2015)

The BBC is the company that has worked the hardest with the algorithmic optimization of PSM values. The work has proven to be more challenging than originally thought, especially for news but continues nonetheless.

Some quite advanced conceptual thinking about public value in a many-to-many news environment has been put into practice in the form of the news avatar that NHK of Japan has created: a clearly non-human animation figure called Yomiko, who besides reading the news can even interact individually with members of the audience.

SOMETIMES THE CAUSALITY IS HARD TO PROVE

Roberto Suárez Candel, Head of Strategy and Media Intelligence of the EBU: “What I think is very important is first to identify the need we want to meet, the problem that we want to solve, and then set goals. For example, for a public broadcaster, you might want to have the role of raising awareness, you will want to have the role of helping people to understand better, you want to have the role of motivating people to do something.”

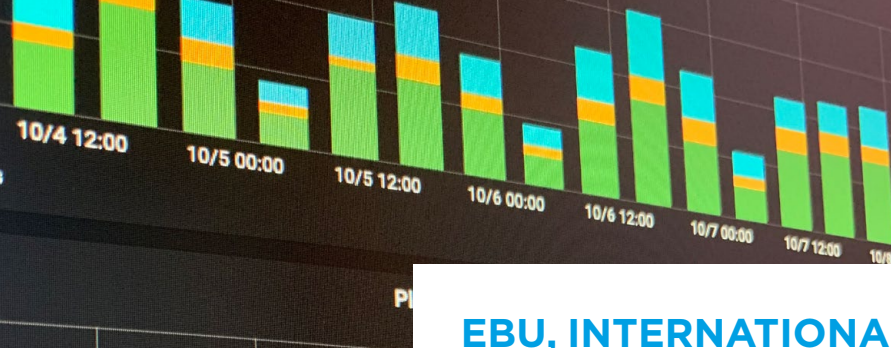
He would not focus too much on direct, proven causality. “I think our individualistic approach needs to change. Because as individuals or as organizations, we need to be aware that alone, we can’t accomplish the desired impact. Public service media play their role, but so do libraries, schools, etc. We should cooperate, sometimes leading, sometimes following. If the goal is met, in the end, it doesn’t matter who did what.”

A good example of a public service broadcaster with a clear impact goal is RAI in the project Femicide. They wanted to raise awareness of hate crimes against women, while at the same time researching the language being used to describe such crimes. They then built an algorithm that is constructing a database of real-life cases, thereby providing a rich source for stories.

An example of a value-based goal is bias-free and gender-neutral reporting. A research team at Stanford University is dissecting television news to detect these biases.

RECOMMENDATIONS FROM THIS CHAPTER

1. Individual targeting of content, web pages, mobile apps or email newsletters is an efficient way to serve people better. There are examples in privately owned media showing that when this is done in a clever way, people really appreciate the service.
2. In personalization, the bottleneck is typically the lack of available content. Traditional news organizations don’t produce enough news suitable for personalization. Therefore, robot journalism, in which several versions of stories from datasets are made, may be a solution, although it is not a silver bullet and has much more limited scope than some thought a few years ago.
3. To be successful in targeting, PSM journalism has to find a way to offer the considerable amount of content that personalization engines need so that they have something worth recommending to each user. That will require changes in content production, and PSM newsrooms may even have to consider distributing content other than their own.
4. Public service algorithms in news are still under development. However, there have been serious attempts to take public interest into account in designing personalized news services. This field will probably develop fast.



EBU, INTERNATIONAL PEACH



Sébastien Noir
Project Manager PEACH
noir@ebu.ch

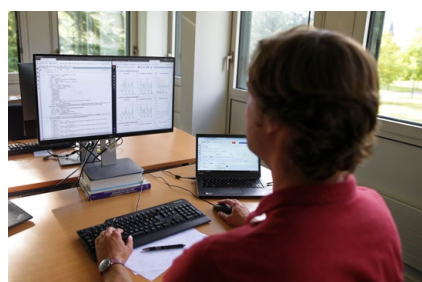
TAKEAWAYS

- PEACH is an example of a data-science project in which public service broadcasters are collaborating. That is an achievement in itself. PSM's resources are small compared to tech giants, and successful collaboration could be a solution.

- This is furthermore part of a broader collaboration framework that is bringing the data-science experts of EBU Members together. - PEACH may be sowing the seeds for something bigger in the future.

- PEACH offers a safe way of 'fast-tracking' capabilities, instead of having to invest in building your own personalization system.

The EBU's PEACH project, Personalization for EACH, unites the forces of public service broadcasters in building a common data-science platform and personalisation tools. One of its tasks is showcasing the paradoxes that PSM organizations have to deal with. For example, should they recommend content that pleases users or try to drive their own agenda? PEACH operates diversified algorithms, which attempt to expand the user's catalogue of content.



"I love bubbles," says Sébastien Noir, product owner of the data-science platform PEACH, which several public broadcasters are already using as a source for their recommendation algorithms. Sébastien's statement is somewhat controversial because breaking down filter bubbles has always been considered a mission for public service broadcasters.

PEACH offers an option for an algorithm that will "broaden a user's horizon". Let Sébastien explain: "With many broadcasters, people are not trapped in a bubble, far from it. If you are in a bubble, that somehow means that we've done a very good job in understanding who you are, and you feel so good that you don't want to escape."

"IF YOU ARE IN A BUBBLE, THAT SOMEHOW MEANS THAT WE'VE DONE A VERY GOOD JOB IN UNDERSTANDING WHO YOU ARE."
(SÉBASTIEN NOIR)

Moreover, do public broadcasters really produce content that is not suitable to be recommended? In his own words: "If there is content that is deemed good and content that is deemed bad, why keep the bad content?"

Here's what Sébastien Noir is actually saying: Don't try too hard to create a public service algorithm that tries not to optimize. Because optimizing is what algorithms are for. Instead, you should make an algorithm choose content that has public value. But then you have to define that value so that the mathematics can make sense of it.

"Why not use your editorial judgment instead, and mix that with algorithmic selection, and make it clear for the users why each piece of content is being presented to them?" he asks.



NTB, NORWAY

ROBOTS PRODUCING SPORTS STORIES



Magnus Aabech
Editor, Development Unit
magnus.aabech@ntb.no



Geir Terje Ruud
Head, Development Unit
geir.terje.ruud@ntb.no

TAKEAWAYS

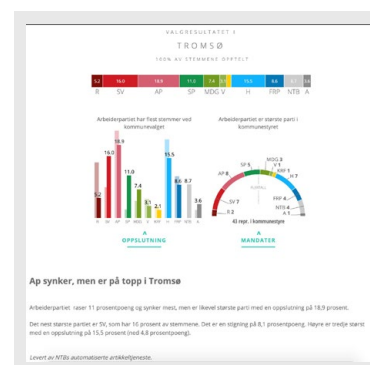
- For common-interest topics, like sports results, finance and election outcomes, it might not be worth investing in your own robots. This technology is becoming standard and commonly available, and it can be wiser to share resources with other companies.

- There are more areas that can be developed with robot journalism. The future also lies in the atomization of information with machines being able to tag and continuously update information which can be used in fact boxes.

- People are still the bottleneck; investing in the right 'data savvy' editorial staff that can communicate with the 'techies' might be even more important than developing machines in house.

Scalable news: that is the potential benefit of the news robot operated by Norwegian news agency NTB. Stories that can be produced faster and for a greater variety of outlets than any human could ever produce. They were the first in Norway to use robots in their sports output and are now searching for new areas of interest – changing their mission from producing stories to producing data.

Not every new technology needs to be fully implemented in every news organization. Sometimes it's better to centralize the efforts and save resources for better journalism. This is at least the view of Norway's news agency NTB. "You need a lot of resources and technical skills when you want to apply robot journalism. And it's really only worthwhile for replicable, non-unique content – not the distinctive stories brands want to be known for. So instead of every news organization investing in robot journalism, they can instead choose to subscribe to it," says Magnus Aabech, editor at NTB's development unit.



"UNDERSTANDING AND IDENTIFYING WHERE AUTOMATION CAN BE USEFUL IS OUR FOCUS. WE BELIEVE THERE WILL BE A LOT MORE AUTOMATED CONTENT IN THE FUTURE, BUT WE NEED TO FIND THE RIGHT NICHES." (MAGNUS AABECH)

Later, he again stresses that news agencies are the most logical employers of robot journalists, as agencies typically provide generic content. "Artificial intelligence is especially useful with scalable topics. You typically work with a template and that needs to be developed every time. This is especially useful when you are using this template a lot – the more 'routine' journalism."

NTB has become a trailblazer in robot journalism because, for the past four years, it has invested heavily in artificial intelligence and specifically robot journalism. They were the first in Norway to use algorithms to produce large volumes of sports reports, especially for football. With local reporters as well as coaches and referees providing the scoring data, the computer is now able to produce a multitude of football reports

16

OKT 2017

MAN 20:54

Sarpsborg rotet bort seieren i sluttminuttene

ID: 14319833

NTB

Sarpsborg ledet 2-0, men Stabæk kom tilbake og sørget for 2-2 i Eliteserien i fotball for menn mandag. Sarpsborg har tatt poeng i ni hjemmekamper på rad.

Vertene tok ledelsen ved Tobias Heintz før det første minuttet var omme. Målet gjør at Sarpsborg-spilleren har tre fulltreffere denne sesongen. Etter tolv minutter doblot Sarpsborg ved Kristoffer Zachariassen. Moussa Njie reduserte for Stabæk etter 27 minutters spill.

Tre minutter før slutt utlignet Morten Skjønsberg for Stabæk.

Sarpsborgs Anders Østli, Matti Lund Nielsen, Joachim Thomassen og Ole Heieren Hansen fikk gult kort. For bortelaget fikk Ronald Hernandez gult kort.

Rohit Saggi var kampleder på Sarpsborg stadion. 4455 tilskuere hadde tatt turen.

22. oktober skal Sarpsborg møte Sogndal, og Stabæk møter Tromsø. (@NTB)

[Levert av NTBs automatiserte artikkeltjeneste.](#)

on various national and regional leagues. NTB is also expanding the range of automatic reporting to include election results, financial news and crime reports. In these fields it is being used to enhance the content produced by their journalists.

NTB is not the only one that uses robots for journalism; so are many news agencies (Reuters, Bloomberg, AP) and news organizations (Washington Post, Yle, NOS). In particular, the progress being made on natural language processing (NLP) seems to be promising, as this technique may make it possible to produce content that is not based on a template.

NTB is currently focused on finding new areas where the automation can be applied. "Understanding and identifying where automation can be useful is currently our focus. We believe there will be a lot more automated content in the future, but we need to find the right niches." NTB is especially looking for opportunities that will free up journalists' time, "so they can focus on the real journalistic work."

One of the areas they are focusing on is the atomization of news, in which stories are broken down into smaller information units. They use machines to tag their data better and keep it continuously updated. The output is information that can be used in structured news stories and fact boxes. "AI can better (and more quickly) update and check the facts than humans can," says Magnus. One area in which they use the structured info are stories on homicide. "We keep the homicide stats constantly updated, so we can always use the latest information for fact boxes. We should also be able to pull out interesting angles by filtering information – for example how many are killed by their partners or what type of weapons are most common. In future, this might be automated as well."

With the implementation of new technologies, NTB is leaning increasingly on data. "We are moving towards offering data, instead of only stories." The idea is that with this data, news agencies can each make their own unique stories, adding value based on their own brand values. But there is a challenge; not many newsrooms know what to do with the data. There is a shortage of editorial staff with the right data skills, those who understand the challenges and don't see machine assistance as a threat. "We need people who can be the intermediates between the editorial and data teams, who understand both worlds and speak both languages."

Another challenge lies in the focus of the content. Most of the 170 customers of NTB are media organizations that have a local or niche focus, whereas most of the NTB content is generic or national. "We need more niche information. We are moving from one-size-fits-all to different sizes for different clients." Automation can help with this process and the head of the development unit at NTB, Geir Terje Ruud, has grand ideas for the future: "Could we perhaps listen to every meeting in every municipality of Norway and make transcriptions of these so that machines can go through the data? This would result in valuable content for local news agencies that don't have the resources to cover all the local political news themselves."

FOR MORE INFORMATION

https://www.ntb.no/om_ntb/robot/

ORGANIZATIONS USING SIMILAR TECHNOLOGIES

Bloomberg Cyborg

Reuters Lynx Insight

Associated Press Wordsmith (2014)

Washington Post: Heliograf (2016)

NOS 2019 elections

Yahoo! Sports

Guardian Australia

The Los Angeles Times Quakebot

Yle Voitto

Press Association RADAR

Mittmedia Rosalinda

Finnish News Agency STT U-bot

COMPANIES PROVIDING FOR TECHNOLOGIES FOR ROBOT JOURNALISM

<https://narrativescience.com/>

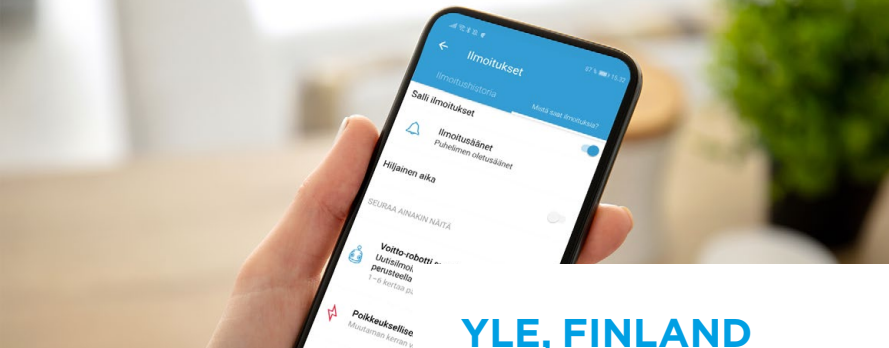
<https://automatedinsights.com/>

<http://unitedrobots.ai/>

<https://urbsmedia.com/>

<https://www.syllabs.com/en>

<https://www4.ax-semantics.com/>



YLE, FINLAND LEARNING TO LEARN FROM FEEDBACK



Jarno Koponen
Head of AI & Personalization, Yle News Lab
jarno.koponen@yle.fi



Jukka Niva
Head of Yle News Lab
jukka.niva@yle.fi

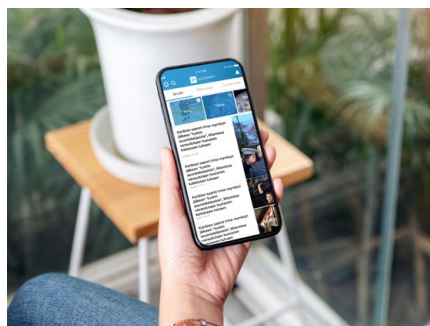
TAKEAWAYS

- Personalization can only be successful if you have an inventory of content that fulfils users' interests and needs. If that's not the case, the whole content pipeline must react.

- The more you personalize, the higher the demand for content diversity. For example, technology is not able to solve the problem of reaching younger audiences on its own.

- Yle has often considered complicated AI features to power its news app but has finally chosen a much simpler solution to create the desired user experience. There's a trade-off between the system complexity and user-oriented value.

Yle, the Finnish Broadcasting Company, was among the first in the public-service world to start experimenting with personalization in its news app. Now, after years in production, the news junkie's daydream app is still a niche product acting more like an R&D platform for Yle. The focus of the developers is already on the next paradigm change: how the feedback received from user interfaces affects the content that the newsroom feeds into the personalization engines.



When it was launched in 2014, Yle NewsWatch, a personalized news app, quickly became a nationwide hit in Finland as well as taking the international market by storm thanks to its novel approach.

NewsWatch picks up data from three sources – user's active choices, other users' behaviour and editorial decisions – and personalizes news for each user by carefully blending these three signals.

Recently, Yle announced that its news app was the first in the world to enable users to give feedback (i.e. whether they like the alerts or not) on smart-news notifications directly from the mobile phone's lock screen, without even needing to open the app. The purpose of the feedback is to give the system more information about the user's preferences; the algorithm adjusts accordingly. "Don't for God's sake send me any more news from NHL hockey league. But yes, more sushi please," according to Jarno Koponen, Head of AI and Personalization at Yle News.

The algorithms developed for the Yle NewsWatch app are now used in other contexts as well. For example, the recommendation engine powering the app provides recommendations in Yle's other news services.

However, the people responsible for Yle's news personalization efforts downplay the relevance of these features. Instead, they emphasize a different angle: personalization actually has limited room to be successful if there are bottlenecks elsewhere, such as a shortage of interesting content to recommend for each user. "What if you simply don't have any more news about sushi? What to do then?" asks Jarno Koponen.

"None of our current products would, even with unlimited investments in algorithms, make us popular among young audiences. More diverse content is needed to be relevant in the future."

“The crucial question is how much you invest in algorithms versus how much you can achieve by working with the user experience holistically, and only from that learn which algorithmic features would actually create value for the user,” he says.

“IMAGINE IF WE PERSONALIZED NEWS ACCORDING TO THE LEVEL OF INFORMATION EACH USER PREVIOUSLY GAINED ABOUT THE TOPIC. THAT MEANS THAT NOBODY WOULD BE WRITING SIMPLE PROSE INTO THE TEXT FIELD OF THE EDITORIAL SYSTEM ANYMORE. THE WHOLE IDENTITY OF A JOURNALIST WOULD HAVE TO CHANGE. THIS IS THE CHALLENGE WE ARE NOW FACING. WISH US LUCK!” (JUKKA NIVA)

Mr Koponen, previously a start-up entrepreneur in Silicon Valley, working within a company aiming to create a new user interface for the whole internet, draws a comparison with his past experience. “Think about the investment needed to bring the whole internet to your system, and then the investment needed to create the algorithms to personalize that content. Investing tens of millions of dollars in each of those functionalities is a crude underestimate.”

While a shortage of available content may limit what you can achieve with personalization, there’s an upside as well. With control over your editorial output, you can start to design the features in the way you want, holistically, focusing on serving the user: “We don’t have to filter anything out.”

Yle-speak is full of slogans. One of them is ‘a smartphone app does not spell the end of human evolution’. Notifications and ‘battle of the lock screen’ have been Yle bets. “We are not a time-consuming business, we are a time-saving business – at least for some people groups,” says Jukka Niva, Head of Yle News Lab. “Our aim is that you won’t have to open the app to stay updated.”

Yle is storing data on single users’ interests in a cloud comprising more than 100,000 keywords of content. Those keywords are clustered with some AI to find links between them. Users’ interests are estimated based on their unambiguous choices and consumption of content, but also on ‘social consumption data’, meaning other users on the app, website or social media. The third source are the editorial priorities decided in the newsroom.



But the new feature on the lock screen has not directly increased usage of the news app significantly. The surge into the ‘hundred thousand’ league, in 2017, represents a decent amount for Finland but that does not make this anything like a mass-market product. The doubling of the number of users at that time stemmed from the introduction of new features on the app such as morning highlights and location-based news recommendations, which had clear value propositions for users. This was facilitated by creating buy-in from the news staff, who had to produce the promised content daily.

The key performance metric became how many people cancelled the feature once they had opted in. Results were very encouraging: for example, with smart personalized notifications, more than 90% chose to remain.

Yle's innovation staff see the next phase of investment as not being in technology but in the newsroom itself: an anthropologist in the newsroom researching the mindset of journalists, for example.

"If you want to create an excellent user experience in the output, the input has to change. Our development has been centred around technology so far, and now the editorial side of things has to react to the feedback we receive," says Jukka Niva, emphasizing that user input is a valuable signal worth adapting to. "This is not easy. You may well end up in a situation in which all your KPIs skyrocket but you wonder if you still have a job. We're changing journalism, and algorithms and tech are just tools for us."



BBC, UNITED KINGDOM

ALGORITHMS PREDICTING CONTENT GAPS



Ben Fields
Lead Data Scientist, BBC
ben.fields@bbc.co.uk

TAKEAWAYS

- It's proven to be difficult to create an understanding of what public service really means when interpreted with algorithms. As machine learning basically is mathematics, public service objectives have to be clearly defined before they can be optimised

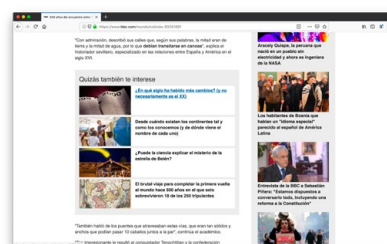
- In news, brand reputation has traditionally been protected and mistakes avoided. When computers are let to make the decisions without human monitoring, the risk of mistakes often overtakes opportunities, especially if evaluators have an editorial background

- Out-of-the-shelf metrics tools are not designed to measure public service value. Creating your own systems may be costly and slow.

Creating public service algorithms in the BBC has proven to be challenging for news. The organisation that invented the term has high ambitions for their use, but so far uses data science and machine learning methods mostly to understand and predict news content consumption and detect gaps in editorial output. The company has been cautious to let algorithms be used in the personalisation of its news content to the audience. With a worldwide reputation as a highly respected news brand, it may be costly to blaze the trail.

A core challenge of the issue for the BBC: How to be different and promote public service ideals, while at the same time provide better news experience for the public - meaning providing them what they are interested in?

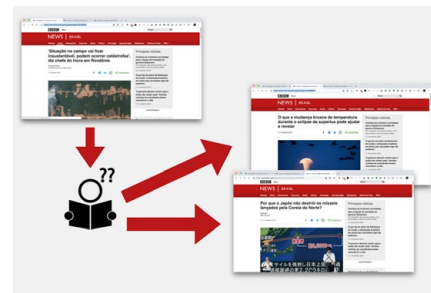
In BBC News, it has proven better to experiment with recommendation features in its global arm, BBC World News, than in national services. Although its global audience is larger than the national one in the UK, it is in the UK that there are more likely to be sensitivities about potential conflicts with the brand image and mission of the BBC.



“THERE'S THIS VERY EARLY EMERGING CONVERSATION AROUND WHAT KINDS OF TECHNOLOGICAL INTERVENTIONS ARE NEEDED TO SERVE THE PUBLIC, THAT ARE DIFFERENT THAN THE KINDS OF THINGS YOU CAN GET OUT OF THE COMMERCIAL.” (BEN FIELDS)

Also, the ultimate aim is different. In the UK, the BBC is known by everyone while in other parts of the world it still has to fight the same fights than any commercial media such as building brand recognition, increasing reach (from the present 400 million), making occasional visitors return.

“Algorithms tested have already shown some progress in offering content that makes people read a second or third article, and significantly increase the likelihood that people come back to the BBC,” says Ben Fields, Lead Data Scientist at BBC.



In the UK, the aim is more to build loyalty and fulfil public service mission. There is the core of the challenge of creating the ‘public service’ in the algorithm.

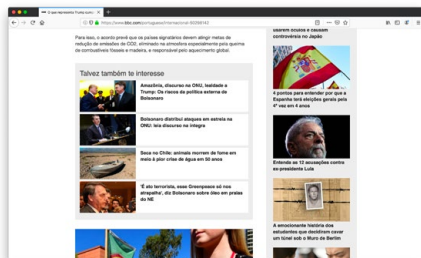
In the UK the BBC does segment its services for different audiences. The BBC homepage looks different for different ages and for people from different regions, but mostly that’s manually curated.

Within these segment buckets, the BBC uses some multivariant testing to check which of the options perform best but makes the decisions always editorially.

Instead of launching (at least so far) individually personalised news sites or apps, The BBC has found a way to use data science and algorithms to help it make better editorial decisions not just by allocating produced content to specific services but by locating gaps in their services that they need to supply.

“As we’ve worked on developing algorithms and recommendations and content aggregation, we sort of realized that actually, we have gaps”, says Ben Fields. By analysing news consumption with data science methods, they found that those gaps may not be just topics, but also tone, angle or medium. The conclusions have been made by analysing both internal and external data sources and comparing. Third-party data sources have been used to look for topics to cover. “That’s about wanting to understand what people who are in the UK, but not necessarily on a BBC property, are interested in seeing.”

“A lot of that is making sure that the gap is something that editorial knows. And if they’re not covering, it’s out of choice, not out of a lack of awareness”, says Fields. So for example with regard to gossip or celebrity news, the BBC will not follow the market.



“There’s plenty of things people are interested in, that does not qualify as news or are not the sorts of things the BBC thinks it should cover.”

An issue the BBC is working on but has not yet settled is the question of how to address the question of social cohesion in its use of these technologies.

The BBC has acted as the respected source of commonly accepted information in a country with a highly partisan press. The challenge for the use of algorithms is can they be targeted to help maintain this status.

Ben Fields thinks that algorithmic work may, ultimately, serve the public good, and become part of the justification for public funding for PSMs. That’s why he, with a PhD in recommender systems, has with his colleagues published several papers on the idea of public service algorithms. They cover how to avoid undermining shared and collective experiences, reinforcing audiences’ pre-existing preferences, and the risk to PSM of ‘becoming more like a goldfish bowl, rather than a window to the world.’

Good news from this research: The audience seems to appreciate diversity and empowerment, shared experience, and understanding why the recommendations have been made. So, there’s perhaps not an ultimate conflict between public service values and content recommendations.

“There’s this very early emerging conversation around what kinds of technological interventions are needed to serve the public, that are different than the kinds of things you can get out of the commercial,” says Ben Fields. But, so far, in practice, he admits: “We’re being quite conservative in our approach on that.”

FOR MORE INFORMATION

<https://www.bbc.co.uk/rd/projects/responsible-machine-learning>

<https://tech.ebu.ch/news/2019/05/can-recommendation-systems-help-fulfil-the-public-service-remit>



NHK, JAPAN

YOMIKO, AVATAR OF JAPAN

Yasunobu Kumad
Deputy General Manager of News Department,
NHK
kumada.y-hc@nhk.or.jp

Kei Yoshida
Editor of News Department,
NHK
yoshida.k-ke@nhk.or.jp

TAKEAWAYS

- Avatar news presenters are able to conduct simultaneously millions of 'personal' conversations with the audience.

- An AI news presenter can fulfil a PSM mission, for example to reach minority audiences with a catered news offering.

- An avatar reading the news might become much less threatening when it's physical appearance is clearly not resembling that of a human being. For the Japanese audience, anime is a logic inspiration.

Virtual presenters, who sound just like humans and conduct simple conversations, will make it possible that millions of people have a new type of relationship with the news. By using this kind of AI, the Japanese public broadcaster NHK is exploring the possibility to attract audiences, especially young people who might find television news too boring. Producers and engineers at NHK are further developing ways to personalize their news and add value to public service news, especially for minority audiences. And it's not all serious; virtual presenter Yomiko also writes her own poems!

In the spring of 2018, the public national broadcaster NHK launched their digital presenter: Yomiko. Looking like a cute anime figure, she is actually an artificially intelligent news presenter. She first appeared in the 11 o'clock television news together with two human presenters and briefly reads news topics, especially online trends. "An animated AI news reader can provide the audience a very unique experience in this digital era, and we thought it would make our news bulletin more attractive," explains NHK.

One of the benefits of Yomiko is the personalization she offers. She has the ability to have a relationship with each individual user as she can, at this stage, respond to their simple requests and have 'a real' (but pre-programmed) conversation. She does this with the anchors in the television studio, but also with news audiences through Google Home and Amazon Alexa. "Yomiko is a new way to deliver the news centered on the concept of 'viewer participation'," NHK says. "When people ask their smart assistant to give them information they need, Yomiko picks it up and reads it out loud. In our view, this can be realized only by using AI technology."

"AN ANIMATED AI NEWS READER CAN PROVIDE THE AUDIENCE A VERY UNIQUE EXPERIENCE IN THIS DIGITAL ERA, AND WE THOUGHT IT WOULD MAKE OUR NEWS BULLETIN MORE ATTRACTIVE." (MR. KUMADA)

They add: "The population of Japan consists of roughly 120 million people, in order to realise a fully personalized services, we need 120 million presenter. But as this is not realistic, AI can find the way. We want to develop Yomiko even further in the future." NHK also has its minority audiences at heart in designing features for the virtual presenter, aligned with their PSM mission: "Yomiko has recently started an attentive news service; written in easy-to-listen wordings, with gentle-slow-sounds, designed for elderly person, handicapped, and foreigners. This is getting a good reputation."

NHK is aware that functionality is not the only important feature for a news anchor; popularity is key as well. So in developing Yomiko's appearance, they asked a famous manga artist (Katsuki Tanaka). To make her 'come alive' even more, there is an augmented reality app that allows people to take pictures with their smartphones featuring Yomiko. But the coolest thing, is that Yomiko can create her own short poems (haiku). "Together with the Hokkaido University, we have trained Yomiko with an enormous dataset of poems, and now Yomiko can generate a new poem automatically. We have a special feature in our evening news bulletin, where she composes a poem on the current topics of that week." This is well received by her human counterpart, one of the presenters, Mr. Yamashita: "Yomiko is comical and charming, and she often writes a very surprising poem."



When it comes to technological advances, Japan is always a good country to look at. The Japanese government's aim to have 'technology and innovation' as a foundation for its competitive edge is reflected in the development of deep learning robots, digital farming, cybernetic medical treatments and other cutting edge technologies. And now Japan is moving ahead with innovation in news.

The automatic reading and voice generation algorithm of Yomiko was developed at NHK's Science & Technology Research Laboratory and is based on deep learning language-processing techniques that the lab made for covering the Pyeongchang 2018 Winter Olympics. The Lab recorded large volumes of news items read by NHK news anchors - humans - and broke down the texts into more than 100,000 phonemes. Yomiko combines these to form words when needed even supplemented with regional accents.

Yomiko is by no means the first virtual news presenter. In 2000, Anonova was created by the Press Association in the UK (<https://www.youtube.com/watch?v=ek-g5A0YTkW>). And the human-like virtual presenters from China caused a lot of stir worldwide, when they were launched in November 2018 (<https://www.bbc.com/news/technology-46136504>).

LINK TO WEBSITE

<https://www.nhk.or.jp/voice/yomiko/>



RAI, ITALY

BATTLING FEMICIDE WITH AI



Sabino Metta
Data Analyst at the Centre for Research
Technological innovation and Experiment
sabino.metta@rai.it+39 0118103190

TAKEAWAYS

- When official institutions are, for whatever reason, not keeping records on important social topics, news coverage can prove to be a valuable source from which to scrape data.

- With a large enough dataset, journalistic organizations can also use data to review their own language and biases and thus acknowledge their responsibility with regard to the issue.

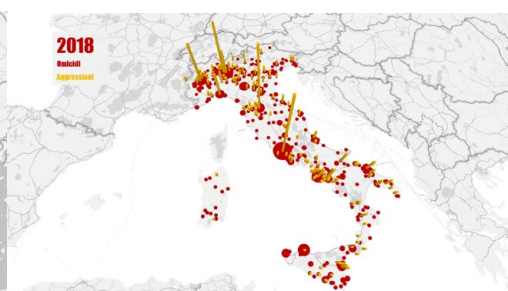
- Journalists do not need to have all the technical skills to review data themselves; by collaborating with a research or technical department, each person can focus on their particular strengths.

Hate crimes against women are a highly politicized topic. Without a collective, neutral reference point, politicians from both the left and the right can assert their own interpretations as the truth. Using artificial intelligence, RAI was able to create – and is still creating – the world’s first official dataset on femicide cases in Italy. This resource, which will soon be accessible to the public, is fed with information drawn from news coverage, and it can be applied to a multitude of social topics on which official records are not kept.

In Italy, like many countries, femicide (the murdering of a woman or girl) has been an issue for a long time. In recent years the public outcry has grown louder and louder. However, the police – for many reasons – do not share an ‘official’ dataset that can form the foundation of a public debate, making it hard to tell whether violent crimes against women are on the increase or in decline. This made it an ideal topic for Sabino Metta, Data Analyst at RAI, to develop a research methodology.

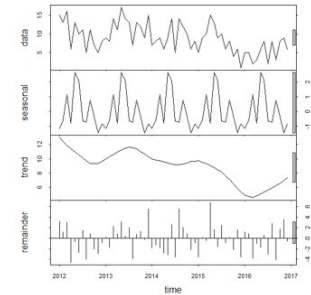
“One of our goals was to collect a neutral dataset that could be used as a reference point on femicide, for all concerned. Up till now, we were confused on how big the problem really was. Some politicians said it was huge; others downplayed the issue. We were missing clear, transparent data on the phenomenon. Our journalists at RAI constantly had to weigh and verify all the information.”

To gather its data, Project Femicide checks text and video files of local, regional and national news outlets, in search for keywords related to femicide. The project is not finished yet, and the machine-learning stage is still ongoing. But the idea is that in the future, automatic tools will be able to aggregate textual items and videos and extract the relevant metadata. “The system still needs to be trained to properly distinguish, among all facts of violence, those that can be labelled ‘femicide events’.”

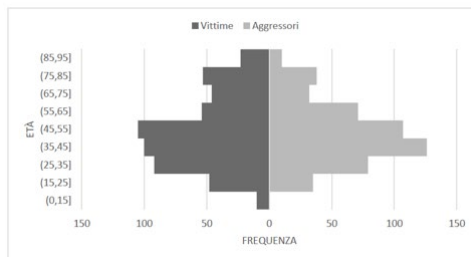


“JOURNALISTS DON’T ENJOY LOOKING AT TECHNOLOGICAL ISSUES. AND THEY DON’T ALWAYS HAVE THE NECESSARY SKILLS. THE IDEA IS THAT WE PROVIDE THE TECHNICAL SKILLS SO THAT THEY CAN FOCUS ON THE EDITORIAL ANGLE.” (SABINO METTA)

The retrieved metadata contains the victims’ names, the geolocation and as many details on the crime as possible. Additionally, information on the perpetrator is recorded: gender, relation to the victim and, where there was a trial, the sentence. “Neither is this a fully automated task. To be sure that information is accurate, you need to follow the news updates. It may happen that for a supposed ‘femicide’ event, the continued investigation reveals a different story altogether, for example, that the woman was killed by a burglar. In that case, the event is excluded from the dataset.”



All the checked data is stored in a database, and journalists can retrieve trends and generic information in dossiers, for example on comparable victims or related types of crimes. But they can also make stories based on particular examples, such as the original television items in which femicide is mentioned and linked to the case which they are investigating.



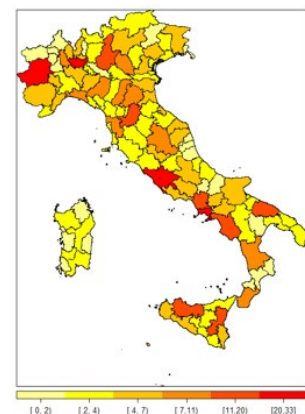
But the project is not only about storing data. “The second goal of the project was to review the language that was used in the news items covering the issue of femicide. Together with researchers from the Department of Social Sciences of the University of Turin, we had a strong hypothesis that this built upon the same socially constructed (unconscious) bias that dominated society; that the women had brought what happened to them upon

themselves. We need to clarify and rectify this bias, so we can start to begin to solve the problem. There are different actors that play a role in this issue, and the media is one of them.”

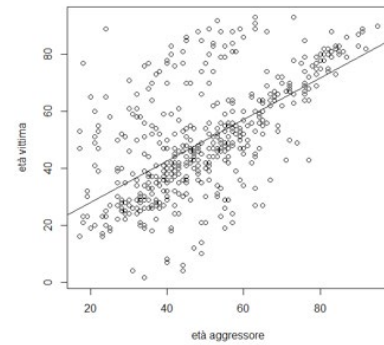
In order to review this possible language bias, RAI joined forces with around 20 researchers from 5 universities. They are working on the same cases that are in the dataset and reviewing the various angles of language bias represented in them. The items and articles are being reviewed and analysis carried out on the language used. “Ultimately, we don’t want to blame the journalists for anything. But we do want to give objective feedback on their performance.”

A third objective of the project is more subtle: to show newsrooms the potential of using AI for stories. Sabino understands that journalists need research support if they want truly to understand the data. “Journalists don’t enjoy looking at technological issues. And they don’t always have the necessary skills. The idea is that we provide the technical skills so that they can focus on the editorial angle.”

With Project Femicide, new contacts between the editorial and research department were initiated, as they could now work together in interdisciplinary teams. “We don’t believe that



journalists should have the same data skills as researchers. That is a different job. But researchers and journalists can combine their skills and work together on projects, which is enhancing the experience. An interdisciplinary approach should be aimed at managing (as much as possible) the unavoidable complexity of the reality. Sometimes we're simply missing or underrating the complexity because we're not able to manage it. With the right team, compelling and constructive storytelling can be supported."



The tool RAI developed can be used on any topic of interest. Sabino thinks it's especially interesting for highly politicized topics: "Like migration, different politicians claim different things and nobody really knows who is right. With a tool like this, you can at least provide a basic starting point for the discussion." He also sees a lot of potential for international cooperation on this topic.

FOR MORE INFORMATION

<http://www.crit.rai.it/CritPortal/?notizia=1553&lang=en>



STANFORD UNIVERSITY, UNITED STATES

DISSECTING CABLE TV NEWS



Geraldine Moriba, JSK Journalism Fellow at Stanford (class of 2019)
gmoriba@stanford.edu

TAKEAWAYS

- AI can be used to reflect on journalistic practices themselves. Tools like these can be used to highlight the values and contribution of PSM to societies.

- By using algorithms and large datasets, assumptions about bias can be tested.

- This tool could be a useful technological addition to the more psychoanalytical approach used by ORF's Reflection Groups. [<link to the case Reflection Groups>](#)

Stanford University has developed a tool that uses artificial intelligence-based image, audio and transcript processing techniques to scan TV news broadcasts for trends in content, bias or polarization. The tool has been developed in an academic setting but can be used by any broadcaster that wants to review its own tendencies in this area.

What biases are present in the news? This is an important question, especially in a country with such a polarized political and social landscape as the United States. In an attempt to find out, Geraldine Moriba set up a research project at Stanford University. "Overall, our mission is to use computational techniques to help improve transparency about editorial choices on TV news. What stories get told? What perspectives are included? Who tells these stories? And what is the race and gender of those telling the stories?"

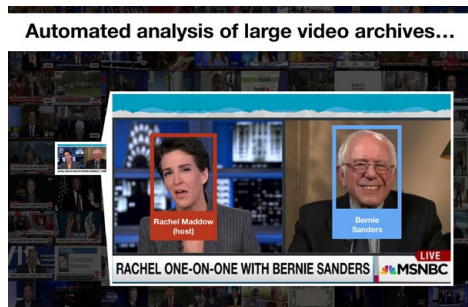
Geraldine is an Emmy award-winning journalist who has worked at both CNN and NBC. In 2019 she was a research fellow of the John S. Knight Foundation in Stanford where she worked alongside Maneesh Agrawala, Director of the Brown Institute for Media Innovation, and Kayvon Fatahalian, Assistant Professor of Computer Science. Together, they set out to develop automated audiovisual analysis for cable TV news.

"We wanted to know if there are ways to automatically analyse television content, so you can look for patterns and trends." They used the Internet Archive to access the audiovisual content of three major 24-hour news channels in the United States: CNN, Fox News and MSNBC. In total, they had access to 200,000 hours of US cable news. They also used transcripts of these news programmes, collected at the TV News Archive.

"THESE KINDS OF TOOLS SHOULD BE AVAILABLE FOR EVERY NEWSDESK, SO YOU CAN REFLECT ON QUESTIONS LIKE 'ARE WE TELLING THE STORIES IN A FAIR WAY?' OR 'DO WE HAVE TOPICS THAT WE OVER-COVER?'" (GERALDINE MORIBA)

Using facial recognition tools, they have identified over 100 public figures about whom they can run any kind of query – for example, how much screen time each presidential candidate had during the 2016 elections and how much that differed between news stations. By combining the transcript data with this analysis, they can further narrow the search, for example regarding how often each candidate was on screen in connection with a specific word, such as 'email'.

The tool uses computer based vision based techniques to determine the identity of well-known public figures like celebrities, politicians, newsmen, and common news contributors. When the identity of a person is not known, the face is classified as male or female. As such, the tool was able to compare hosts and non-hosts of the various stations on various specifics, concluding for example that men dominated on all the stations – with a presence totalling 70%. By using face recognition, some hosts could be excluded from the comparison, resulting in a better gender balance. “We also used this to debunk some beliefs. There is this myth of the typical ‘blonde Fox presenter’. But our analysis showed that it applied just as much to CNN and MSNBC.”



The tool can be used to analyse specific stories and the way they are covered. “These kinds of tools should be available for every newsdesk, so you can reflect on questions like ‘Are we telling the stories in a fair way?’ and ‘Do we have topics that we over-cover?’” The tool does not, however, extend to sentiment analysis. Tests showed it was too difficult to develop an accurate algorithm to correctly indicate specific sentiments, and therefore the team decided not to focus on that right now.

The three news networks which videos are used, are not partners in the project and are not actively participating. Stanford has developed the tool in an academic setting and has used it mainly for historic analysis.

“We are eager to work with newsrooms to see if our tools are useful to them. However, we believe that our tools also stand to have value to media scholars and to people and organizations that act as news watchdogs. Yes, the ultimate measure of success is whether the tools play a part in invoking positive changes newsroom procedures, however the change can result from a number of ways that go beyond a newsroom directly using the tool.”

The tool will be released to researchers and news organisations in the fall of 2019.

FOR MORE INFORMATION

<https://medium.com/jsk-class-of-2019/how-to-identify-bias-in-cable-news-febc6ae0f22e>

HOW TO

UNDERSTANDING PERSONALIZATION

Before you start a personalization project, you should answer some basic questions:

1. What are our journalistic targets, based on our key values?
2. What key metrics for success will we use, based on what we want to optimize?
3. What data is available, and do we have the right tools to collect the data we need?

SEE ALSO: A WAY TO DESIGN METRICS

After that, you can decide on a personalization strategy that will help you hit your targets.

4. Understanding algorithmic approaches

4.1 **Popularity filtering** is based on the idea that if a lot of people find a content item relevant, others may find it relevant, too.

4.2 **Collaborative filtering** predicts a single user's interests based on their similarity to a group which has a user history like theirs.

This approach is based on the idea that users who behaved similarly in the past will behave similarly again and that users will like the same kinds of items as they liked in the past.

This is one of the most commonly used algorithms in the media industry.

Collaborative algorithms can rate:

1. one user's similarity with another user (user-user collaborative filtering) based on users behaviour like usage of content or ratings
2. the similarity of one content item with another content item (item-item collaborative filtering) based on usage or ratings

These algorithms work even if content items are not machine analysable. This means they can accurately recommend complex items such as movies without actually understanding them. The only data it needs to function is consumption data that connects users to content items.

Item-to-item filtering, created 20 years ago by Amazon, tends to be more stable in cases when there are more users than items.

The challenge with this kind of algorithm is the 'cold start' when the system has no previous data on a user or a content item.

There are two types of cold start:

Visitor Cold Start: if the system has no history of a user, it doesn't know their preferences. Instead of using a collaborative-filtering algorithm, the system usually just offers them popular content.

Product Cold Start: When there is no interaction data on a new content item, content-based recommendations can be used instead.

4.3 Content-based recommendations analyse content items to offer you content similar to what you've consumed or liked in the past.

Content-based filtering methods are based on a description of the item and a user preference profile. This profile may be based on user behaviour or explicit choices.

The drawback to this kind of algorithm is that it can only recommend types of content that you have already consumed or shown interest in.

User profiles can be created temporarily, so this method doesn't need users to sign in to anything for it to work – although that does improve the quality of the recommendations.

4.4 Contextual recommendations detect use context from data and recommend content that the system thinks will be relevant in that context.

Most successful recommender systems do this by combining different data to predict what use context the user is in.

This approach helps you to avoid upsetting users by pushing recommendations in certain contexts (e.g. early in the morning, late at night).

This approach can be used to enhance other kinds of recommendations. For example, recommending news articles based on a user's browsing history is helpful, but it would be even more helpful in combination with other data that predict the user's interests and needs. Therefore, most content-based recommender systems now use some form of hybrid system.

4.5 Public service algorithms recommend content because consuming it is believed to serve the public interest.

A successful public-service personalization strategy often combines several objectives.

- There has to be a certain amount of diversity in recommendations so users don't see the same thing all the time, although sometimes recommendations have to be shown several times.
- Typically, there also has to be some serendipity, because recommendations that are too obvious recommendations aren't interesting.
- Possible objectives for public service algorithms include diversity of content, novelty, diversity of opinion, cultural diversity, and ability to create cohesion and common ground.

5. Collecting the data you need

5.1 Analyse the types of data you need

- Content data
- User data
- Data about use context (usually a combination of data on the user, device, location, content and interactions)

5.2 Collect the data

Examples of explicit data collection

- Asking a user to rate an item
- Asking a user to search for something
- Asking a user to give feedback
- Asking a user to choose between two items

Examples of implicit data collection

- Tracking history, clicks, user behaviour
- Tracking the items that a user views or reads
- Tracking how long a user stays on an item and if they finish the item
- Getting data from third-party sources
- Tracking facial activity, vocal expression, choice of words, eye movement or emotional state

Implicit data collection involves multiple ethical concerns.

6. Choosing a segmentation or targeting strategy

Recommendations can be:

- Tailored to each individual user
- Matched to a group of users that share several features
- The same for all users

HOW TO

NEWS ROBOTS

The earliest news robots were used by news agencies to put together stories on companies' financial results. For financial news agencies, where speed and right numbers are key, the robots provided a clear benefit: they could produce the texts much faster and more accurately than human journalists.

If you're not familiar with news robots, a good first step might be to what other news organizations have already done. Here are some examples of robots already in use:

- [Bloomberg Cyborg](#) transforms any company's earnings results into a "mini-wrap with all the numbers and a lot of context"
- [Reuters Lynx Insight](#) generates parts of a story for journalists to edit
- [Associated Press Wordsmith](#) turns earnings data into a publishable article in AP style
- [Washington Post: Heliograf](#) covers elections, local sports and business
- [NOS elections](#) uses a robot to fill in an election template with 76 fields
- [Yahoo! Sports](#) creates hundreds of millions of personalized automated reports for fantasy football players
- [Le Monde and Syllabs](#) automatically presented and commented on the election results of 30,000 municipalities
- [Guardian Australia](#) created its own experimental bot called ReporterMate
- [The Los Angeles Times Quakebot](#) can automatically create and publish an earthquake report in three minutes with some quick editing from a human reporter
- [Yle Voitto](#) follows nightly NHL hockey matches while reporters sleep
- [Press Association RADAR](#) produces text for more than a thousand regional publications in the UK
- [Mittmedia Rosalinda](#) creates local stories on sport, weather, the real estate market and small businesses – very popular on personalized web pages
- [Finnish News Agency STT U-bot](#) is a news robot funded by Google
- [Tamedia's Tobi](#) can write stories using reporter-generated templates.

Some of these robots have been around for more than five years, indicating that news automation is entering a more mature phase. And there's much more.

In March 2019, a group of researchers published an in-depth report on news automation in a report with WAN-IFRA, the World Association of Newspapers. They concluded that previous predictions for news automation had not yet become a reality.¹

¹ (Lindén and Tuulonen 2019)

The report's recommendations (adapted for PSM) include:

- Support from the top

The editorial management team must show support for the automation project, or the newsroom may kill the project.

- Consider combining with product development

Automated content should be connected to new product development. Does it offer more advanced local content? Superiority in a specific topic area? A personalized newsletter? Special sites? Does it push notifications?

- See strengths realistically

Successes in news automation are usually in the area of local content, faster publication and a larger volume of stories, combined with personalized content recommendations. Robots can't yet handle complicated reporting.

- Review tasks and schedules

Machines can take over important tasks that are difficult for humans to do. If you need more content online in the morning, as most newsrooms do, you might consider using machines to process news during the night, such as accidents, sport results, traffic or weather.

BUILD OR BUY?

Robots may be churning out stories in bulk, but technology is becoming a commodity, too.

One of the biggest decisions facing newsrooms that want to move to automated news production systems is whether to buy from a service provider or build a system in house.

It may be that only very large news organizations, such as the biggest public service companies, have the scale to invest in building proprietary AI systems. For others, it may be more feasible to wait until decently priced services or products come on the market.

If you can't build a system yourself, there are alternatives.

First, there are **companies focusing on language processing for automated news production**. Tailoring may be needed to adapt to any languages not previously supported.

Companies that provide technologies for robot journalism:

- <https://narrativescience.com/>
- <https://automatedinsights.com/>
- <http://unitedrobots.ai/>
- <https://urbsmedia.com/>
- <https://www.syllabs.com/en>
- <https://www4.ax-semantic.com/>

Second, **news agencies** with lots of news media clients see a clear opportunity for scale benefits by investing in robots even for general news. Often regional publishers – typically the key client group for national news agencies – complain that readers aren't interested in the national and international stories produced by news agencies. Using public datasets to write automated stories targeting different areas is a clear value proposition for these agencies.

Of course, buying from news agencies does not differentiate you, but perhaps neither does robot journalism, even if performed in house. It is already becoming a crowded area.

The [Press Association of the UK](#) is running a hugely interesting robot journalism project. It employed five reporters to work with templates that turn out thousands of regional and local stories based on public health, traffic, school and real estate data. Because human reporters are heavily involved in the writing process, these stories often end up as top stories in local papers. This project also addresses two key challenges in robot journalism: checking data quality (since the quality of the story depends on the data behind it), and making sure the writing quality is high for front-page stories.

HOW TO

PUBLIC SERVICE

PERSONALIZATION

How can a newsroom design a useful public service media (PSM) algorithm?

How can a newsroom design a useful public service media (PSM) algorithm?

FIRST, UNDERSTAND THE SOCIETAL IMPACT OF ALGORITHMS

Tarleton Gillespie, a senior researcher at Microsoft Research New England and affiliated associate professor at Cornell University, highlighted six aspects of what he calls 'public relevance algorithms' in his article 'The Relevance of Algorithms':¹

1. The choices behind what is included, what is excluded, and how data is made algorithm-ready
2. The implications of algorithm providers' attempts to thoroughly know and predict their users, and how the conclusions they draw can matter
3. The criteria by which algorithms determine what is relevant, and how those criteria enact political choices
4. The promise of algorithmic objectivity: how the technical character of the algorithm is positioned as an assurance of impartiality
5. How users reshape their practices to suit the algorithms they depend on
6. How the algorithmic presentation of publics shape a public's sense of itself

Gillespie's list can be used as a tool to evaluate proposed qualities of public service algorithms.

SECOND, CONSIDER PUBLIC SERVICE OBJECTIVES

Diversity

This is the value most often mentioned when discussing PSM algorithms. However, as professor Natali Helberger notes, it's unclear what diversity means in this specific context.

It's often used to refer to a diversity of supply (source and content), but digital media can also monitor diversity of real exposure and actual consumption. In PSM, diversity can also refer to the broader societal issue of reflecting different cultures and experiences through representing diverse communities.²

Diversity of exposure to programming has also been flagged as a core challenge for PSM automation systems by Danish associate professor Jannick Kirk Sørensen and Dr Jonathon Hutchinson, who is currently a fellow at the 'Algorithmic Public Spheres' project at the Hans-Bredow-Institut for Media Research at the University of Hamburg.³

James Bennett, in his article 'Public Service Algorithms', discusses diversity in the sense of connecting diverse audiences to shared content. This concept is very similar to a value discussed elsewhere as 'cohesion': the ability to gather people together

1 (Gillespie, 2014)

2 (Helberger, 2015; Helberger, Karppinen, & D'Acunto, 2018)

3 (Sørensen & Hutchinson, 2018a)

through shared experiences. Bennett even proposes connecting audiences with external services and content.⁴

Karin van Es, an assistant professor of Media and Culture Studies at Utrecht University, considers recommender systems to be a threat. She sees them as serving customers rather than citizens, and even views interactivity of services as bad for public service media. According to her, while diversity contributes to a varied media diet, it does not produce shared reference points.⁵

Novelty, distinctiveness and serendipity

Another feature of PSM algorithms is that they are designed to balance success metrics and metrics linked to the quality of the content.

Bennett suggests using metrics that 'optimize for new experiences' and would require systems to connect audiences with new content and forms.⁶

Sørensen and Hutchinson propose balancing popularity with distinctiveness, while Van Es adds that an effort should be made to either boost or minimize minority or controversial viewpoints.⁷

Serendipity is already being used as a design principle for commercial media, as it has been found to increase audience satisfaction with recommendations.

Transparency

Another widely mentioned objective for PSM algorithms is to make the logic behind the recommendations transparent, in order to build trust and accountability. Sørensen and Hutchinson include transparency in their list, while Van Es would require genuine accountability for algorithms to maintain trust.

A group of BBC researchers involved with a radio recommendation algorithm project based on public service values found that users appreciated transparency and the ability to give feedback and adapt the recommendations. Although there has been no conclusive evidence proving the existence of filter bubbles, this continues to be a thorny issue. Participants in the BBC study voiced concerns about a lack of diversity in existing personalized recommendation systems because of such bubbles and feared they would not see potentially relevant content.⁸

Transparency is also closely connected with the ideas of impartiality and fairness.

Sovereignty

Sørensen and Hutchinson outline two kinds of sovereignty as important for PSM algorithms: user sovereignty (the ability to make decisions about what is recommended and how) and independence from commercial intermediaries. Van Es repeats Bennett's call for PSM to use different performance indicators than commercial media.⁹

THIRD: MAKE IT HAPPEN

At the BBC, much research has been done on creating an algorithm which optimizes for the public interest.

Ben Fields, Rhianne Jones and Tim Cowlshaw explored this issue in their article 'The Case for Public Service Recommender Algorithms',¹⁰ much of which revolves around the core questions of ethics, transparency and accountability in algorithm design.

However, actually creating an algorithm around these values may be difficult. As Fields, Jones and Cowlshaw note: "Whilst some progress has been made, core challenges remain around the implementation of these values and ideals in practice and importantly how to determine 'success'."

However, the researchers think that diversity metrics in addition to serendipity, novelty and distinctiveness are promising avenues for further research: "Diversity in recommendation also affords serendipity or surprise in encountering new and unusual content which one would not otherwise have encountered, which was an experience

4 (Bennett, 2018)

5 (van Es, 2017)

6 (Bennett, 2018)

7 (Sørensen & Hutchinson, 2018b; van Es, 2017)

8 (Cowlshaw et al., 2018)

9 (Bennett, 2018)

10 (Fields, Jones, & Cowlshaw, 2018)

frequently cited as being a particularly pleasant aspect of listening to broadcast radio.”

Their list of practical questions on this issue includes:

- What are useful metrics for which to optimize (e.g. diversity or serendipity)
- How should the importance of different metrics be balanced in different PSM contexts?
- How much accuracy loss is acceptable in pursuit of new metrics, e.g. diversity?
- How should transparency work – when and to whom is it useful, e.g. regulators?
- To what extent should we be transparent about how we are resolving metric and optimization complexity (the trade-offs we are making)?
- How do we design for interpretability and explainability to enable appropriate oversight of how recommenders are making decisions and ensure due accountability?
- What type/level of explanation will be most useful?
- Will explanations produced for editorial need to vary from the type of explanations PSM may provide to audiences?
- How will we determine the value of different potential approaches?

CHAPTER 05

FIND THE NEW PUBLIC SERVICE IN AN ERA OF AI

DEFENDING CORE DEMOCRATIC VALUES IS AN ETHICAL RESPONSIBILITY FOR PUBLIC SERVICE MEDIA

We are seeing more and more ways in which new technologies are transforming societies, professions and even the distribution of power inside democracies. In conjunction, the PSM mission of journalism is evolving, for example by holding those in power to account or bringing the nation – or even many nations – together to foster common understandings and public debates. The values we choose to express through our journalism, including how we use new technologies, are increasingly defining the relationship between news organizations and their audiences.

Is journalism in crisis? If so, it may have been that way for decades. In the past glory days of mass media and high-end journalism, the profession had the power to dictate the tone of the public debate by setting the agenda for public discourse. Those days are gone. In fact, they were gone even before automation took over.

“I like to say that journalism as we know it, the journalism that governs the work of modern professional reporters and editors, is not much older than I am, which makes it roughly the same age as satellites and rock and roll,” notes John Nerone, now an emeritus professor of communication. “When the ‘crisis’ came, of course, everyone blamed the internet, but this neglects just how deep the problems of hegemonic journalism have run.”

All technologies embed and encode human values, reflecting human choices. AI systems are tools built by humans to serve human means and ends. And it can be argued that all the decisions made in building them reflect the values of their designers.

“FOR OTHER COMPANIES, THE DATA ABOUT YOUR AUDIENCE IS MONEY AND A WAY TO CREATE AI. FOR US, THE DATA IS ABOUT TRUST. PEOPLE TRUST THAT WE WORK WITH DATA IN A VERY RESPONSIBLE WAY AND KEEP IT FOR OURSELVES. SO, IT’S A TRUST ISSUE BETWEEN THEM AND US.” (HANNA STJÄRNE, CEO OF SWEDISH TELEVISION SVT)

People define their relationship with a news brand increasingly through the values it conveys. Especially in news, the relationship is rarely purely transactional, based only on the question 'is it worth the price'. It is more an expression of shared values and building one's identity, based on the statement 'this represents something I believe in'.

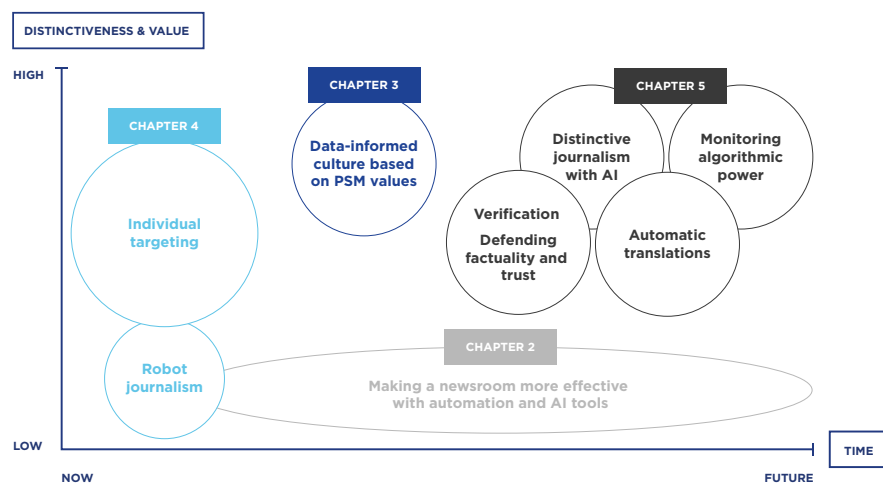
Therefore, there are many fundamental value-based ethical questions related to the use of AI in society and in news.

Taking the EBU General Assembly decision on core values in 2012 and the editorial principles derived from them as our starting points, we can then ask what values and principles such as universality, diversity, impartiality and fairness in journalism really mean to audiences, and how they should be expressed in everything we do.

Do algorithms change the way PSM understands its role? How does it relate to public service values? What is the strategic use of artificial intelligence for the newsrooms of EBU Members?

Starting from very basic rules on strategic management, PSM has to – in order to fulfil its mission and stay relevant – be valuable and distinctive. In other words, it has to be appreciably different. And that is especially true in these data-informed times.

We see several ways forward to make public service journalism distinctive in the new era. These are future-oriented ideas that can be seen as supplementing the worthwhile AI applications discussed in chapter 2 (practical AI tools) and chapter 4 (targeting and robot journalism), 'which are more "here and now", as illustrated in the following diagram:



What could be the real distinctive value that new technologies add to public service journalism in the future?

First, AI can be a new source of idea creation for journalism with a high degree of public relevance. AI-aided data journalism can detect important stories and provide a rich source of distinctive content.

Second, as more and more decisions in our societies are being made using algorithms and data, a new interesting field for investigative journalism has arisen to scrutinize this form of decision-making and persuasion, which sometimes can even be classed as 'manipulation'. By investigating how the algorithms have been designed and how they function in decision-making processes, journalists can hold them and their creators to account.

Third, truth is under attack from false news, misinformation and deepfakes, created either automatically or manually. These phenomena challenge how people perceive evidence. As we know, for societies to function, there is a huge need for trust-building and common ground.

Fourth, AI offers incredible new opportunities to help people understand the world's diversity. One example is making language barriers a thing of the past.

And fifth, AI can help newsrooms diversify between the content they produce and which audiences consume, making both the producers and the consumers of the content less prone to staying in bubbles of their own beliefs.

LET'S REMIND OURSELVES ABOUT ETHICS TOO

Before we can do anything, however, we must stop using public service values as an excuse not to use AI. In ethical discussions, fears tend to dominate, which is only natural. But PSM should remember that doing good to society and individuals is an ethical responsibility as well. "The virtuous journalistic public service media culture says: Our journalism is human, it has human values where people are serving democracy. I think that can be used as a kind of excuse to say: Let's not trust the algorithms," says Charlie Beckett of LSE.

As journalists develop skills to monitor algorithmic decision-making, it's important that they are responsible in the use of algorithms themselves and trustworthy and respectful in the use of personal data. PSM organizations are keenly focused on accountability, and AI can assist in being more transparent about choices. At the same time, however, AI can conceal information about the origins of a story. Any decisions made by machines need to be clearly communicated.

In the guidelines on AI ethics proposed by the EU High-Level Expert Group, about half of the principles overlap with public service values: an informed citizenry, diversity, fairness, excellence and accountability. This takes on a concrete meaning in the new digital context, as ethics always does when applied, but in the end it's the same ethics as elsewhere. Hanna Stjärne, CEO of Swedish Television SVT says: "For other companies, data about your audience is money and a way to create AI. For us, the data is about trust. People trust that we work with data in a very responsible way and keep it for ourselves. So, it's a trust issue between them and us."

To use AI ethically, you should adhere to the following five points:

- Reinforce public service values and apply them in the algorithmic era.
- Consider the individual user and keep your promises.
- Be transparent: communicate whenever you use AI.
- Educate the newsroom on the ethics of AI.
- Consider value-based threats like filter bubbles and weakening societal cohesion.

FOR MORE DETAILS, CHECK OUR CHECKLIST BASICS OF AI ETHICS.

AI ENABLES MORE DISTINCTIVE AND NEW TYPES OF JOURNALISM

Artificial intelligence can help find hidden stories. With machine learning and data science, deep investigations into data can be conducted. "Data journalism is the practice of finding stories in numbers and using numbers to tell stories," says New York University professor and data journalist Meredith Broussard in her book 'Artificial Unintelligence'. She sees a new development in this field: from 'simple' data journalism, to data journalism aided by artificial intelligence.

Stories relying on huge datasets that are impossible for humans to scan and analyse by hand are where the true value of data science and machine learning lies for public service investigative journalism.

With the Panama papers and Bellingcat's investigations, interest for this kind of original reporting has grown. It seems that more and more newsrooms are delving into data to unearth stories. Tom van de Weghe, a Stanford fellow and reporter at VRT, is convinced that with machine learning and data science, "there will be more stories and we get faster funding and better results". Professor Philip Napoli of Duke University agrees: "It's driving a lot of the experimentation that's happening."

Nicholas Diakopoulos, Director of Northwestern University's Computational Journalism Lab: AI approaches enable more original and unique storytelling. Stories you would not have found otherwise. Because the internet is an environment that requires everybody to differentiate, AI is a means to differentiate in several ways. What we're going to want is more diversity in the AI that's being used in journalism," he predicts.

In using data to make sense of the world, we have all kinds of bias to deal with. The obvious risk is to focus only on areas where data can be obtained: “We need to have journalists on the lookout for gaps, for blind spots, for places where there are no data,” noted a group of researchers in a Tow Center report on AI and Journalism in 2017. This risk even has a name of its own: data bias.

And yes, most investigations in huge datasets are expensive, take a lot of time, are difficult to plan, are unique and have a high fail risk. Research by six Norwegian newsrooms concluded: “We find little evidence that computational journalism increases the efficiency of doing journalism.” Newsworthiness is difficult to code. But equally, these large-scale investigations often have huge societal value and therefore are clearly relevant to the remit of public service media.

IN OUR CHECKLIST ON ALGORITHMIC JOURNALISM WE DESCRIBE MORE ON THE ISSUE.

Because of the scale and costs involved, large investigations on datasets are often conducted in collaboration with researchers and NGOs, like with the projects of Forensic Architecture and OCCRP.

ALGORITHMIC ACCOUNTABILITY: A NEW WAY OF MONITORING THE POWER BROKERS

New technologies are transforming many areas of societal life and decision-making. That sheds light on a totally new area that journalism should monitor. What is the power of algorithms that are making decisions? Do we understand them? What are their shortcomings, and who designed them that way?

Therefore the next evolutionary step in AI-aided data journalism will be algorithmic accountability: the reporting on the algorithms and data that are used when computers make decisions about us. Who, for example, decides whether you are creditworthy, and how is this risk calculated? How is a medical treatment designed? And how is social housing allocated to new residents? Who influences the decision-making in elections?

Reporting on these algorithms is especially important when the so-called ‘black box’ dilemma is in action, i.e. even those who designed a deep-learning model don’t know how it reaches its conclusions because of the many superposed layers of variables and decision-making structures. In addition, many algorithms are trained with biased data. Even the smallest bias can have detrimental effects when unnoticed.

Journalists investigating algorithmic accountability ask questions like: What are the impacts of that algorithm for the public? Does the algorithm have the potential for discrimination? Do errors arising from the algorithm create risks for the public? Is the algorithm distorting the expected outcome? If so, why? This line of investigation aims to reveal and debunk fake data and biased algorithms and the biased decision-making that originates from both, especially when used by government authorities, large corporations and even the media itself.

Algorithmics can also be used to reverse-engineer human biases. Same questions can be asked about humans, with the help of computers: What kind of discrimination or biases does the human decision-making have? Why are court decisions different in cases that are similar, systematically?

As with data journalism focusing on big datasets, this kind of journalism needs resources. For one, it needs people, and so far not many capable of doing this type of work. Unsurprisingly, these people are in high demand. Uli Köppen: “Good data scientists are rare and wanted, which means they can earn a lot of money at big companies. That you have to balance with your mission and public value.”

It may be that only the large news organizations are able to invest what’s needed. Mixed teams with both technical skills and journalistic skills may be needed. On the other hand, the payoff will be original reporting and providing a key watchdog role for journalism in a field that is becoming extremely relevant for societies.

Algorithmic journalism is much more developed in the US than in Europe. One of the few examples of far-reaching use of machine learning and data science in data journalism comes from BR in Germany.

STRENGTHENING THE CORE TASK OF JOURNALISM: VERIFICATION

At the very core of the journalistic process, data and AI are being used for verification at several organizations. Automatic verification is closely related to automatic fake news detection, and both share the same challenges if they are tried to be done without human involvement. Technology alone can't solve the issue, but works best when combined with human effort.

At City University on London, the research project DMINR aims to develop a web-based tool to assist in verification in a way that is easy to use and – in their own words – ‘sympathetic’ to the way journalists work. The project's tool has been developed and tested in collaboration with dozens of newsrooms in Europe.

The BBC, among others, is joining forces with technology firms in fighting fake news, for example by using an early-warning system. And the European collaborative project InVID has brought together a consortium of nine partners to collectively find new solutions.

Deutsche Welle is another public service organization developing tools for verification. The system they have created in collaboration with various technology partners is also used by organisations like Amnesty International. In their deep fake detection attempts they are focussing in audio, partly because of its relevance but also because there are fewer bytes than in video and therefore it's cheaper to process and analyse with computers.

“We dive into the DNA of the audio. We are looking at electricity frequency patterns, for example, which allows us to maybe locate the video to a certain area”, says Ruben Bouwmeester, Innovation Manager at DW.

Sometimes the goal may be to educate audiences about the new social-media reality and how information works therein. ‘Troll factory’, a game developed by Finnish Yle, lets you play a nasty troll, using authentic social-media content. It helps social-media users understand their role in spreading false information through ‘likes’ and ‘shares’, potentially fuelling fear and suspicion. The ultimate aim is to promote digital-media literacy.

IN THE ERA OF DEEPPAKES, THE LIARS HAVE AN ADVANTAGE

Take the advanced language technologies out there today, produce a synthetic voice imitating someone famous and add some very advanced neural networks technology to swap the faces around. This is a deepfake. The most famous example of this was a video of President Obama altered to deliver a fake message. And this is not the only threat: journalists should be prepared that those in power will label reliable witness material as fake. Deepfake attacks are even being constructed against journalists themselves. In India, for example, a journalist covering corruption became the target of rape threats and had to go offline for months.

Factbox: What is deepfake?

Deepfake (‘deep learning’ and ‘fake’) is an AI technique to alter video or audio so that it realistically shows something that didn't occur in reality. Because of advanced video-processing technologies that use deep neural networks, fake celebrity pornographic videos, revenge porn as well as fake news and hoaxes, in which decision-makers or journalists appear to say or do something they actually didn't, can be produced quickly and easily.

For journalistic organizations, verification is a key capability. Making sure this capability is up to date is a strategic priority, and that is not easy. Although even venture capitalists are pouring money into AI to detect deepfakes themselves created by AI, most experts don't believe technology alone will deliver the final solutions. Those doing the faking are not bound by the same ethical rules as journalists, academics or even the intelligence services trying to hunt them down. This is called the “liar's dividend”.

"I'm afraid that one day we will have a digital wildfire in which a tsunami of fake messages is spread. That can happen in a crisis or before elections. Some media will make mistakes, and that can destroy some trust," says Tom van de Weghe.

But the most dangerous point of the deepfake phenomenon for journalism may be that audiences don't believe their eyes – or ears – anymore. Tom van de Weghe: "As long as there is no real message in a deepfake video, you can think of it what you want. But as soon as there's a believable message created, then there's a big problem." Transparency therefore becomes an even bigger focus. "In a deepfake world we need to be showing exactly why we reached the conclusion that we've reached," says Sam Dubberley, Investigator at Amnesty International's Evidence Lab.

In order to fight the rise of deepfake videos and audios, several universities and start-ups are developing technologies to detect them. The US universities University of California-Berkeley and University of Southern California have analysed certain individuals' style of speech and created an algorithmically detectable 'soft biometric signature' for top politicians. They argue that the very technology used to create a fake is the one used to detect it.

The best way to prepare is to combine automated analysis with human perception. Humans are much better than computers in detecting nuances such as humour or satire and can run a manual check or make a phone call, if needed. There's a small bright light in scepticism, though, as this scepticism might actually be healthy. People have become much pickier about whom they trust. And journalists are learning about deepfakes as well. "There's no need to panic, as there is time to prepare," says Sam Dubberley.

US-based non-profit Witness.org has organized events and drawn up proposals about how to best prepare for these threats in news organizations.

TEAR DOWN LINGUISTIC WALLS WITH AUTOMATED TRANSLATING

One of the most promising areas of automation is language technologies, from understanding to generating language. Journalists encounter these technologies through tools like automatic transcription, which transfers recorded interviews from speech to text. We, as authors of this report, have used these tools. AI-based proofreading is another example (which we used too). These tools are developing fast, especially in major languages, as people are eager to use them: they help to get rid of the boring work and free up time for inspiration.

Currently, the technologies are rarely able to produce fine-tuned translations but these are good enough for monitoring purposes. And it's much faster to edit automatic raw-data translations than to start from scratch. When the automatic translations are good enough to be published for the public, that will represent a significant advance. Such developments could potentially allow all the news content of EBU Members to be available in all languages, which would be an exciting prospect.

Because major technology giants are investing enormously in language technologies, there should be significant developments in just a couple of years in English, German, French, Italian and Spanish. For smaller languages, public service media have a significant role to play in making these services available, as part of both their remit and their duty to serve minorities in their communities.

The BBC, Deutsche Welle and SRG-SSR in Switzerland are among the public broadcasters working intensively with language technologies. Automatic translations are becoming the norm and offer considerable benefits for monitoring what's going on the world, or simply what's going on inside a large multilingual broadcaster.

CASE 24: Automatic translations in the BBC and DW

RECOMMENDATIONS FROM THIS CHAPTER

1. AI-aided investigative journalism is an effective way of finding relevant issues to cover in society. The result may be both relevant and distinctive, as this type of journalism requires special, rare skills.
2. The growing use of algorithms in societies requires a fuller understanding of their impact. The media has an important role to play in holding algorithms to account.
3. Misinformation and deepfakes represent a huge risk for truthful reporting. As verification lies at the core of news journalism, organizations should be equipped with both the technology and knowledge needed to separate the facts from the lies, in an environment where the human eye or ear can no longer differentiate fake video or audio from the real thing.
4. Deepfake videos and audios have the potential to erode trust in the witness materials used by newsrooms. They are even used in attacks against journalists. News organizations should be prepared for when this happens.
5. In multilingual Europe, automatic transcription and translation technologies combined with synthetic voice generation are giving rise to interesting new opportunities for delivering news to new audiences and to minorities who previously had only limited access to content in their language.



OCCRP, BOSNIA HERZEGOVINA

USING MACHINES TO FIGHT CORRUPTION



Emma Prest
Chief Technology Officer
emma@occrp.org

TAKEAWAYS

- It's better to talk about specific techniques than to use the generic 'artificial intelligence', as this can mean many different things.

- When there is a clear societal advantage, it's relatively easy to obtain interesting datasets. To make these accessible and searchable is the real challenge.

- For journalists it's ideal to work in close cooperation with organizations like OCCRP, which have the knowledge and the data to build data-search engines. This way each can focus on their own strengths and goals.

What do you do with the inbox of a politician you suspect is corrupt? How do you make sense of a set of bank transfers that you think might be suspicious? One option is to send this data to the OCCRP, who use a 'leak ingestion machine' to upload all of your content to a huge database that is searchable and makes connections with the other 2 billion records they have.

With its headquarters based in Sarajevo, the Organized Crime and Corruption Reporting Project is a non-profit media organization that uses technology-based approaches to expose organized crime and corruption worldwide. They have 45 investigative centres in 34 countries and in 2018 alone produced more than 90 cross-border investigations. Their aim is 'to give citizens and governments the information and tools they need to bring about a fair system in which criminality and injustice are fought with transparency, knowledge, and empowerment'. Historically, their focus has been on stories in Eastern Europe and the Balkans, but now their scope is worldwide.



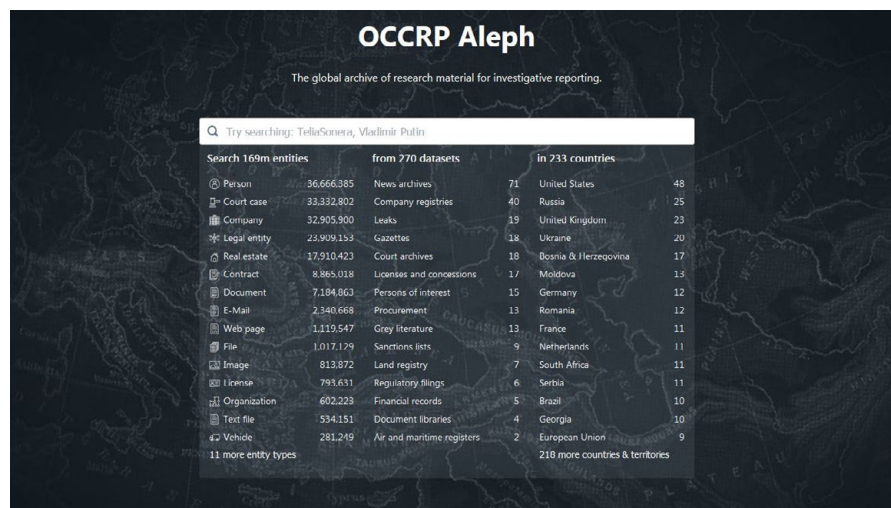
Among their recent projects is the Troika Laundromat. This is an investigation to a financial money-laundering scheme that channeled billions of dollars out of Russia and allowed Russian oligarchs and politicians to secretly acquire shares in state-owned companies, to buy real estate both in Russia and abroad, to purchase luxury yachts and much more. The investigation led to a [decrease in the share price](#) of various involved European banks and has been used by those [pushing for a new EU anti-money laundering agency](#).

"THE AMOUNT OF DATA WE HAVE IS A DATA SCIENTIST'S PARADISE." (EMMA PREST)

Yet the OCCRP is surprisingly little known in the West. "I think this is mainly a geographical issue," says Emma Prest, Chief Technological Officer at OCCRP. "Many people in the West aren't familiar with the political dynamics in places like Ukraine, Serbia or Kazakhstan. It's our job to expose large-scale crimes like money laundering, but also to explain how it affects people across various regions and countries." The

investigations are also not the easiest to cover: “It’s hard to make good stories on offshore banking and money laundering, as most of it is rather abstract and can be difficult to understand.”

Tech is at the core of the OCCRP mission and they are often mentioned as a leading organization in the way they apply artificial intelligence.¹ Yet Emma would not use this word herself: “At OCCRP we’re doing sophisticated data work, but no artificial intelligence.” Whatever you call it, the work OCCRP is doing is very impressive.



Emma has been working with OCCRP for 6 months now and is impressed herself with the dataset that the organization is dealing with, consisting of around 2 billion records. “The amount of data we have is a data scientist’s paradise.” The organization obtains the data mainly through leaks and scraping. “Every month we get new troves of data as big as the Panama Papers in. And our main focus is to make this data accessible to journalists.”

And they are especially good at that. On the OCCRP Aleph platform, for example, journalists can browse all documents in the OCCRP database, searching for emails, documents and web pages mentioning a specific person, for example. “It does two things. For an external user, it’s just like a search platform with datasets behind it. You can pull out any kind of record on people or companies, especially for ‘follow the money’ style investigations. But it’s also a ‘leak ingestion engine’. We have built a data model that categorizes all the information that comes in – bank transactions, email accounts and company servers. The information then gets mapped onto our data structure, so you can make connections with other types of data in the datastore.” There is a separate dashboard for bank transaction data, and another one for flight tracking. “These are all built in house, and are open source.” This gives the opportunity for external users to verify or challenge the veracity of the content.

To improve the search functionality, one of the gnarly problems OCCRP is now working on is suggesting journalistic leads. “When you are searching for a specific entity, how can you be sure the one you find is the correct one, the same one you are looking for? We are exploring a system that will give you a probability score on how likely it is that it’s the same entity. But always as a suggestion to a journalist, we never want to automate that kind of thing.” For that, context understanding is needed, and this is very much a human trait. Therefore the last judgement call is always made by the journalist using the data.

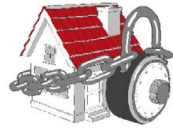
One of the answers lies in their own community. “We have thousands of journalists using our platform. If they could be reporting back to us confirming that this is indeed the right entity, they can train the data and thus help us.”

Emma also remarks that techniques associated with artificial intelligence, like deep learning, could be relevant but they’re not there yet. “Normally a journalist starts with a hunch or a lead and then begins their investigation. In our case, we often start with a large dataset and need to work out what the question is. What are we even looking for here? It’s a different way of doing investigations, and one where data science techniques could help to identify relevant documents and leads faster and better.”

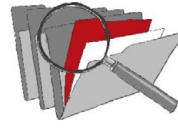
¹ <http://jonathanstray.com/papers/Making%20Artificial%20Intelligence%20Work%20for%20Investigative%20Journalism.pdf>

The hybrid nature of the organization, with both journalists and data scientists working side by side, make the OCCRP a leading player in embedding technological innovations in journalistic investigations.

OCCRP's Growing Impact



Over **US\$ 6,300,000,000** in fines levied, assets frozen or seized by governments



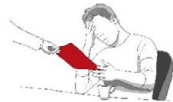
More than **290 criminal investigations** and government inquiries launched as a result of its stories



More than **330 calls for action** by civil, public or international bodies



More than **380 arrest warrants** issued



More than **45 major sackings**, including a president, a prime minister, and CEOs of major international corporations



More than **1,400 company closures, indictments and court decisions**

FOR MORE INFORMATION

<https://aleph.occrp.org>

www.occrp.org

<https://www.occrp.org/en/troikalaundromat/>

SIMILAR ORGANISATIONS/PROJECTS

International Consortium of Investigative Journalists: <https://www.icij.org>

Forensic Architecture: <https://forensic-architecture.org>



FORENSIC ARCHITECTURE, UNITED KINGDOM

SCANNING DATA TO FIND HUMAN RIGHTS VIOLATIONS



Nicholas Zembashi
nz@forensic-architecture.org

TAKEAWAYS

- Machines are useful in determining correlations, but when high accuracy and moral judgement is at stake (as in this case, when collecting evidence), it's wise not to let the machines make the decisions.

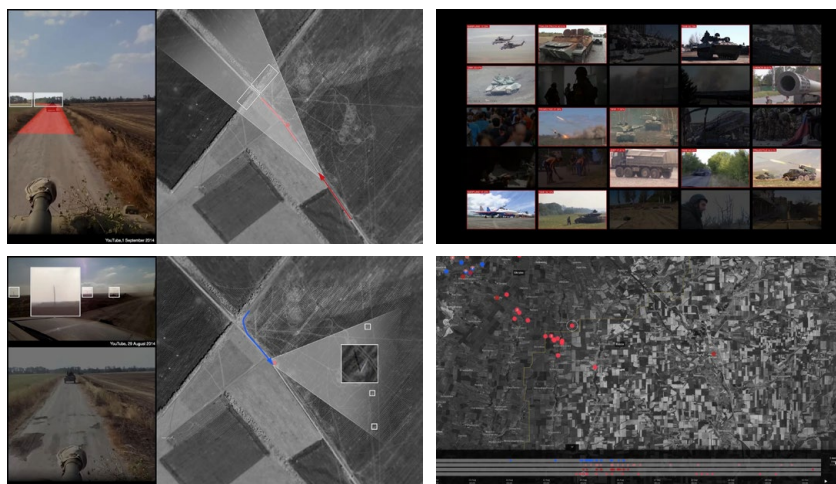
- Mathematical models can be seen as team members with specific talents and, likewise, shortcomings. Their role and decision-making authority should be based on these considerations.

- By being transparent and working with open-source technology, the process of verification is demystified for the audience. Artificial intelligence is not showcased as a magic wand but as a tool that also has shortcomings.

Using computer vision classifiers as part of its investigative toolkit gave Forensic Architecture the ability to scan large amounts of data for possible evidence of human-rights violations. In a project commissioned by the European Human Rights Advocacy Centre (EHRAC) and the Ukrainian Legal Advisory Group (ULAG), Forensic Architecture was asked to gather available evidence to verify the presence of the Russian military in Eastern Ukraine, in support of a legal claim against the Russian and Ukrainian states.

How can you prove something happened? Although once it would have sufficed to show a single picture or video, in an era of deepfakes and digital manipulation, this is not always enough. Therefore the London-based research agency Forensic Architecture was asked to verify Russian military presence in Eastern Ukraine in 2014 by providing a spatial and time-based representation of accounts. They were able to gather evidence by using computer vision classifiers as research team members and present their findings alongside existing open-source evidence in an online interactive platform.

"The work that classifier has done for us would have been too time-consuming and difficult for us humans to do," explains Nicholas Zembashi, project coordinator of the 'Battle of Ilovaisk'. "We used a pre-trained image classification model to scan thousands of hours of video material and locate Russian military equipment on Ukrainian soil."





Their data set comprised of a of image frames extracted from YouTube videos focused on a specific region and within a certain time frame. They then let loose a classifier that checked the images for specific military hardware, including tanks. The model made a comparisons between the image it was looking for and the images it saw in the video frames. It then determined the likelihood of the correlation: is this indeed a tank? Each hit received a probability score and was sent to the researchers for a final check.

“The Battle of Ilovaisk was our first project to which machine learning was applied, and for this we used a pre-trained classifier. But this kind of deep learning requires a training set composed of thousands of annotated images. Not only is this costly and laborious, but building such a set also relies on there being enough images available in the first place and that they are appropriately licensed. As this is not always the case, we focused on the use of synthetic data to train a classifier in the next project to which we applied machine learning, the Triple Chaser Investigation.”

Nicholas sees the classifier as a team member with exceptional processing skills, but very poor moral judgement. “We know that the classifier is great in showing correlations, but we also know it’s really bad at inferring causations and therefore making decisions. For these reasons, we never completely relied on the machine. The humans were the active agents.” The humans working on the investigation were specialist researchers (specialising in various topics), journalists, lawyers and architects, each with their own expertise in evaluating the images.

“WE WANT TO DEMYSTIFY AI FOR OUR AUDIENCE - IT’S NOT SOME KIND OF MAGIC, IT’S A TOOL WITH INNER WORKINGS THAT CAN BE EXPLAINED. WE ALSO ACKNOWLEDGE ITS LIMITATIONS AND PROBLEMATICS AND THEREFORE ENSURE THAT HUMAN RESEARCHERS ARE AUGMENTED EVERY STEP OF THE WAY TO MAKE THE ACTUAL DECISIONS.”
(NICHOLAS ZEMBASHI)

The team presented all of their data on an interactive, cartographic platform, rendering a visual and spatial representation of events. On a map of the region, one can trace back in time events related to Russian military presence and the whereabouts of military hardware. The map has different layers of information and is currently being used as evidence in a case against the Russian and Ukrainian states brought before the European Court of Human Rights.

“What this tool can do is provide a constellation of events, pinpointed as precisely as possible in time and space. It’s a synthesis of evidence. An agent with an agenda, like a government, can always argue that there is a different truth, but with this volume of connected evidence, they will have a hard time.”

Nicholas does not see artificial intelligence as something objective or as something that will provide ‘the truth’: “The truth is something that is constructed in dialogue, in civic society. It’s a collective process. A machine can’t supply the truth, no matter the level of abstraction and obscurity, as an outcome doesn’t mean anything unless we collectively give it meaning.”



Machine-learning algorithms have their own biases and often function as black boxes, making it hard to understand how they come to certain conclusions. “They are not objective tools, although people often think they are. They are, at some point, developed by humans, who are biased. Someone has, for example, decided what is a tank, and what isn’t. Also, in our case, the model evaluated images, and images are never objective. So a tool is quite often a compilation of black boxes – it will not lead to One Truth.”

The team actively discusses the possible biases that are in the material. They also make sure that they articulate these biases clearly to a wider audience. Their focus is on open verification and the open sourcing of the evidence. “We want to demystify AI for our audience - it’s not some kind of magic, it’s a tool with inner workings that can be explained. We also acknowledge its limitations and problematics and therefore ensure that human researchers are augmented every step of the way to make the actual decisions.”

The project team consisted of between 4 and 12 people, and it took about 6 months to make a case suitable to take to court. After its submission to court, the platform was adapted for public release and a short film describing the project was released to accompany it. “We want to find the best forum for each of our cases. Sometimes we do this through journalism, sometimes through court while also making an exhibition out of the material.”

FOR MORE INFORMATION

<https://ilovaisk.forensic-architecture.org>

<https://forensic-architecture.org/investigation/the-battle-of-ilovaisk>



BR, GERMANY

HOLDING AI TO ACCOUNT

Ulrike Köppen
head of BR Data
Ulrike.Koeppen@br.de

TAKEAWAYS

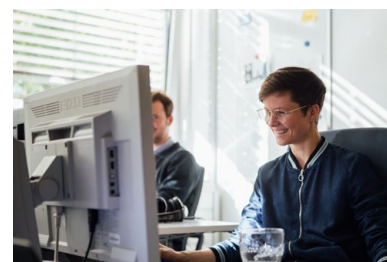
- Even though AI is not yet omnipresent in your society, chances are it will be in the near future. Being capable of reporting on AI and holding those controlling it to account also means starting to use AI and algorithms yourself.

- When developing new tools and algorithms, it's worth sharing them openly with other newsrooms, especially as many are doing so themselves as well.

- More than anything else, you need the right people and the right data-savvy culture to validate the opportunities arising from AI.

What if a scoring company calculating your creditworthiness for a bank loan is saying that you are a risk case, based on a faulty algorithm? The data journalism department at Bayerischer Rundfunk (BR) is developing a new skill set in machine learning and artificial intelligence to achieve a better understanding of the power and influence of algorithms – and hold the organizations that use them to account.

Bayerischer Rundfunk understands the power of artificial intelligence: that it can be used for great journalism but can also harm society. “We have to use machine learning and algorithms in order to make better stories, but also to gain a clearer understanding about how they work. In Europe, we might not yet depend on algorithms as much as in the US, but they are our future. We need to understand the principles behind them and challenge their accountability as early as we can,” says Uli Köppen.



Uli is the head of BR Data, the data journalism department at Bayerischer Rundfunk, and a 2019 Nieman Fellow at Harvard. Her team, consisting of 9 people, is using machine learning to conduct investigations they could not do otherwise. “Like in our project on political trends, here we used a training dataset, tagged by scientists, and an algorithm that was trained to detect liberal or conservative words in political manifestos. The algorithm translated the count of a particular manifesto into a score. We used this algorithm to analyse the manifestos that were drafted for the elections in 2019 and the ones of previous elections. As a result we obtained a certain score per

“WE CONSCIOUSLY SHARE OUR KNOWLEDGE BY PUBLISHING OUR METHODS AND DATASETS, WHENEVER POSSIBLE. WE REALLY HOPE PEOPLE WILL USE IT.” (ULRIKE KÖPPEN)

party, per year. This showed the trends and also that three parties moved together on this axis ‘liberal versus conservative’. Those parties actually were the first to sit down for coalition talks. And they probably would have succeeded if the human politicians hadn’t messed up the actual formation...”

By using algorithms and machine learning themselves, BR Data is also acquiring some much-needed skills to hold other organizations that are using these techniques to account. “I think it’s relevant for the future that we educate ourselves and acquire the skills to identify and scrutinize these algorithms,” Uli says. Together with newspaper Der Spiegel, BR Data challenged a major scoring company that calculates people’s creditworthiness. By evaluating the data of more than 2,000 consumers, the team showed that the organization “knows far less about many people than one might think – and yet dares to calculate an exact score. Many people are declared a risk case through no fault of their own. The story caused a big discussion on social media about the methods the scoring company is using, but unfortunately the company has not changed its methods up till now.”



Uli also emphasizes that the knowledge they obtain is being fed back to society. “We consciously share our knowledge by publishing our methods and datasets, whenever possible. We really hope people use it.” In some cases this does indeed happen, as in the investigation which revealed ethnic discrimination in the rental housing market. These methods have been used by various newsrooms in Germany and other countries.



Algorithmic accountability reporting is Uli’s main interest. “Where are algorithms making important decisions and how can these be challenged?” She sees large differences between Germany and the US, where she has recently spent a lot of time: “In the US, algorithms are all around; they are used for criminal sentencing, loans, pay rises. In Germany, we have much stricter privacy and data protection laws, making it less attractive to use algorithms in place of human decisions. Which may be good for society, but not for finding stories.”

Uli hopes more news organizations will set up a data department and focus on the risks and possibilities of AI. But she doesn’t know if this is yet feasible. “I’m not sure. You need to start with the right skills, and for these, resources are needed. Good data scientists are rare and in demand, which means they can earn a lot of money at big companies. That you have to balance with your mission and public value.”

You also need good journalists to cooperate with the data scientists. At BR Data, editors from other departments come to pitch ideas and ask questions, leading to collective projects in which the data scientists focus on the technical challenges and the journalists provide in-depth knowledge. “This collaboration works very well in our organization, as people can focus on their strengths while at the same time learning new skills and acquiring additional knowledge.” These skills and insights are needed to hold the growing number of organizations and institutions that are using machine learning or AI to account.

EXAMPLES OF THE PROJECTS:

<https://web.br.de/interaktiv/erhoehtes-risiko/english/>

<https://www.hanna-und-ismail.de/english/>

<https://web.br.de/madeira/english/>

<http://web.br.de/interaktiv/winnti/english/>



CONSORTIUM, LEAD BY THE CENTRE FOR RESEARCH AND TECHNOLOGY HELLAS (CERTH), GREECE

INVID

Dr. Vasileios Mezaris

InVID Project Coordinator, and Research Director at the Information Technologies Institute (ITI) of CERTH
bmezaris@iti.gr

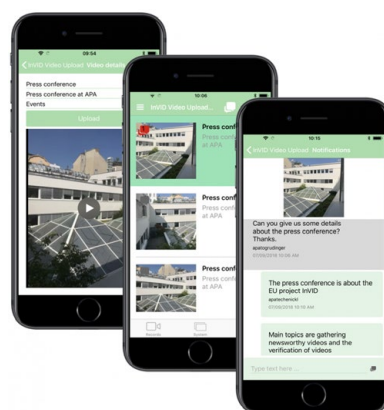
TAKEAWAYS

- As there are many ways to manipulate a video, a tool that tries to combat manipulation also needs to combine various antidotes.
- Flagging possible manipulation is good enough, as journalists themselves want to make the final decision as to whether or not a video has been manipulated.
- Computers are better equipped than humans to process large volumes of data, for example in comparing videos for similarities.

Artificial intelligence plays a crucial role in checking the reliability of social-media videos. By processing huge amounts of information, algorithms are able to detect if videos have been tampered with or are similar to previously posted videos. This information is extremely useful to journalists who need to decide whether or not to trust videos for purposes of their news production.

Bringing together a consortium of nine partners from Austria, Greece, Spain, France and Germany, InVID is a truly European effort designed to battle disinformation and manipulation in social-media video content as used by the news industry. It applies academic research directly to tools that can be used to detect, authenticate and check the reliability of newsworthy video files.

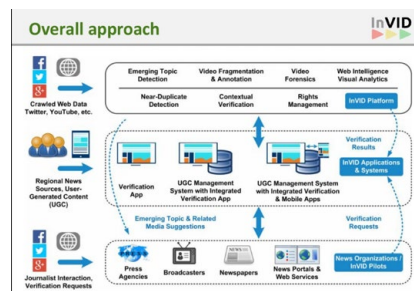
This project's approach is unique as it does not focus on just one type of technology but combines various tactics. Project coordinator Dr Vasileios Mezaris, explains how this works: "There is not one method being used in disinformation, so there can't be just one solution. Sometimes a video has been digitally manipulated; sometimes it's been taken out of its context. You need different techniques to signal these problems."



"THERE IS NOT ONE METHOD BEING USED IN DISINFORMATION, SO THERE CAN'T BE JUST ONE SOLUTION. YOU NEED DIFFERENT TECHNIQUES TO SIGNAL THESE PROBLEMS."
 (VASILEIOS MEZARIS)

Artificial intelligence lies at the heart of all of these techniques. "We could not have done it without the use of AI," says Mezaris. "For example to check if a video has been altered, we need the power of algorithms. There are simply too many elements to check, and the differences can be too subtle for humans to detect. As humans, we can see something. But a computer can see so much more."

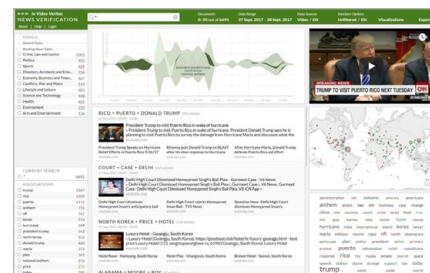
Various new deep-learning algorithms were developed by the team. “One algorithm automatically describes the content of videos. Another can assess similarities between two videos in order to find effective duplicates. And a third can assess if videos have been tampered with.” The algorithms are combined in a browser plugin and app that are free for journalists to use.



The tools don't work fully automatic. “They will only give the user cues on the various functionalities; they do not draw any conclusions. It can, for example, indicate that video A shows a great resemblance with video B, which was uploaded six months ago. But it will not say that it is for sure a duplicate. This is for the journalist to decide.” The journalists using the tools prefer it this way: “Journalists want to make their own informed decisions, and have their own judgement calls. Otherwise, the tool would be like black box.”

The project officially ended in December 2018, but its impact is snowballing. “In December 2018, there were 10,000 downloads of the plugin. Now, in November 2019, we already have 15,000 downloads,” says Mezaris. “The users are mainly journalists, but also amateurs use our tools. And we also see usage in the field of human rights, where user-generated videos are also often used, either as evidence in cases of violations or as disinformation.”

InVID has also developed an app through which audiences can upload their user-generated content. This footage automatically gets the stamp of approval of being verified content. “By using this app, news organizations actively involve their audience in the process of verification, which makes them more engaged. This in turn improves the trust-relationship between the news brands and their audiences.”



Vasileios sees more and more applications for AI in the news. “Evidently in verification, where the power of the machines are simply indispensable. But as AI is also used to create fakes, especially deepfakes, you also need these techniques to combat them. Then there's also robot journalism and the reusing and repurposing of content with the use of artificial intelligence that is very interesting.” For these applications, new research projects have already been designed.

FOR MORE INFORMATION

<https://www.invid-project.eu>

https://www.invid-project.eu/wp-content/uploads/2018/10/InVID_Leaflet_v4.0.pdf



WITNESS.ORG, UNITED STATES

PREPARING FOR DEEPFAKES AGAINST JOURNALISM



Sam Gregory
WITNESS.org
sam@WITNESS.org

TAKEAWAYS

- Journalists, especially women, may be targets of humiliation attacks with deepfake technologies. Organizations should prepare themselves for this.

- There are a lot of universities, start-ups and even public service companies developing technologies for deepfake detection. However, in the technology race, the fakers have an advantage.

- Journalistic organizations need strategies to defend their role as distributors of trusted, verified content, and know how to use new anti-deepfake approaches in their workflow.

- Often it will be about proving something someone claims is a 'deepfake' is real, rather than verifying that content is falsified

Celebrities and politicians are already victims of deepfake technology, which can be used to produce revenge porn and other manipulating and often humiliating fake videos. The next in the line of targets are journalists. US-based nonprofit WITNESS.org has held workshops with media houses, technologists and academics to build a strategic approach for fighting this phenomenon.

It might sound like a joke: To put words in someone's mouth, in a video, and see if someone believes whether he or she actually said it.

But this is no joke. Deepfake videos and audios are threatening journalism in two opposite but equally harmful ways: you trust in something you should not, and you don't when you should.

With deepfake technology, you can paste a person's face onto somebody else's body and put a person somewhere where he or she has never been, or insert them into an event that never happened. For journalism, the ability to use machine learning to produce fake audios and videos represents a major paradigm shift. So far, 'seeing and hearing' has been 'witnessing and believing'. Television and radio – by broadcasting real-life events – have over time become trusted sources of news.

Therefore, deepfakes need a strategic approach to be countered, concludes US-based non-profit organization WITNESS.org. It has pioneered and initiated collaboration with news media, with whom it shares a common interest in being prepared for the time when both find themselves in the 'eye of a storm' with a massive attack – which could happen at any time.

Sam Gregory, program director at WITNESS.org, underlines the need for awareness and preparedness, and has coordinated action among those who share the same concern. In his opinion, deepfakes are in fact both a new wave of reputation-based attacks targeting journalists as well as a new form of manipulating online content. "People who've worked with UGC or content like that have had a decade at least of dealing with the same kind of challenges."



And how does one prepare? Like the BBC, DW, The Washington Post and The New York Times have done: by familiarizing themselves with the deepfake technologies. The BBC, for example, has tested faking their own BBC World News presenter, Matthew Amroliwala, speaking Spanish, Hindi and Mandarin Chinese – languages he does not speak. These results, featuring a familiar newsreader's face, were frighteningly good and reminded the organization itself about the dangers of deepfakes.

“DON'T MOVE TOO FAST TO A 'YOU-CAN'T-BELIEVE-ANYTHING-APPROACH', TO A CONVENTION THAT WILL FUNDAMENTALLY CHALLENGE BOTH TRUST IN JOURNALISM AND TRUST IN COMMUNICATION THAT IS NOT YET JUSTIFIED.” (SAM GREGORY)

Sam calls for a common approach involving sharing training datasets and technologies among researchers, platform companies and journalists. Detection should be built into daily processes in newsrooms. In practice, that means toolkits, browser extensions and access on platforms to systems like reverse search for videos. Add lots of training and stress-testing of newsroom processes.

While there are effective technologies for detecting each type of deepfake, often built on using the same technologies with which the fakes were created, it's already a cat-and-mouse game in which the fakers have an advantage. While journalists have to build credibility and trust, fakers don't have that burden and often also have the technological upper hand.



As detection technologies get better, so do faking technologies. Eye-blinking, for example, was found in summer 2018 to be unnatural in fake videos. Once that finding became public, fake algorithms were shortly thereafter taught to blink naturally.

In the near future, it will become impossible to the naked eye or ear to distinguish real video or audio from synthetic versions.

That's why shared detection and authentication technologies are needed, providing strong signals for human reasoning and good journalistic practice and for sorting through content as it gets easier and easier to make fakes at volume. Journalistic organizations should also prepare themselves for attacks on their teams and journalists who are attacked should be supported and protected.

“Non-consensual sexual images are already a problem – an issue which is often underplayed. This has been used against journalists in a number of cases already,” says Sam Gregory.

Sam underlines the role of audio in fakes. “Audio synthesis has been improving, I think, more rapidly than people expected a year ago. And it's more vulnerable in some senses because it has less semantic clues around it,” says Gregory.

The key message for journalists from him?

“Don't move too fast to a 'you-can't-believe-anything-approach', to a convention that will fundamentally challenge both trust in journalism and trust in communication that is not yet justified.” “It's not that every image is fake.”

“We have to avoid playing into the hands of people who want to call everything 'fake news' and to technology solutions that will completely substitute a technical signal for human judgement, rather than complement human judgement. Yet we do have to prepare.”

LINK TO WEBSITE

<https://blog.WITNESS.org/2019/06/deepfakes-synthetic-media-updated-survey-solutions-malicious-usages/>

<https://lab.witness.org/projects/osint-digital-forensics/>

<https://lab.witness.org/projects/synthetic-media-and-deep-fakes/>

<https://blog.WITNESS.org/2018/07/deepfakes/>

[http://openaccess.thecvf.com/content_CVPRW_2019/papers/Media%20Forensics/](http://openaccess.thecvf.com/content_CVPRW_2019/papers/Media%20Forensics/Agarwal_Protecting_World_Leaders_Against_Deep_Fakes_CVPRW_2019_paper.pdf)

[Agarwal_Protecting_World_Leaders_Against_Deep_Fakes_CVPRW_2019_paper.pdf](http://openaccess.thecvf.com/content_CVPRW_2019/papers/Media%20Forensics/Agarwal_Protecting_World_Leaders_Against_Deep_Fakes_CVPRW_2019_paper.pdf)

<https://arxiv.org/pdf/1806.02877.pdf>

AN INTERESTING EXAMPLE OF BBC FAKING THEIR OWN NEWSCASTER, INCLUDING VIDEO OF HIM SPEAKING DIFFERENT LANGUAGES:

[https://www.bbc.co.uk/blogs/internet/entries/814eee5b-a731-45f9-9dd1-](https://www.bbc.co.uk/blogs/internet/entries/814eee5b-a731-45f9-9dd1-9e7b56fca04f)

[9e7b56fca04f](https://www.bbc.co.uk/blogs/internet/entries/814eee5b-a731-45f9-9dd1-9e7b56fca04f)



DEUTSCHE WELLE, GERMANY
BBC, UNITED KINGDOM

AUTOMATIC TRANSLATIONS



Peggy van der Kreeft
 Head of Research and Cooperation projects, DW
peggy.van-der-kreeft@dw.com

Andrew Secker
 Language Technology Lead for BBC News
andrew.secker@bbc.co.uk

TAKEAWAYS

- Automatic translation is one of the most promising applications of machine learning for PSM. It can be used to dramatically increase efficiency and multiply reach and coverage for organizations operating in several languages.

- Even low-quality translation is better than no translation at all, as it can be used for monitoring the news scene – especially in languages that are rarely spoken outside their indigenous area.

- High-quality standards in PSM mean that, in most cases, household-name broadcasters still need human editors to check and correct the translations. This stance could place them at a disadvantage, since the audience may value even a low-quality translation instead of no content at all.

- If a automatic translation could come close to the PSM standard of writing, it would open up considerable business opportunities among EBU Members for the providers of such a service.

Automatic translation tools are giving international broadcasters like BBC or DW the ability to understand the varied multilingual content that their own organizations are producing. It could be said that these tools have almost magical capabilities. Audiences will soon be able to experience more content, for example listening to or viewing programmes that have been translated into minority languages. However, human editors will still be needed to ensure clarity in public-facing versions.

When you first see and hear it happen, it feels like magic. A news video is downloaded. Shortly thereafter, it is automatically translated, and subtitling in another language appears on the screen. Or, even more amazingly, an artificial voice narrates in the new language.

That may be a bit exaggerated. In reality, a human editor is still needed to make sure that the machine has not made any silly mistakes. However, according to Wilfried Runde, Head of Innovation at Deutsche Welle, a piece can be versioned in another language in a fraction of time compared to doing it all manually.

“FEEDBACK REGARDING THE SPEED OF TRANSLATION HAS BEEN OVERWHELMINGLY POSITIVE BUT THE STORYTELLING ASPECT, AND TONE OF VOICE, I WOULD SAY, IS SOMETHING THAT MACHINES WILL NOT ACHIEVE ANYTIME SOON.”
 WILFRIED RUNDE, HEAD OF INNOVATION PROJECTS, DEUTSCHE WELLE

While initial feedback has been overwhelmingly positive, Runde manages expectations: “We are still very hands-on when it comes to implementation.” Open-source technologies and libraries for automatic translation are growing rapidly all over the world. The benefits of automating complex language-related tasks are obvious for international broadcasters like BBC or DW. Moreover, individual journalists are already using commercially available tools in their daily work. DW currently monitors its in-house content in 8 languages, fully automatically capturing the information, transcribing it and translating it into English. The goal is to expand this range to other source languages. The platform doing this is fully

automated, without human intervention or correction, and is mostly for in-house use, to allow DW staff to track what is produced internally across its 30 languages. In addition, another platform goes further and provides a complete toolkit for computer-aided transcription, translation, subtitling and voice-over, with post-editing opportunity in all phases, from virtually any language into any other. This platform is currently still under development and an initial version is already in beta-testing at DW.

“Basically a lot of opportunities for content adaptation and exchange are lost because people cannot get the full picture on what is produced within all DW languages,” says Peggy van der Kreeft, the project lead who has been involved with DW’s language projects from the beginning.



At BBC News, the approach is more cautious. “We’re experimenting, both around the technology and how it could be placed within people’s workflows to actually help them out,” says Andrew Secker, BBC News Labs’ Language Technology Lead.

A BBC News Labs pilot project, Alto, designed to publish automatically translated pieces on social media using an artificial voice, was trialled but not further developed as the underlying technologies did not provide output of sufficient quality for the BBC. “The technology was not mature enough,” says Secker.

For Secker, the most promising project is the one run jointly with DW as part of the EU-funded projects SUMMA and GoURMET, which aims to increase visibility and coverage for under-resourced languages. BBC World Service operates in 40 languages but automatic translations supplied by business partners would not be available in all of them, or the output is not yet good enough. Broadcasters have special requirements that have to be dealt with to make the system work in practice, and that’s what these projects are focused on. Video and audio have to be fed in and played out. The process should be as smooth as possible for journalists working in hectic environments.

The broadcasters also have to manage costs relative to the benefits achieved. Translation of large video files requires “an awful lot of computing power” and some of the commercial components are not cheap either. According to Secker, it is crucial to approach the issue by asking the question: How can we make journalists’ lives easier rather than pushing new technology into a workflow just because that technology exists.

“We want to solve problems rather than saying, here is a tool.”

When users are journalists hungry for news, small mistakes don’t matter too much. If needed, facts can be checked from the field with own correspondents or freelance journalists.

But when will we have a fully automated process that produces language that can be put in front of audiences?

“Not anytime soon,” says Wilfried Runde.

LINK TO WEBSITES

[BBC News Labs Language projects](#)

[DW Innovation projects](#)

EXPLAINER BOX

SUMMA is an EU-funded project that integrates speech recognition and machine translation with deep language understanding capabilities (including entity-relation extraction and semantic parsing) for open-source applications.

GoURMET, another EU-funded project, standing for Global Under-Resourced Media Translation, aims to develop a sustainable machine-translation platform for under-resourced language pairs and domains as well as accompanying tools for analysts and journalists.

news.bridge, a Google-funded project, is building an environment that will recognize and transcribe speech from different audiovisual sources into various languages, translate transcriptions, produce voiceovers of transcriptions, generate summaries of transcriptions and allow users to edit their transcriptions and translations. It will also allow for the publication of audiovisual content with subtitles and voiceovers.

HOW TO

THE BASICS OF

AI ETHICS

OLD RULES IN NEW CONTEXTS

Are AI ethics different from general ethics?

The answer is simple: no. But AI ethics is a field of applied ethics, so the general rules have to be applied to AI-specific issues. Additionally, AI systems can have unintended consequences, since they are sometimes 'black boxes' – even the people who trained them may not know what logic they use to draw their conclusions.

The main ethical problems in AI are:

- A **lack of transparency**, due to the 'black box' issue
- The **data used to train the model** can be used unethically or contain unintentional biases (for example, a postcode may seem to be just a number – an unbiased piece of data about where people live – but it could be used in an unethical and discriminatory way to predict where crimes will occur)
- The machines can be **designed** in such a way that they produce biased results, e.g. a machine may only be able to detect faces if they belong to white people
- AI can be used for **unethical purposes**, such as unmanned weapons
- Data can be **collected and handled** unethically, e.g. by breaking privacy laws

One key challenge is that people tend not to accept errors from machines, even when they would accept the same errors from humans. Blind spots in the systems also undermine trust and prevent AI from being used to produce positive outcomes.

High-level experts with the European Union have published guidelines for the ethical use of AI¹ that are very much in line with the core public-service mission and its values. The guidelines state that AI should empower humans to make informed decisions, that algorithm usage should be transparent, and that AI systems should take into account the values of diversity, non-discrimination and fairness. Its key points are summarized below.

- **Human agency and oversight:** AI systems should empower human beings, allowing them to make informed decisions and upholding their fundamental rights. At the same time, proper oversight mechanisms must be put in place.
- **Technical robustness and safety:** AI systems need to be resilient and secure as well as safe, accurate, reliable and reproducible to ensure that unintentional harm is minimized and prevented.
- **Privacy and data governance:** Privacy and data protection laws must be fully complied with. Adequate data governance mechanisms must also be put in place.
- **Transparency:** Data, AI systems and AI business models should be transparent. Traceability mechanisms can help achieve this goal. In addition, AI systems and their decisions should be explained in such a way that the stakeholder concerned can understand. People should always be told when they are interacting with an AI system and made aware of the system's capabilities and limitations.

¹ High-Level Expert Group on Artificial Intelligence. (2019). Ethics guidelines for trustworthy AI. Commission.

- **Diversity, non-discrimination and fairness:** Unfair bias must be avoided as it could have multiple negative implications, including the marginalization of vulnerable groups and an increase in prejudice and discrimination. AI systems should foster diversity, be accessible to all regardless of disability, and involve relevant stakeholders throughout their entire life cycle.
- **Societal and environmental well-being:** AI systems should benefit all human beings, including future generations of people. Therefore, they must be sustainable and environmentally friendly; their impact on other living beings, human relationships and society should be taken into account.
- **Accountability:** Mechanisms should be put in place to ensure accountability for AI systems and their outcomes. Auditability, which enables the assessment of algorithms, data and design processes, should play a key role.

FURTHER READING:

[“We need to talk”](#), by Ellen Schuster of Deutsche Welle, in Medium

[“As AI enters newsrooms, journalists have urgent responsibility”](#), by Nausicaa Renner in the Columbia Journalism Review

[“Robot Wars. How artificial intelligence will define the future of news”](#), by James Ball in the Ethical Journalism Network

[Ethics Guidelines for Immersive Journalism](#), by Ana Luisa Sánchez Laws and Tormod Utne

[What are the ethics of using AI for journalism? A panel at Columbia tried to tackle that question](#), by NiemanLab

[Incorporating Ethics into Artificial Intelligence](#), by Amitai Etzioni and Oren Etzion

[Big Data, Big Problems: Emerging Issues in the Ethics of Data Science and Journalism](#), by Joshua Fairfield and Hannah Shtein

[As technology develops, so must journalist's codes of ethics](#), by Paul Chadwick in The Guardian

[Oxford Handbook of Ethics of AI \(forthcoming\)](#)

HOW TO

ALGORITHMIC

JOURNALISM

DEFINITIONS AND EXPLAINERS

From data journalism to algorithmic journalism: there is a full range of journalism categories that use data in some form or another. To distinguish between computer-assisted reporting, data-driven journalism, data journalism, machine learning in data journalism and algorithmic journalism, visit datajournalism.com. Here you will find definitions of all categories and a history on the development of the genre. For an explanation of the difference between small and big data, watch [this talk](#) by Bahareh Heravi, given at [The Conference](#) 2019.

WHAT YOU NEED

Most algorithmic journalism projects are carried out by an interdisciplinary team consisting of multiple people. In this team, different skills need to be represented. But as people can possess more than one skill - coders can for example also be trained journalists - it's hard to say what should be the minimum size of such a data team.

The below skills are all very useful in a data journalism team that is also implementing machine learning and automation for its investigations:

- Journalism – to choose the topic, do offline research, draw conclusions and provide story context, storytelling and narrative construction.
- Statistics and Data Analysis – for the experimental setup, for analysing large data sets, for the statistical models used in machine learning.
- Coding – to scrape data, run the databases, write code and develop interactive visuals (both front-end and back-end developers).
- Design – for the visualisation and interaction design.
- Project Management – for managing the team and the different projects.

Ulrike Köppen, a 2019 Nieman Fellow and head of the data department at Bayerischer Rundfunk, gives these ten tips to interdisciplinary teams working on data stories:

- Define your mission
- Grow your team organically
- Fit a model to your newsroom
- Hone a digital strategy and mindset
- Work on a common language
- Manage your skill sets
- Live in beta mode
- Develop your own tools (based on practices like dailies and weeklies)
- Engage in job development for news nerds
- Experiment A LOT and learn from your failures

In [this clear and compelling article](#), Ulrike also describes different models for how a data team can function within a newsroom. Some newsrooms take a centralized approach (i.e. an independent team with permanent members), while others prefer an embedded approach (where data journalists work primarily inside other teams).

Whatever the approach, however, internal communication is crucial. Projects should be embedded within the rest of the newsroom as much as possible: a story's impact will be much greater if it also features in broadcasting, for example.

For this to work, the newsroom needs a common language to talk about data journalism projects. When everybody understands the basics of what is going on, different departments will start to contribute ideas.

Data projects also might need completely different planning than 'normal' journalistic investigations, as they are often long term and bring challenges that are hard to predict. New tools may need to be built, for example, or datasets may not be available.

EXAMPLES

In the full range from data journalism to algorithmic journalism, there are many different end results possible. Which ones are good, depends on many factors. Assistant Professor in Information & Communication Studies at the University College of Dublin, Bahareh Heravi, wrote about the [success factors of data stories](#) and explains a few elements that she finds crucial.

It's also worthwhile to take a look at the nominees and winners of the [Data Journalism Award](#), hosted by GEN. Some outstanding examples from 2019 include:

- A story by Quartz, Puerto Rico's Centre for Investigative Journalism and Associated Press, on the 487 people killed by Hurricane Maria, won the [investigation of the year](#).
- Reuters' story on the Indonesia plane crash, developed within 36 hours of the event, won [best use of data in a breaking news story](#).
- Reuters won an award for [data visualization of the year](#) with its compelling story on the pollution of the Ganges river in India.
- This year's [public choice award](#) went to the Dutch Broadcast Foundation for its story on the waist measurements of fashion models.

For the algorithmic journalism stories, we compiled the below list with inspiring examples. This is by no means an exhaustive list, but rather a starting point for inspiration:

- Several inspirational stories are from the Data team at Bayerischer Rundfunk:
 - Increased Risk delves into the way scoring companies calculate people's creditworthiness by using algorithms.
 - In Hanna and Ismail, discrimination in the rental housing market is explored.
 - Another story explains how Madeira is actually a tax haven approved by the European Commission.
 - Hackers are the target in this story on Winnti.
- Forensic Architecture also has several interesting cases:
 - In the [Battle of Ilovaisk](#) they used open source methods to identify tanks in video footage. Also read our case on this project. <link>
 - The [Triple Chaser project](#) was aimed to train 'computer vision' classifiers to detect a specific kind of tear gas canisters, used in situations against civilians.
- In those cases machine learning was used for the investigations:
 - In ICIJ's [Implant Files](#) AI was used for pattern recognition.
 - [ProPublica](#) taught a machine what Congress is talking about.

- BuzzFeed trained an algorithm to recognize flight patterns of [Spy Planes](#).
- Ukrainian [Texty](#) determines the scope of amber mining with a machine learning model.
- And here's an overview what kind of stories [Quartz](#) is doing with its AI Studio.
- And some stories on algorithmic accountability:
 - Harvard Business Review wrote an interesting story on how algorithms are also used by hackers and [challenge cybersecurity efforts](#).
 - Reuters reveals that the recruiting machine that Amazon was using, favored [male over female job applicants](#).
 - The Washington Post also writes about gender bias, in their report on how Google's algorithms [shows prestigious job ads to men, but not to women](#).
 - The Guardian writes about [how the recommendation algorithm of YouTube promotes divisive clips and conspiracy videos](#).



CHAPTER 06

IMAGINE YOURSELF IN A BETTER FUTURE

IT'S HUMANS, NOT MACHINES, THAT MAKE POSITIVE CHANGE HAPPEN

In their essence, data and AI represent the next stage of digitalization, so the old rules of digital transformation still apply. This means the industry as a whole shrinks but the first movers collect the paycheck. For PSM news, adapting quickly requires fundamental cultural changes to be made in the newsroom, but legitimate concerns on evolving journalistic identities also need to be addressed. The winning strategy is to combine unique, insightful and creative reporting with the best that tech can offer.

Although the data and AI trends are these days partly hype, it's taking a huge risk not to be part of this next wave. "In the long run, I don't think there's an option of staying away from these technologies," says Erik Brynjolfsson of MIT. Remember Amara's law: "We tend to overestimate the effect of a technology in the short run and underestimate the effect in the long run."

"The companies that use AI, for at least part of the tasks, will succeed. And managers who know how to use AI will replace managers who don't know how to use AI," Brynjolfsson says, adding the following obvious yet dramatic remark: "Some existing companies are going to disappear because they don't use the technologies, new companies will be created, and others will change and transform. That is a process Joseph Schumpeter called creative destruction."

When looking at the European PSM scene, we see companies that are adopting these technologies and ones that are not. But even the forerunners are only in the beginning stages. The transformation is important for all, not just early adopters and companies with large resources. These technologies are becoming important for all strata of society, and it is vital for the future of the industry and society at large that PSM have the right strategies. Here are some proposals.

CREATE A LONG-TERM STRATEGY AND BE PREPARED TO INVEST

Digitalization typically makes the production of services more efficient and cheaper while shrinking the industry in question. That's what we've been experiencing in news media in recent decades, especially in privately funded media. Global consultancy McKinsey makes it clear: "Early digital movers grow while the whole industry shrinks, so there's a clear benefit of being among the first to adapt."

However, while innovation with data and AI has accelerated, productivity growth has slowed in almost all industrialized countries over the past 15 years. The technology is very real, but additional innovations and investments are required in support of the technology, and this takes time. Concurrently, old resources are tied to old ways of working. It's hard to build the future while running today's newsroom. And it may take several years of investment before the benefits become visible.

“THE PREDICTION MIGHT BE INCORRECT, OR MIGHT BE JUST, NOT RIGHT. BUT THEN THE INTERPRETATION OF THE EDITORS MIGHT ALSO BE INCORRECT. SO HOW DO WE MAKE A DISTINCTION BETWEEN THE TWO AND BUILD A TRUST RELATIONSHIP BETWEEN THE METRIC AND THE JOURNALISTS?”
(MAGDA PIATKOWSKA, BBC)

Take for example the automotive industry and self-driving cars. The investments in the latter are estimated at tens of billions of dollars, but so far no actual drivers have been substituted. Moreover, one of the companies investing most in the self-driving future is not a legacy automotive company but Google. One lesson that PSM could take away is that they need to consider what part of their core mission they want to handle themselves and what part should be left to the tech giants.

A couple of years ago, it was believed that computers would save money and make human work superfluous in newsrooms. In most cases, this has proven untrue. AI is expensive and depends on highly skilled work performed by highly paid specialists. For it to be economically feasible, use cases must be scalable (and the data has to allow for that).

In other words, either AI is used to handle a mass of simple tasks or to produce something extraordinary, or it's a waste of time and resources.

Philip Napoli emphasizes that data journalism and AI should not be considered a means to save money. Quite the opposite in fact: it's an investment. He adds: “Doing this type of journalism requires investing in a way that nobody in the US wants to invest anymore. And so it has become the province of the select few still well-resourced news organizations.”

Let us be gently reminded with this warning: most AI projects fail to meet expectations, first because expectations may be inflated and second because they are not strategic enough and lack support from top management. If there is no strategy, all you get in practice is small pilot projects that fail to make it to production, or projects aimed solely at saving money (i.e. doing old things a bit more cheaply). Nowhere is value created.

David Caswell of the BBC says that there is a huge gap between expectations and actual tangible products and projects, so far. “There doesn't seem to be much to choose from, which is surprising. What happens is an almost immediate focus on recommendation engines. And that is unfortunate because there just isn't much left to do there that hasn't been done.” David Caswell's comment may be true from a global media perspective where recommendations have been the norm for some time, but not in European PSM. Many of the PSM organizations have not yet started servicing the audience's individual needs in a responsible ‘public service’ way.

So the need for urgent change is clear. But how can we make it happen? How can we truly implement change in our organizations? We believe that the true success factor for navigating the next wave of the digital transformation depends especially on the ability to develop human skills. Therefore we stress the following issues:

- Supporting cultural change in the newsroom
- Cultivating the best possible creativity in human-made, AI-augmented journalism
- Managing the paradoxes and tensions of the new era
- Creating the scale needed, especially in data

SUPPORT CULTURAL CHANGE BY CONNECTING TO THE HIGHER PURPOSE

The history of journalism has always been interconnected with technology. Yet journalism is fundamentally a human endeavour. Whenever technologies change, journalism's people-centric role has to be redefined. This naturally applies to public service journalism as well. The roles, tasks and workflows in the newsroom might change, but the mission of telling stories about people does not.

Do not confuse this with keeping everything the way it was. It is perhaps merely romantic to believe that humans should always make the choices of news offered to audiences, or do the editing. Reflect, and find the real value of human choices and storytelling, a 'human-centric perspective' on AI for journalism. Nicholas Diakopoulos claims that AI is "a new medium through which journalists can express and exercise their ethical and normative values."

But retaining this focus does not preclude change. It is perhaps idealistic to believe that humans should always decide what news will be offered to audiences, or do all the editing. Reflect, and find the real value of human choices and storytelling, a 'human-centric perspective' on AI for journalism. Nicholas Diakopoulos claims that AI is "a new medium through which journalists can express and exercise their ethical and normative values."

To be among the frontrunners, innovation alone is not enough. Most change initiatives collapse because the organization's culture is not able to adapt. Lucy Küng, Senior Research Fellow at the Reuters Institute, University of Oxford, has been studying newsroom cultures and concludes that well-functioning collaboration between tech and journalism should build on commonalities between journalistic and engineering cultures. For journalists, this means that "data should be at their service; it should help them do what they really in their inner hearts want to do."

Professor Lucy Küng, Senior Research Fellow, Reuters Institute, University of Oxford notes that change initiatives inevitably run up against organisational obstacles - rigidities arising from hard factors like roles that are defined too tightly or missing skills, and also soft ones like fixed cultural beliefs. Blending the cultures of technology and journalism can be problematic - outwardly these are very different, the one based in the liberal arts, the other in the natural sciences, but in fact there are important areas of commonality, for example a deep commitment to the product itself, a respect for craft skills, and high levels of intrinsic motivation.

For journalists, this means that "data should be at their service; it should help them do what they really in the inner hearts want to do."

Data dashboards are complex to produce and that production is done by tech experts. The challenge is ensuring they are clear enough, but also subtle enough to inform journalistic decision making and really succeed in bringing news creators closer to the needs of news users. This is no easy task and can require years of fine tuning and feedback loops. Data needs to inform decision making, not direct it.

Journalism's roots are in the liberal arts, and some journalists may even be afraid of data, finding it intimidating and confusing. "Many of these journalists had chosen a career that they mistakenly thought would allow them to avoid mathematics," says James Robinson of the Tow Center. And it does not help that machine learning is sometimes - like actually some IT systems, as well - a 'black box' that only gives output, with no understanding of the processes that lie within. If journalists are to feel more comfortable with the algorithms, they have to understand the reason (the 'why') for the processes and not just use the outcome (the 'what').

Professor Mark Deuze says that journalists generally see technology as something that comes in from the outside. "They want nothing to do with it. They don't see that technology is and has always been an essential part of how they do their work."

And there are new challenges ahead, for example as data becomes more predictive. "How do you culturally convince journalists to believe in predictions?" asks Magda Piatkowska, Chief of Data Science at the BBC. "The prediction might be incorrect, or might be just...not right. But then the interpretation of the editors might also be incorrect. So how do we make a distinction between the two and build a trust relationship between the metric and the journalists? (...) I think some of the conclusions have to be aligned with what people think. If the outcome of the machine is going to be always contradictory to what people think or thought about the audience, the metrics, the success of the stories, that is going to be very hard."

Lucy Küng notes that one of the core tasks of journalists is to make decisions based on expert judgments. “Do we just feed the public desire for news, or do we actually lead by infusing our news with the public-service values of informing democracy and holding power to account? I do think that those cultural issues around potential conflicts and data need addressing carefully and I don’t think you’ll get real cultural acceptance unless you seem to be taking those concerns seriously.”

TRAIN MANAGEMENT, STAFF AND AUDIENCES

Who are needed in this machine-intelligence environment, journalists or data engineers? These are quite different people with different skill sets. Bahareh Heravi, who is Assistant Professor in Information & Communication Studies at the University College Dublin, says that newsrooms usually have just a handful of data-literate people: “Most newsrooms still either rely on one or two people to create work for them, or work with people outside of the newsroom. If you look at all the newsrooms around the world, there’s only a very small percentage that have teams bigger than five data journalists.”

So, it seems important to hire the right people with the right skills. But this is not always possible, as “good data scientists are rare and wanted, which means they can earn a lot of money at big companies,” says Uli Köppen. However, the research published by MIT indicates that despite these challenges, the most successful and game-changing projects are run by devoted in-house teams.

Journalists themselves then need to be educated. Many of them are actually interested in that. One of the most striking imbalances is that while more than 40% of journalists would like to learn more, only 5% of newsrooms offer that kind of training (Owen, 2019).

But this is not the case for all. “Journalists are usually not very open to data and machine learning,” says Mark Deuze: “From what I know about schools for journalism and my education, I say that maybe one student in the five years has ever done something with math once. So it’s full of anthropologists and sociologists and political scientists.”

And it should not be forgotten that journalists have been asked to develop their skills constantly in recent years: “Journalists are expected to reskill, deskill, and upskill their practices and working routines, generally without any direct say in the way the organisations they engage with operate. In doing so, they vulnerably move inside and outside of newsrooms and news organizations large and small, trying to both make a difference and to make ends meet in an exceedingly competitive market,” explain Professors Mark Deuze and Tamara Witschge.

Bahareh Heravi thinks that it’s important though that journalists acquire at least some basic data skills. “I don’t think journalists should become computer programmers. But I do hope that all journalists will become more data aware, data literate, so they can actually go and look at the data instead of relying on something that they heard from some minister or something.”

It’s this technological literacy that Charlie Beckett of LSE emphasizes: “The key thing that we saw in our survey was how important it is that journalists have the technological literacy, to make editorial judgment calls, that they’re able to understand the process.”

And the training should not be aimed only to the editors: “There is definitely a huge appetite for a more strategic approach. And that appetite includes a very strong sense that there needs to be more training. Not just on skills for the tech itself, but also on increasing knowledge in key management areas. (...) The knowledge should spread throughout the news organizations so that people can augment journalism in it and improve it.”

Some newsrooms have custom-made training.

EBU Academy is partnering with EBU news to develop data journalism capacities within EBU Members. Currently, EBU Members are facing different situations in terms of availability of public data and the level of skills in their newsroom. EBU is pursuing its efforts to offer adequate training to each level:

- For the most advanced Members: a computer assisted journalism course on Python programming, web scraping and data visualisation. This course is offered in partnership with Cardiff University.
- For the middle level Members: data journalism courses focusing on Excel and Google spreadsheets offered by EBU Academy regional Learning Hubs.
- For the least advanced: A combination of management awareness sessions and training of trainers on entry level data journalism tools.

In our checklist Courses on AI and data, we have collected examples of training offerings available externally, from universities or online. Though typically not about journalism specifically, they still might be useful.

One way to become more literate is visiting events where data skills are exchanged, like on the annual festival Dataharvest in Belgium. Here journalists and 'data nerds' can acquire new skills on various levels by working together on real-life projects.

THE BEST OF BOTH ORGANIZATION CULTURES: BOTTOM-UP AND TOP-DOWN

What we see in both media and elsewhere is the emergence of organizations that are more autonomous. Lean and agile are concepts that have been around a while but they are spreading from tech departments to whole organizations.

However, Lari Hämäläinen, McKinsey's Lead of Advanced Analytics in Nordics and Benelux, says that agile models are not enough to transform an organization. Transformations always have to be led from the top. Otherwise, the desired autonomy will be killed by hierarchies. "Observers don't lead. You need to lead through vision" is his first piece of advice. Second: "Think big, transform top-down". And only the third piece of advice is about agility and autonomy: "Break the silos and implement bottom-up."

When the level of uncertainty rises in a profession, even more focus is needed on correct 'change management'. Professor George Kohlrieser has taught leaders at IMD about his 'secure base' concept. He claims that people do not naturally resist change; they resist the pain of change and the fear of the unknown. He says that people are hardwired to be curious, explore, learn and change. By providing security, leaders can free this natural curiosity. Another trick is to let people make the changes themselves because they usually resist being changed by others.

All autonomous organization models have one thing that prevents them from becoming anarchies: clear objectives and accountability for results. Autonomous decision-making is crucial at the operative level, where the content has to respond quickly to changing habits and needs. And finally, the most successful projects are not led by technologists like Chief Technology Officers or Chief Data Officers. They are led by the top management, which in news organizations may mean News Directors or Editor-in-Chiefs.

Responsibility for successful innovation should belong to everyone in a news organization. By focusing on digital audio as a strategic goal, Swedish Radio has distributed to the editors the ownership of creating even better stories and products.

Autonomy can be practiced systematically by allowing individuals to take on greater responsibility, increasing the scope for an organization to deliver the expected results, as with holacracy as practiced in the Swiss national broadcaster's archive department.

COLLABORATE INTERNALLY, WITH PARTNERS, OR EVEN WITH COMPETITORS

CEO Hanna Stjärne announced in September that five units where employees work with data will be put together" at Swedish Television SVT. Why? "We think that to be more data-driven is a really big question for the future for all media companies. We put them into one so that they could work more closely and they could find connections, and to see how we could use data more than we do today, but responsibly," she says. The model already works fine, and has been extended.

What's happening in Sweden is a move in the right direction. Combining data experts for common benefit is a typical move in a data strategy. Data becomes more valuable when there's more of it and when it provides different angles. Therefore, news organizations should team up around data. The more we know about users, the more holistic an understanding we will have about their needs.

Breaking down silos is just as necessary in other departments. For example, in investigative productions, this means bringing in AI specialists while letting the journalists be journalists. PSM organizations cannot do this alone. They need partners to catch up with developments in this field. Partners can be found in different places: in the international PSM community, within other news organizations, or in public agencies, research institutes and universities that share the same objectives.

In several countries, media companies have understood that they can't compete alone with tech giants because they don't have enough data. Many alliances have emerged among privately owned media, but in some countries public service organizations are also involved.

Quite often, though, R&D doesn't make it to the production stage. One reason may be that sometimes the real value of the projects isn't analysed before the experimentation phase. Sometimes external funding may even have a negative effect. "A lot of the AI innovation we're seeing in Europe is stimulated through the Google Digital News initiative," says Nicholas Diakopoulos. And he sees a risk in this, as this external funding gives project an impetus that may not be driven by solving real challenges that connect to business needs. "If someone walks in and says, we'll give you EUR 500,000 to think about AI and news production, you'll say, sure! And at the end of the year you have a nice prototype. But maybe you wouldn't have done that prototype if it hadn't been for that money. And are you going to invest another EUR 500,000 to bring this into a real production? Or, more likely, EUR 5 million?"

And, of course, there are issues too with media being so dependent on one tech multinational for funding. Another risk is the potential conflicts of interest associated with such partnerships.

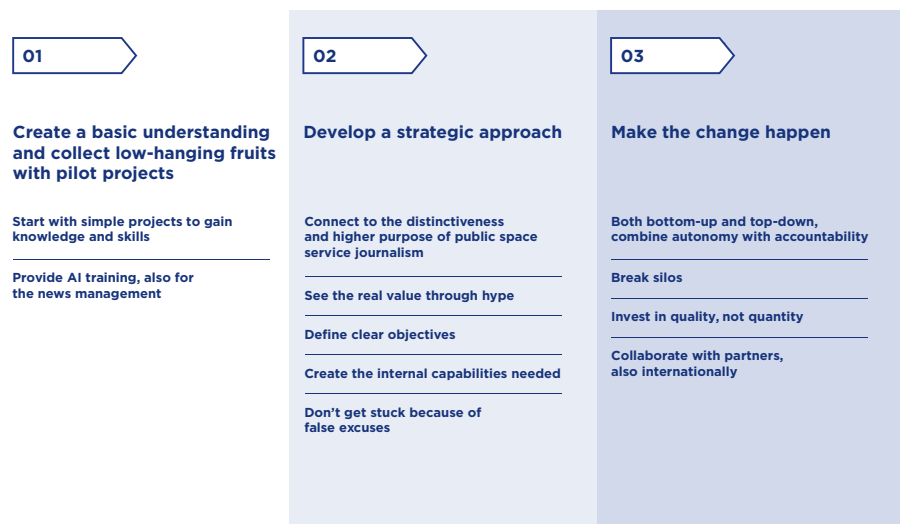
In Switzerland, the PSM broadcaster has teamed up with private media and universities to found an initiative to foster groundbreaking innovation and research.

The Irish public service company RTÉ collaborated with Al Jazeera in a project organized by IBC, the tech and media trade show of Amsterdam.

ROADMAP

Based on Andrew Ng's transformation playbook 5-step model and Tom de Weghe's application for newsrooms, insights from leading consultancies and the additional research we have conducted, we propose the following roadmap in order to successfully implement AI in your organisation:

1. Create a basic understanding of AI and collect low-hanging fruit through pilot projects
 - Start with simple projects to gain knowledge and skills
 - Provide AI training, including to newsroom management
2. Develop a strategic approach to AI
 - Connect to the distinctiveness, values and higher purpose of public service journalism
 - Distinguish the real value from the hype
 - Define clear objectives
 - Create the essential internal capabilities needed
 - Don't get stuck because of excuses and false pretexts
3. Make the major strategic change happen
 - Lead both bottom-up and top-down, combine autonomy with accountability
 - Break down silos
 - Invest in quality, not quantity
 - Collaborate with partners, including internationally



RECOMMENDATIONS FROM THIS CHAPTER

- New AI technologies are still largely hype, but in the long run, they will transform the way journalism is practised. The next transformation will again reduce the number of organizations and journalists involved, but those leading the change wisely and who are willing to meet the challenge head on are the ones who will come out on top. Being among them requires adopting a strategic approach and investing in what is really valuable.
- To combine technology and journalistic identity successfully, news leaders should address the legitimate concerns and fears that journalists have. Changing journalistic culture to adapt to the new technologies will be hard, but connecting to the higher purpose and mission will help explain why the change is needed in the first place.
- There's a greater need for more autonomous organizational models, which are better at adapting to the fast-paced changes affecting audiences and broadcasting environment. However, lean or agile, or even more radical models, are not enough. The change also has to be led from the top. Otherwise, hierarchies will ultimately bury the initiatives.



ARENA FOR JOURNALISM IN EUROPE, BELGIUM

DATAHARVEST, A FESTIVAL FOR DATA JOURNALISTS



Brigitte Alfter
Director of Arena for Journalism in Europe
brigitte@journalismarena.eu

TAKEAWAYS

- Mid-career, on-the-job training that focuses on real projects and is executed in collaboration with (international) colleagues is valuable for people working in data journalism.

- Collaboration seems to be a keyword in many aspects of the data-informed newsroom and is also useful for acquiring skills. Learning from peers is a powerful experience.

- Data providers are often willing to share their data sets and insights with journalists.

To acquire the needed skill set to handle data, you can take a BA or MA degree. But you can also attend a yearly festival like Dataharvest, where you will acquire the necessary skills in a real-life setting. Dataharvest is not your average conference. Yes, there are keynote sessions and presentations, yet it is rather a get-together of like-minded souls, working on open datasets and learning 'on the job'. The festival leaves attendees not just with fresh skills but also with new contacts and concrete ideas for stories.

With a name that reflects the origin of the festival – a cross-border project scrutinizing the EU's generous farm subsidies – Dataharvest attracts annually close to 500 journalists and data developers from all over Europe. Every year they meet in a venue in Belgium, in a traditional conference setting as well as in a so-called 'nonference': rooms where 'data nerds' and journalists meet either to exchange skills or work collaboratively on data sets.

Brigitte Alfter, festival director since 2011, explains why the event has become so successful: "All the top-level people come to Dataharvest to talk about trends. But at the same time, both 'data illiterates' and 'data geeks' come for mutual knowledge sharing and to work collaboratively on stories."

"IT'S LIKE A DATA BUFFET, WHERE YOU CAN GET BOTH THE BASICS AND THE REALLY NERDY STUFF."
(BRIGITTE ALFTER)

The festival is not interested in speakers who are only sharing their own success stories. It's the methodology that matters. "We want to know how you did the investigating so others can learn and apply those lessons to other projects." The methods designed and tools developed are almost always open for everyone to use unless there are privacy issues or embargoes. In line with the



philosophy of autonomous teams, a part of the festival can be shaped by the visitors themselves. “We have rooms that can be used for all kinds of meetings, and people fully take responsibility and design parts of the programme themselves. Sometimes behind closed doors, as some investigations can be confidential.”

Festivals like these are extremely important to the data community. Not so much for the talks that are held on stage, but for the much-needed exchange of skills. “The data people are often rather isolated within their organizations, as they are unique and there often aren’t many of them. This festival gives them an opportunity to teach each other. The networking element is key for this success.”



There are not too many classes and courses on data and/or AI journalism around, and there are considerable differences between countries. It’s also difficult to prepare for data journalism work when one is studying journalism. “Journalism students are already swamped as it is. Also, you don’t know yet which skills you will need: is it data scraping? Is it to be good in Google Search? To run open-source or social-media investigations as Bellingcat does?” So it’s the ‘on the job’ training that usually makes the most sense.

The festival began with a practical focus: collaboration and skill sharing around a particular dataset each year, like the EU farm subsidies. “Afterwards it became more conference-like, with keynote speakers and plenary events.” But it still offers classes on data skills, divided into three levels. “It’s like a data buffet, where you can get both the basics and the really nerdy stuff.”

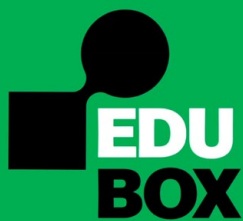
The focus is very much European. The conference brings together both stories and people from across the continent. “Visitors leave with knowledge, datasets and new colleagues throughout Europe.” Important guests are the providers of data. And quite often, these come from the EU. “For example, we invite the Director-General for Budget of the EU so he can give insight into the tool they have openly available to access their data. We find that data providers are quite willing to cooperate, as it’s in the common interest to have the data transparent and accessible.”

Brigitte believes data skills are important to everyone in the newsroom. “Basic skills are needed for all. You need to know what is possible, how the internet works, how a database is structured. If you have some basic understanding, you will be less scared of data and instead see the opportunities. And you need to know how journalists and data experts communicate to get the best out of such cooperation.”

She also points to the considerable opportunities that lie ahead in interactive journalism, in which citizens are involved to contribute to datasets that journalists can then check, analyse and process. “There is great potential in gathering data together with the audience. Also in light of the whole debate on trust, there might be a key here, as people are made part of the truth-seeking.”

FOR MORE INFORMATION

<https://dataharvest.eu>



VRT, BELGIUM EDUBOX ON AI

Tom van de Weghe
Research Fellow Stanford University & Foreign
Correspondent, VRT NWS
Tom.vandeweghe@stanford.edu

TAKEAWAYS

- For people of all ages, but especially young people, it's important that they learn how AI works and what the possible implications of AI are. A public service broadcaster can offer educational materials for this purpose.
- For children, addressing AI in an entertaining manner, in an offline setting, makes it more concrete and understandable.
- Journalists might be another target group for whom training material should be developed, as they often find the subject matter too complicated and in any case irrelevant.

The news department at Flemish public broadcaster VRT (VRT NWS) is explaining the complicated subject matter of algorithms, predictions and data by presenting secondary school pupils with an apocalyptic game world in which they must distinguish between humans and zombies. With the EDUbox on AI, the students learn – mainly in an offline setting – how AI works and what the possible implications are. This box has also value for adults, as we are all sometimes confronted with zombies, eh, AI.

Learning about AI is not only important for journalists; all people in society are faced with algorithms and their consequences every day. VRT NWS considers it their responsibility to offer education on this topic and so has developed a so-called 'EDUbox' on AI. "Few people are actually aware of how artificial intelligence works. Furthermore, many people that believe that they understand how artificial intelligence functions base their 'understanding' on unfounded urban myths," explains VRT on its website.

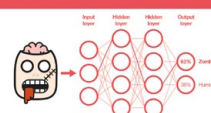
"WE CONSCIOUSLY DESIGNED MOST ELEMENTS OF THE BOX TO BE EXECUTED OFFLINE. THIS MAKES AI MUCH LESS VAGUE AND DISTANT, SHOWING INSTEAD THAT IT'S SOMETHING TANGIBLE."
(TOM VAN DE WEGHE)

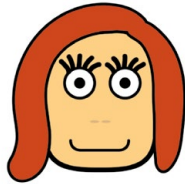
Tom van de Weghe, a foreign correspondent at VRT NWS, is one of the developers of the EDUbox. He spent the past year as a Fellow at Stanford University exploring the use of deep learning and blockchain technology in detecting deepfake video and authenticating audiovisual content. "It's important that children learn how algorithms work, so they better understand the impact they have on our lives and can prepare themselves for a future in which AI will play an even bigger role."

The EDUbox is an actual cardboard box that secondary school teachers can order from VRT. The material covers the basics of AI in a two-hour course. It contains videos, reading materials and group exercises designed for a classroom setting. "We consciously designed most

3.5. NEURAL NETWORKS, WHAT ARE THEY?

Neural networks are relatively recent. But that is no reason not to understand them. Via the video below we briefly explain how a neural network arrives at its final decisions.



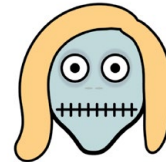


elements of the box to be executed offline, so that students would actually sit down together and work on it in real life. This makes AI much less vague and distant, showing instead that it's something tangible."

The material is designed to be entertaining: "The pupils are confronted with a zombie apocalypse and need to use various AI systems to distinguish real people from the zombies."

The EDUbox is aimed specifically at young people, but there are currently discussions on how it can be applied to cater to older audiences as well. "I'm not sure if the example of the zombie apocalypse would remain in a kit aimed at grown-ups, but the general idea of understanding the way AI works is also very much of interest for an older audience."

Internally, there is also interest in the topic. "At VRT NWS, we are currently focusing on how to improve the tech literacy of the journalists." Tom is part of the Data Unit at VRT NWS, where journalists and developers from various departments work together on stories focusing on data, disinformation and AI. "Perhaps it might be a good idea to give the EDUbox on AI to our colleagues as a Christmas gift. Because for many journalists, AI is still very much a distant matter, not something that they actually understand."



Here Tom sees the importance: "AI can have a positive impact on society but can also be misused. It's our responsibility to monitor and report on this." In that respect, he sees opportunities for collaboration on a European level: "We should cooperate, as public service broadcasters. An algorithm doesn't stop at a border, so we shouldn't either."

ARTIFICIAL INTELLIGENCE IS OMNIPRESENT

Believe it or not, we have all already come into contact with AI, including you!

It is used, for example, in:

- ☐ various apps, such as Waze or Google Maps
- ☐ social media, such as Facebook and Instagram
- ☐ search engines, such as Google, Yahoo and Bing
- ☐ streaming services, such as Netflix and Spotify
- ☐ mechatronic systems (robots, self-driving cars)
- ☐ ... and so much more

? AI is therefore used far more than you might think.
Do you know any other applications that use AI?

The EDUbox on AI is the fourth themed EDUbox produced by VRT NWS. The previous three focused on data, democracy and fake news. "They are all interesting, but I feel this one on AI is spot on what people are talking about."

FOR MORE INFORMATION

<https://www.vrt.be/vrtnws/en/2019/10/23/vrt-news-launches-artificial-intelligence-edubox-in-english/>

<https://www.vrt.be/vrtnws/nl/2019/09/17/artificiele-intelligentie/>



SR, SWEDEN

WE ALL INNOVATE

Josef El Mahdi
Head of Editorial Development
josef.el_mahdi@sr.se

TAKEAWAYS

- Enacting a huge strategic shift within an organization is not a one-person job – it takes people from all departments and positions to join forces.

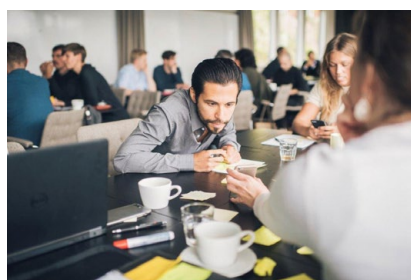
- Innovation can happen much more easily when more people take ownership of their work. Professionals often know quite well what is needed, and by helping them in their efforts rather than taking the innovation away from them, much more can be achieved.

- Sharing innovative methods rather than finished products enables departments to execute new projects on a stand-alone basis.

Having a strategy is important for any innovation. But being able to execute that strategy is imperative. To do so, it's not just good leaders that are needed: everybody in the organization needs to understand their role in contributing to the end goal. By sharing the ownership of their innovation towards digital audio, SR inspired editors in all of its newsrooms to create even better stories and products, thus realizing the company's key objective.

If public-service news organizations want to become better at adopting new technologies and possibly even become frontrunners in digital, it's not enough to just appoint leading technical staff and a few good managers. They need to change the very structures in which technological innovation can take place. Innovation needs to be everyone's responsibility.

Josef el Mahdi, Head of Editorial Development: "At SR, we have a national newsroom, 25 local news stations and several language-specific newsrooms, plus Science, Sports and Culture. To get all of them to move in the same direction is not an easy job – but combined these newsrooms constitute an extraordinary force for intelligent journalism." Yet moving the company as one is what they have been doing



"MANAGEMENT KICKSTARTED INNOVATION AND SET THE RIGHT TARGETS, BUT IT WAS UP TO THE REST OF US TO FIGURE OUT THE EXACT 'WHAT' AND 'HOW'." (JOSEF EL MAHDI)

since 2017, when SR decided on a new strategic course: "We wanted to change the way we delivered news online, as this was too focused on text. And this needed to be changed to audio. After all, we are Swedish Radio." By implementing this change of course, SR hoped to attract more people to listen to SR's audio content on digital platforms.

The new strategy implied introducing a completely new way in which news stories are told. "On FM, audio clips begin and end with the presenter. But for digital, we needed to develop clips that could be played stand-alone. Because only then can you



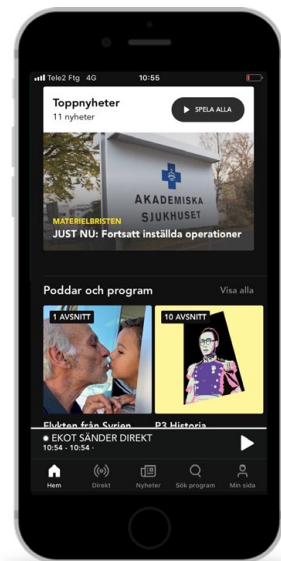
mix and match the offering, creating unique and more personalized content for social and smart media apps,” explains Josef. The shift in focus was huge for the organization: “1,000 people needed to change the way they told stories and needed to embrace innovation as part of their jobs.”

Instead of implementing this radical change only top-down, SR consciously placed the ownership of the innovation with the people who were actually creating the content: the editors. “Management

kickstarted innovation and set the right targets, but it was up to the rest of us to figure out exactly ‘what’ and ‘how’. The innovation then happened within the newsrooms, in collaboration with our digital specialists, where people are in close contact with their audiences.” One result is “Topnyheter”, Top News, a news playlist that in a short time has gone from zero to 100,000 + clips played weekly.

Josef and his colleagues in Editorial Development have played an important role in the transformation, as they are helping and coaching the news departments. “We literally travel up and down Sweden to talk about the new mission, with emphasis on the ‘why’ and to get as many on board as possible. We offer guidance and help, for as long as it’s needed.”

The interesting thing is that they are not just giving advice; the news departments also inspire them. “Actually, most of the major insights on how to actually do this digital way of audio storytelling came from the local newsrooms. They lead the way on how to best make digital clips.” In line with the new strategy, newsrooms and digital specialists have continued teaming up. “We focus on making it easier for the audience to choose SR’s audio: push notifications to sound, social tools for audio platforms, new visual presentations that drive listening rather than reading, ditching the page view logic for more qualitative data and a podcast curation desk.”





RTS, SWITZERLAND HOLACRACY, IN SWISS ARCHIVES



Léonard Bouchet
Head of Data and Archives
leonard.bouchet@rts.ch

TAKEAWAYS

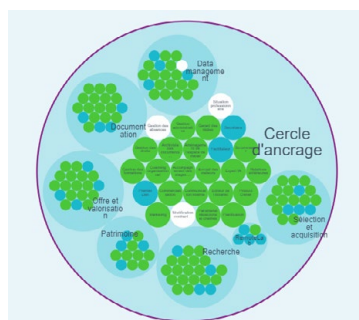
- One of the key challenges of AI projects is to move from research and experiments to actual operations. Most projects are cancelled before the production phase. The reason is that the challenges in rolling out the systems for production are often underestimated while expectations might have been overestimated.

- In highly demanding AI projects, it's important that the key people are given room to solve the problems autonomously. Their intrinsic motivation to solve problems is high. Micro-management can sap motivation.

- Self-adapting organizational models enhance the innovation process.

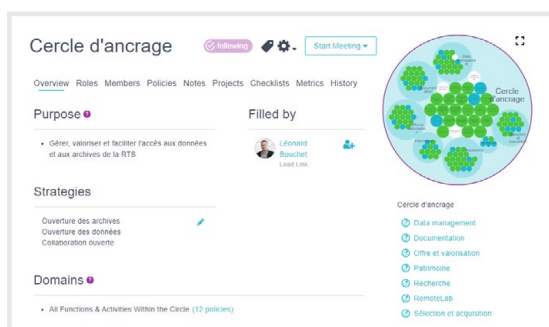
The French-speaking arm of the Swiss public broadcasting RTS has quickly emerged as one of the forerunners in adopting AI in media. What made it possible for RTS to move from experiments and R&D to real-life applications, while others are often still stuck in pilot phases? This success is a combination of two things: the right people in the right place at the right time, but also innovative organization models, like holacracy, explains Léonard Bouchet, one of the architects behind the success story.

Leonard Bouchet, Head of the Data and Archives department of RTS, says that the real difficulty in AI projects is not AI, but access to relevant data and integrations to the production systems. In other words: To make the system work in its real environment. "Our experience is that it's 40% data, 40% integrations and 20% AI." Despite these challenges, RTS has rolled out several functioning deep learning systems in less than two years. That's more systems in production than there are projects in the planning phase in most public service companies in Europe. Many of these deep learning systems are world premieres in use in everyday production.



Among them is facial recognition of more than three thousand Swiss politicians and other locally known public personalities, something that even global tech giants are not able to provide for the moment. RTS is also able to recognize a speaker's identity from audio. Gender is identified as well, as are different kind of scenes, objects and buildings. All this means that RTS is able to provide visual search in its archives, and if decided, an almost real-time search of interesting footage and audio

"THE FIRST REACTION WAS THAT IT'S COMPLETELY CRAZY. THERE'S NO WAY WE CAN IMPLEMENT IT HERE. THIS IS SUCH A HIERARCHICAL ORGANIZATION"
(LÉONARD BOUCHET)



recorded from the field, inducing dramatic changes in news workflows and allowing fast detection of interesting footage and publishing. To overcome the challenges of bringing AI in production, RTS has rolled out new features in small pieces and changed the way changes were typically introduced in the legacy systems. And focused always on people, both in building the models and adapting to the new workflows on the editorial side. They even have adapted a self-

organising, non-hierarchical model called holacracy.

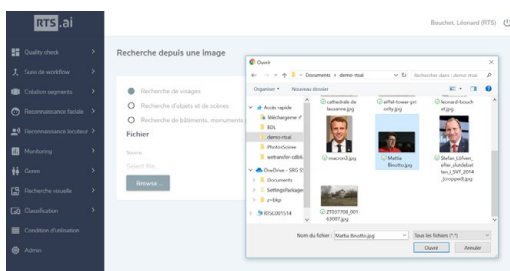
Léonard Bouchet has led the RTS Data and Archives department of 60 people, in Geneva and Lausanne, since January 2016. A few months after starting, he noticed and measured, with the help of the RTS organizational coach, that something had to change in the organisation. It lacked clarity, the people were not really satisfied and the structure was clearly not flexible enough for the upcoming digital and AI tsunami.

The coach introduced the concept of holacracy to Bouchet. “The first reaction was that it’s completely crazy. There’s no way we can implement it here. This is such a hierarchical and industrial organisation”, he says. But a deeper look at the idea revealed that it would address exactly the challenges Bouchet had: to shift to a dynamic organisation with clarity and using the collective intelligence of the teams. Self-adapting organisational models, like holacracy, are not anarchies, explains Bouchet. Autonomy is balanced with clearly defined roles and processes. Each of them is accountable for delivering certain aims.

In other words, authorities and duties are fully delegated and distributed in different roles in holacracy. Defining strategies and priorities, in turn, are a strong way to lead the organisation. Within the roles, there’s full autonomy. Micro-management is a total no-no in this model. . “There is no hierarchy of people. But you really have the power of leading because the roles have the obligation to align themselves with the strategies and the priorities.” Change is driven by tensions. If there is a problem, something to improve, or an opportunity, there is a reason to change. And substantial changes can be decided and enacted in a few minutes, rather than months, as it often happens in more traditional systems.

Of course, the new organisational model is not the only explanation for the successful roll-outs in RTS. When asked what were the key factors for success, Léonard explains the importance to have people organized in a way that actually helps them find the best solutions themselves. But regardless of the model, he says, the most important factor in his opinion is to “not be married to any model” and have a constant reflection, staying open to any possible change to get better results. And that’s the role of the organizational coach.

For Léonard Bouchet, the AI revolution has only just begun. “We are at a very early stage in this technology tsunami which grows exponentially by nature, and we can’t imagine all the work that can be avoided.” “The development will be strong and fast. Faster than we expect.”



Léonard Bouchet calls for international co-operation between public broadcasters. “The only way public broadcasters can survive is to work together. There’s no way to gain a critical mass that’s needed to compete with Netflixes and Amazons of this world otherwise. It’s a hard way, but I think it’s the only one.”

FOR MORE INFORMATION

<https://www.rts.ch/archives>

<https://www.facebook.com/lesarchivesdeRTS/>

<https://www.youtube.com/user/rtsarchives>

<https://www.instagram.com/rtsarchives/>



INITIATIVE FOR MEDIA INNOVATION, SWITZERLAND JOINING FORCES FOR MEDIA INNOVATION



Mounir Krichane
Director
mounir.krichane@epfl.ch

TAKEAWAYS

- Public and private media organisations are able to collaborate in the field of innovation if the decisions are made at a strategic level within organizations. Successful collaboration requires a clear governance structure, clear commonly agreed objectives and principles, and clear decisions on funding.

- Universities that are used to applying their research to various industries, in conjunction with specific companies, can add value to these current projects aimed at helping media organizations.

- Collaboration can be interdisciplinary, combining expertise in the fields of technology and social sciences, for example.

- The IMI is a young organization. It remains to be seen what kind of disruptive innovation it can foster through this kind of financing policy.

In Switzerland, public media SRG-SSR has joined forces with two private media houses and five universities to jointly fund research for break-through innovations in media. With the funding provided, six projects have already started, four of them applying artificial intelligence. Each project has one media partner and one university partner, but results are shared with all alliance members.

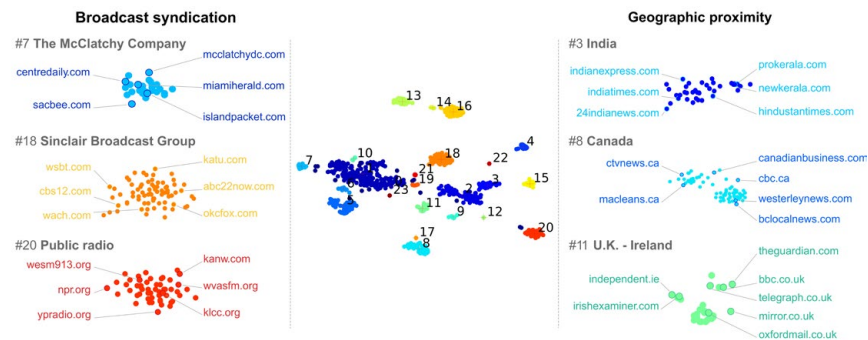
“Ten or five years ago this would not have happened. People were not ready to collaborate this way. This has now changed,” says Mounir Krichane, Director of the Initiative for Media Innovation (IMI), Switzerland. The initiative is a consortium between, on the one hand, public and privately-owned media companies in Switzerland and, on the other, some of the most prominent research universities in Europe. The aim is to combine scientific research with professional practice to create breakthrough innovations that have real value for media organizations.

IMI was founded in January 2018. The first year was spent building the initiative and launching the first call for proposals, on misinformation and public trust. There's money available too. The founders are financing CHF 700,000 (EUR 635,000) worth of research project costs per year, with some additional funding from investors and even public sources like Ofcom, the Swiss communications watchdog. The media companies also provide expertise and a first-hand understanding of the media environment. Two people work full time at the initiative's Media Centre in Lausanne.



“WE WERE VERY LUCKY TO HAVE THE RIGHT PEOPLE WHO WERE SMART ENOUGH AND OPEN-MINDED ENOUGH TO COME TOGETHER, SIT AT THE TABLE, AND START TO COLLABORATE.”
(MOUNIR KRICHANE)

“The big challenge is that you are bringing in people from very different backgrounds to work on an initiative they are co-financing and from which they all want to gain something. It’s one of the challenges but also one of the success factors,” says Krichane.



A similarity map, where distances between channels represent the similarity of their coverage.

“We propose a model of the media landscape that automatically maps news sources based on their selection of subjects. By tracking sources over time, we can identify driving forces– from the influence of ownership to large-scale content diffusion patterns. We have produced a dynamic map of the media landscape over the past 3 years, based on coverage from 500 million articles shared by around 8,000 sources.”

“We were very lucky to have the right people who were smart enough and open-minded enough to come together, sit at the table, and start to collaborate.”

Mounir Krichane, a former applied researcher who built cutting-edge distributed wireless weather sensors at EPFL and who also has experience with RTS and its Digital Lab, underlines the applied nature of the projects IMI is funding.

“Any project funded has to come up with something that’s beneficial to media.” All partners share the results of the research, even if only one of them acts as the media partner.

Founding members include SRG-SSR, a private media group Ringier, technical university EPFL of Lausanne and the universities of Geneva, Lausanne and Neuchâtel, through their common platform Triangle Azur. Others have since joined. EPFL competes with the likes of Cambridge, Oxford and ETH Zurich for the top European spots in university rankings in the field of engineering and technology.

According to Krichane, both private and public media now understand that this kind of common initiative is needed. Digital transformation has resulted in both societal and business pressures that threaten the future of established organizations. “People are more inclined to collaborate because of that. Everybody has understood that they can’t do it on their own.”

The initiative’s first call for proposals was on the subjects of fake news and public trust. Close to 30 came in; 10 candidates were then asked to submit a full application. Six projects have now been launched, and four of them use sophisticated machine-learning models to deliver practical solutions for partner media organizations.

SRG is media partner for two of the published projects, both of which use machine learning and data science. The first uses deep-learning techniques to reduce the risk of picking false or irrelevant content in automatic aggregations. In the second, the aim is to understand social-media activities for which the data is not transparently available. The researchers are analysing networks with machine learning to discover hidden patterns.

The funding decisions are for a single year, so the initiative is only at the very beginning of reaching its goals. To achieve really substantial results, longer projects with additional funding may be needed. The second call for proposals has already been published, specifically looking for projects to strengthen the relationship between media organizations and their audiences. Projects involving new approaches for content recommendations and predicting audience needs would be especially welcome.

FOR MORE INFORMATION

<https://www.media-initiative.ch/>



**RTÉ, IRELAND
AL JAZEERA, QATAR**

**ASSOCIATED PRESS, USA
IBC, UK**

COLLABORATING IN R&D WITH OTHER COMPANIES

TAKEAWAYS

- Mittmedia had clear goals based on deep contextual understanding: that it is important to know the audience and their news consumption in order to select relevant content.

- Mittmedia's data strategy and data architecture made it possible to use different kinds of data in complex algorithms.

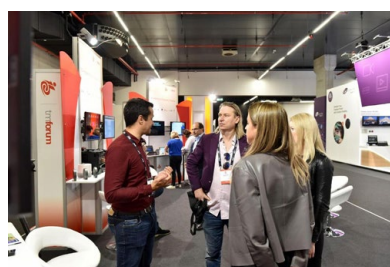
- Journalistic values were carefully taken into consideration.

Deirdre Temple,
Head of Solutions, RTÉ Technology
deirdre.temple@rte.ie

Mark Smith,
Technology, Telecom and Media Advisor, IBC
msmith@ibc.org

Grant Franklin Totten
Head of Cloud & Analytics, Al Jazeera
totteng@aljazeera.net

The Irish broadcaster RTÉ needed to automatise monitoring of regulatory requirements relating to its election programming. It does not have an R&D unit or resources to build complex AI systems on its own. As a result, it responded to a 'catalyst' challenge, an initiative organized for the first time for media by IBC, the media and telecom show which takes place in Amsterdam. In doing so, RTÉ joined forces with Al Jazeera, which also had a monitoring issue, to specify a challenge that would be compelling enough to encourage technology providers to develop a solution and a proof-of-concept to solve the problem.



The idea emerged when Irish broadcaster RTÉ was reporting to the regulator about political representation in its output, to ensure balance in national election programming. Records about the airtime allotted to the various politicians had been kept manually. "Initially, we used stopwatches to do that. We've then moved on to multi-timers where we have people signing in with iPads and capturing how long someone is speaking," explains Deirdre Temple, Head of Solutions at RTÉ.

Using AI was possibly a way to eliminate the manual labour-intensive process and move from a system that has built-in delays when collating data to live reporting – and have the ability to address in real time if needed.

One problem, however: RTÉ did not have an R&D department or an R&D budget to develop the AI solutions needed for these kinds of cases. And even if it did have the budget, for a single small broadcaster facing financial constraints, the investment would be very difficult to justify.

“VENDORS AND BROADCASTERS WORKING COLLECTIVELY ENSURED WE MOVED THIS PROJECT AT A PACE THAT WOULD NOT BE DELIVERED IF WE TRIED TO DO THIS ON OWN. THERE IS NO WAY WE WOULD HAVE GOT THIS FAR IN BASICALLY SIX TO EIGHT MONTHS. IT’S JUST NOT POSSIBLE. IT WOULDN’T HAVE HAPPENED”

(DEIRDRE TEMPLE)

At the same time, in Qatar, Al Jazeera was trying to solve a similar issue. It had to report internationally, for example in the UK to Ofcom, about specific aspects of its coverage. “We have a whole team of people within the news organization that looks at editorial quality issues,” says Grant Franklin from Al Jazeera. These quality issues can be visual, auditory or textual. “We started to say, look, instead of having 500 human beings sitting here, looking for all these problems and logging them, why don’t we see which ones we can automate. So we went from an analytics approach to an AI automation approach.”

What brought RTÉ and Al Jazeera together was the international tech and media show IBC, with its new approach to accelerating innovation.

“IBC came to the conclusion that the speed at which digital transformation is happening around the industry requires new models and approaches to innovation within the industry,” says Mark Smith, Advisor at IBC.



For broadcasters, the idea is almost irresistible. Instead of starting their own R&D project, with all the costs, risks, time and resources, the Catalyst project offers the opportunity to present collective problems as challenges to technology providers. Broadcasters don’t have to pay at all to get to the proof-of-concept level. Instead, technology providers help fund the project to receive valuable guidance to help them solve the case, serve the customers and possibly open a new market for themselves. They also receive a highly visible slot at the major IBC conference, attended by some 55,000 participants.

To serve the media industry with this concept, IBC joined forces with TM Forum, a telecom industry’s association, for the first time. The first three projects were presented in May at Digital Transformation World -conference and at the IBC show in Amsterdam in September 2019.

The solution for RTÉ and Al Jazeera was created by technology companies V-Nova, Metaliq, Tech Mahindra and the Qatar Computing Research Institute, who together introduced an innovative solution with a contextual algorithm that enabled the monitoring for targeted issues across multiple content categories. The list is extensive: as well as political representation, there are questions like fact-checking, issues of taste such as use of graphic content or protection of privacy, and questions of accuracy in terms of validity of maps and names.



The solution contextualized the metadata by combining object, action, face recognition and text detections: plainly speaking, to detect not only what is in the footage but what’s going on, contextually. For example not only a group of people but a protest; not only a sign but a protest sign and the context what the demonstration is about. Grant Franklin says the requirement to operate almost live has been met. “Let’s say 10 times the speed at which we’re able to detect it now using existing algorithms on the market.”

The telecom industry's TM Forum has been running catalyst programmes for around 20 years. Hundreds of projects have been delivered through this initiative.

"We estimate that the speed at which you can do R&D is about five times faster than typically done in a telco or a broadcast organization. So it's very, very fast innovation, open innovation," says Mark Smith.

Although the broadcaster partners have yet to proceed from proof-of-concept phase to actual live systems, Grant Franklin testifies: "It would have taken us probably two years to do what we did within six months."

Deirdre Temple is not short of praise for the concept. "It's just such a great idea. I mean, even if your ideas come to nothing, just the relationships that you're building with your peers within other organizations or the learning that you get is pretty phenomenal. We don't do it enough as an industry."

"Even if we had all the money in the world, there's no way we would have got this far in basically six to eight months. It's just not possible. It wouldn't have happened," she concludes.

FOR MORE INFORMATION

[IBC Website & Telecoms-Media Catalysts Projects](#)

MOBILE NEWSGATHERING USING AI-POWERED COMPRESSION

[RTE \(Ireland\), Associated Press \(USA\), BBC NEWS \(UK\), BT \(UK\), V-Nova \(UK\),
Awiwest \(France\)](#)

[This Catalyst project aims to improve the efficiencies of bandwidth utilization using
advanced compression and encoding to empower journalists in the field.](#)

HOW TO

COURSES ON AI AND DATA

How can a newsroom design a useful public service media (PSM) algorithm?

TRAINING IS NEEDED

Journalists working in data-informed newsrooms need to have the skills to either communicate with data experts or carry out certain data journalism and/or machine learning tasks themselves. If journalists don't yet have these data skills, then it's crucial that they receive training, either in-house or externally.

EXAMPLES

Some newsrooms offer their own custom-made data journalism classes; the [BBC Academy](#) has been doing this since 2016. The EBU Academy also regularly offers training on various aspects of data journalism, like this workshop on its [Regional Learning Hub](#). And when this report went to press, the EBU Academy is working with the EBU News Unit in creating a specific data training program for PSM newsrooms.

There are also training opportunities outside the news organizations. Bahareh Heravi, an assistant professor of Information & Communication Studies at the University College of Dublin, has [researched](#) data journalism courses worldwide, from short courses on 'small data' to full-time master's courses on complicated machine learning. Here are the highlights:

- The University College of Dublin's short [postgraduate professional certificate degree](#) with only three modules taught in the evenings, designed for working journalists;
- The Ray Juan Carlos University of Madrid's well-known [master's programme](#) in investigative journalism, data and visualization;
- The University of Amsterdam's [bachelor-level course](#) in New Media & Digital Culture.

All the courses from the study are plotted on this [map](#).

Renowned expert Nicholas Diakopoulos has set up a [free online course](#) on the impact of automation and AI on journalism which debuted in February 2019.

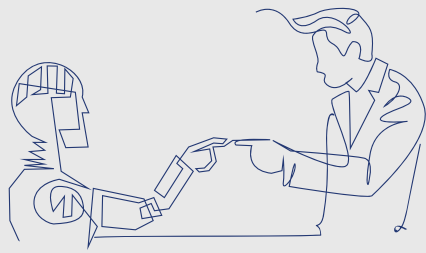
Below is a list of other online courses on AI, which are not specifically about journalism AI but still useful:

Basics of AI and machine learning

- 'AI for Everyone', by Andrew Ng at Coursera
- <https://www.coursera.org/learn/ai-for-everyone> or DeepLearning.AI
<https://www.deeplearning.ai/ai-for-everyone/>
- 'Elements of AI', a free online course by University of Helsinki and Reaktor
<https://www.elementsofai.com/>

More advanced courses

- 'Intro to Machine learning', by Udacity: <https://www.udacity.com/course/intro-to-machine-learning--ud120>
- 'Introduction to AI', by EdX: <https://www.edx.org/course/artificial-intelligence-ai-columbia-csmm-101x-0>
- 'Machine Learning', by Andrew Ng at Coursera: <https://www.coursera.org/learn/machine-learning>
- 'Natural Language Processing with Deep Learning', by Stanford University: <http://web.stanford.edu/class/cs224n/>
- 'Convolutional Neural Networks for Visual Recognition', by Stanford University: <http://cs231n.stanford.edu/>
- 'Statistical Machine Learning', by Carnegie Mellon: <http://www.stat.cmu.edu/~larry/=sml/>
- 'Artificial Intelligence', by MIT: <https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-034-artificial-intelligence-fall-2010/>
- 'Neural Networks for Machine Learning with Geoff Hinton', by Coursera: <https://www.coursera.org/learn/neural-networks>
- 'Practical Deep Learning', by Fast.ai: <https://course.fast.ai>



CHAPTER 07

PUBLIC SERVICE: UNITE TO DEFEND YOUR FUTURE!

One of the striking claims that we encountered during our investigative journey looking into the future forms of journalism was that we might – right now – be entering communication's final paradigm. It started thousands of years ago as one-to-one. Then in the modern age came one-to-many. Now the trend is towards many-to-many.

Is there anything for newsrooms to progress to, or is this the end of the road? A striking thought.

That's why the challenge faced by public service media is so fundamental. Its future mission will depend on the kind of lives that people want to live and how shared 'public life' will fit into that. In turn, this will define the future mission of PSM.

This upcoming shift may well spell the end of PSM as we know it. But it could also be the start of a new era in which PSM organizations are more than ever called upon to act as curators of shared experiences. In a decentralized networked society, trust and ethical data use will be the most precious currencies.

We can't avoid the obvious conclusion that PSM organizations in Europe will not be able to travel this road successfully on their own.

Their own data sets are not big enough. R&D resources are not big enough. They can sit back and wait for the free solutions. But the real difference will be made by striking out with strategic initiatives, which of course require substantial investment. Those PSM organizations that wish to stand out will therefore need to team up with others.

At a societal level, there has to be a new consensus about the shape and form of PSM in the new era.

So our message to European PSM is 'Unite! Work together to make it happen!'

And finally, while data and AI are certainly transforming the way in which journalistic output is produced and distributed, there is value that can't be extracted from data.

If you have a precise end goal in mind measuring, optimizing and automating is a perfectly good way forward. But journalism at its best entails creating something that has never been done before. The world changes, people change and journalists change too. Exploring the unknown and making sense of it is the highest aim.

The world changes, yes, but mostly for the better. It's our duty to make sure it does.

ABOUT AUTHORSHIP



ATTE JÄÄSKELÄINEN

Is a Professor of Practice at LUT University, Finland, and a Senior Visiting Fellow at LSE, London. He is also a Senior Advisor at Fourkind; an advisory company on artificial intelligence and digital business models.

He is the former Director of News and Current Affairs at the Finnish Broadcasting Company, Yle, having previously served as CEO and Editor in Chief for the Finnish News Agency, and as Political Editor and Business Editor for Finland's leading daily paper. He recently served as a Visiting Fellow for the Reuters Institute in Oxford.



MAIKE OLIJ

Is a journalism consultant and media innovator at [./bureaumaike](https://bureaumaike.nl). With a background in Communication Science, Maike has worked in journalism for 12 years, primarily at the Dutch public service broadcasters NPO and NOS. digital business models.

She has done extensive research into people's motivations for consuming news, which culminated in her 2016 book, *Why we need news?* Maike now develops new journalistic formats and gives strategic media advice to media organizations, journalism foundations and (local) governments.

THE TEAM

(in alphabetical order)



JEFF DUBIN

is Head of News and Events at the EBU, overseeing the EBU's news output. Prior to joining the EBU, he was an Emmy award-winning producer at CBS News, based in New York and London.



LUIS JIMENEZ

is the EBU News External Relations Manager, Digital News Lead, and project manager for this report. Before joining the EBU he worked for 8 years at RTVE News in Spain where he headed the News and Sports Exchange department.



JUSTYNA KURCZABINSKA

leads the Eurovision News Exchange and News Strategy at the EBU. Since 2017, she has also overseen the Public Service Journalism Initiative aimed at strengthening and supporting independent trusted news that includes the Report.



ED MULHALL

is an editorial advisor and consultant with 40 years' experience in broadcast and news media. He was previously Managing Director of RTE News, Ireland, as well as a television and radio producer, and editor. He now advises the EBU, leading their EVN Rules project.

INTERVIEWS FOR THIS REPORT

Besides the interviewees for the cases, the following people, among others, have contributed with their insights and opinions (in alphabetical order):

- Benoit Balon-Perin - Digital Workflows Manager at RTBF
- Charlie Beckett - Media and Communications Professor at the London School of Economics
- David Bunker - Head of audiences at BBC
- Erik Brynjolfsson - Professor at MIT
- David Caswell - Executive Product Manager at BBC News Labs
- Mark Deuze - Professor Media studies at University of Amsterdam
- Nicholas Diakopoulos - Assistant Professor in Communication Studies and Computer Science, Director of the Computational Journalism Lab (CJL) at Northwestern University, Tow Fellow at Columbia University School of Journalism
- Sam Dubberley - Special Advisor, Evidence Lab at Amnesty International
- Bahareh Heravi - Assistant Professor in Information & Communication Studies at the University College of Dublin, Co-chair of the European Data & Computational Journalism Conference, Director of UCD Data Journalism CPD
- Lucy K  ng - Senior Research Fellow at Reuters Institute University of Oxford
- Jaakko Lempinen, Head of Customer Experience & Data & AI at Yle
- Philip Napoli - Professor of Public Policy at Duke University
- Magda Piatkowska - Head of Data Science at BBC
- Mika Rahkonen, Head of Strategy at Yle
- Hanna Stj  rne - Director General at SVT
- Roberto Su  rez Candel - Head of Strategy & Media Intelligence at EBU
- Tom van de Weghe - John S. Knight Fellow at Stanford University and Foreign Correspondent at VRT NWS

REFERENCES

- Agrawal, A., Gans, J., & Goldfarb, A. (2016, Nov 17). The Simple Economics of Machine Intelligence. Retrieved from HBR.org website: <https://hbr.org/2016/11/the-simple-economics-of-machine-intelligence>
- Agrawal, A., Gans, J., & Goldfarb, A. (2018). Prediction Machines: The Simple Economics of Artificial Intelligence. Harvard Business Press.
- Arnecke, P. (2018, March). Dive into AI and machine learning. iTech, pp. 6–7.
- BBC News. (2019, September 7). BBC and tech firms join forces to fight fake news. BB Retrieved from <https://www.bbc.com/news/technology-49615771>
- Beckett, C. (2019, November). New powers, new responsibilities: A global survey of journalism and artificial intelligence. Retrieved from <http://blogs.lse.ac.uk/polis>
- Bodó, B., Helberger, N., Eskens, S., & Möller, J. (2019). Interested in Diversity. Digital Journalism, 7(2), 206–229.
- Bremme, K. (2019). AI in the Media Sector. A mapping of Artificial Intelligence applications in the media sector. Retrieved from France Televisions Directorate of Innovation and Prospective website: <https://www.meta-media.fr/files/2019/06/Cartographie-IA-dans-les-medias-ENG.pdf>
- Broussard, M. (2018). Artificial Unintelligence: How Computers Misunderstand the World. MIT Press.
- Broussard, M., Diakopoulos, N., Guzman, A. L., Abebe, R., Dupagne, M., & Chuan, C.-H. (2019). Artificial Intelligence and Journalism. Journalism & Mass Communication Quarterly, 96(3), 673–695.
- Burri, M. (2013). Contemplating a “Public Service Navigator”: In Search of New (and Better) Functioning Public Service Media. <https://doi.org/10.2139/ssrn.2364951>
- Burt, C. (2019, July 1). Researchers develop “soft biometrics” from facial expressions to detect deepfakes. Retrieved October 14, 2019, from Biometric Update website: <https://www.biometricupdate.com/201907/researchers-develop-soft-biometrics-from-facial-expressions-to-detect-deepfakes>
- Chadha, K., & Wells, R. (2016). Journalistic Responses to Technological Innovation in Newsrooms. Digital Journalism, 4(8), 1020–1035.
- Coddington, M. (2015). Clarifying Journalism’s Quantitative Turn. Digital Journalism, 3(3), 331–348.
- Deepnews.ai FAQ. (n.d.). Retrieved October 25, 2019, from Deepnews.ai website: <https://www.deepnews.ai/#faq>
- Deuze, M., & Witschge, T. (2018). Beyond journalism: Theorizing the transformation of journalism. Journalism, 19(2), 165–181.
- Diakopoulos, N. (2014). Algorithmic Accountability Reporting: On the Investigation of Black Boxes. <https://doi.org/10.7916/D8ZK5TW2>
- Diakopoulos, N. (2019). Automating the News: How Algorithms Are Rewriting the Media. Harvard University Press.
- European Broadcasting Union EBU. (2019, September 10). Public Service Media Contribution to Society Project. Retrieved October 1, 2019, from European Broadcasting Union website: <http://www.ebu.ch/psm-contribution-society>
- European Broadcast Union. (2012). Public Service Values. Editorial Principles and Guidelines. Retrieved from European Broadcast Union website: https://www.ebu.ch/files/live/sites/ebu/files/Publications/EBU-Public_Service_Values.pdf
- Ferrer-Conill, R., & Tandoc, E. C., Jr. (2018). The audience-oriented editor: Making sense of the audience in the newsroom. Digital Journalism, 6(4), 436–453.
- Fields, B., Jones, R., & Cowlishaw, T. (2018). The Case for Public Service Recommender Algorithms. Retrieved from <https://piret.gitlab.io/fatrec2018/program/fatrec2018-fields.pdf>
- Fraunhofer-Gesellschaft. (2019, February 1). Software that can automatically detect fake news. Retrieved October 16, 2019, from Phys.org website: <https://phys.org/news/2019-02-software-automatically-fake-news.html>
- Graefe, A. (2017, September 1). How algorithms and human journalists will need to work together. The Conversation. Retrieved from:

<http://theconversation.com/how-algorithms-and-human-journalists-will-need-to-work-together-81869>

- Graefe, A., Haim, M., Haarmann, B., & Brosius, H.-B. (2018). Readers' perception of computer-generated news: Credibility, expertise, and readability. *Journalism*, 19(5), 595–610.
- Graves, L. (2018). Understanding the Promise and Limits of Automated Fact-Checking. Retrieved from Reuters Institute for the Study of Journalism website: <https://agency.reuters.com/content/dam/openweb/documents/pdf/news-agency/report/reuters-institute-graves-factsheet-180228.pdf>
- Grothaus, M. (2019, October 15). Machine learning isn't effective at identifying fake news. Retrieved October 16, 2019, from Fast Company website: <https://www.fastcompany.com/90417625/machine-learning-isnt-effective-at-identifying-fake-news>
- Hämäläinen, L. (2019, September 18). Change in the Era of AI. The What, The How and The Unknowns. A presentation at AI & Data Intelligence -seminar, Helsinki.
- Hansen, M., Roca-Sales, M., Keegan, J. M., & King, G. (2017). Artificial Intelligence: Practice and Implications for Journalism. <https://doi.org/10.7916/D8X92PRD>
- Hanusch, F., & Tandoc, E. C., Jr. (2019). Comments, analytics, and social media: The impact of audience feedback on journalists' market orientation. *Journalism*, 20(6), 695–713.
- Hao, K. (2019, June 28). An AI app that “undressed” women shows how deepfakes harm the most vulnerable. MIT Technology Review. Retrieved from <https://www.technologyreview.com/s/613898/an-ai-app-that-undressed-women-shows-how-deepfakes-harm-the-most-vulnerable/>
- Helberger, N. (2015). Public Service Media. Merely Facilitating or Actively Stimulating Diverse Media Choices? Public Service Media at the Crossroad. *International Journal of Communication Systems*, 9(0), 17.
- Helberger, N., & Burri, M. (2015). Public Service Media | Introduction ~ Public Service Media and Exposure Diversity. *International Journal of Communication Systems*, 9(0), 5.
- High-Level Expert Group on Artificial Intelligence. (2019). Ethics guidelines for trustworthy AI - Digital Single Market - European Commission. Retrieved from European Commission website: <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>
- Jääskeläinen, A., & Olij, M. (2018). News Report 2018: 50 ways to make it better. Retrieved from European Broadcasting Union website: <https://www.ebu.ch/publications/news-report-2018-50-ways-to-make-it-better>
- Karlsen, J., & Stavelin, E. (2014). Computational Journalism in Norwegian Newsrooms. *Journalism Practice*, 8(1), 34–48.
- Kohlrieser, G. (n.d.). Unleash potential as a Secure Base Leader - George Kohlrieser. Retrieved October 14, 2019, from George Kohlrieser website: <http://www.georgekohlrieser.com/unleash-potential-as-a-secure-base-leader/>
- Köppen, U. (2019, July 9). Working Across Disciplines: A Manifesto for Happy Newsrooms | Nieman Reports. Retrieved October 1, 2019, from Nieman Reports website: <https://niemanreports.org/articles/working-across-disciplines-a-manifesto-for-happy-newsrooms/>
- Lamot, K., & Paulussen, S. (2019). Six Uses of Analytics: Digital Editors' Perceptions of Audience Analytics in the Newsroom. *Journalism Practice*, 1–16.
- Laskowski, N. (2018, September 26). AI models could help companies overcome human bias. Retrieved October 16, 2019, from SearchCIO website: <https://searchcio.techtarget.com/news/252449410/AI-models-could-help-companies-overcome-human-bias>
- Lewis, S. C., Guzman, A. L., & Schmidt, T. R. (2019). Automation, Journalism, and Human-Machine Communication: Rethinking Roles and Relationships of Humans and Machines in News. *Digital Journalism*, 7(4), 409–427.
- Lindén, C.-G., & Tuulonen, H. (2019). News Automation. The rewards, risks and realities of “machine journalism.” WAN-IFRA
- Linden, T. C.-G., & Dierickx, L. (2019). Robot Journalism: The damage done by a metaphor. *Unmediated: Journal of Politics and Communication*, 2, 152–155.
- Lomas, N. (2019, September 13). This game uses troll tactics to teach critical thinking. TechCrunch. Retrieved from <http://social.techcrunch.com/2019/09/13/this-game-uses-troll-tactics-to-teach-critical-thinking/>
- Marconi, F., & Siegmund, A. (2017). The Future of Augmented Journalism. A guide for newsrooms in the age of smart machines. Retrieved from AP website: <https://insights.ap.org/industry-trends/report-how-artificial-intelligence-will-impact-journalism>

- McBride, K. (2019, May 17). The "Liar's Dividend" is dangerous for journalists. Here's how to fight it. - Poynter. Retrieved September 17, 2019, from Poynter website: <https://www.poynter.org/ethics-trust/2019/the-liars-dividend-is-dangerous-for-journalists-heres-how-to-fight-it/>
- Mele, N., Skibinski, M., & Spector, M. (n.d.). How Today's News Publishers Can Use Data, Best Practices, and Test-and-Learn Tactics To Build Better Pay-Meters - Lenfest Institute for Journalism. Retrieved October 12, 2019, from Lenfest Institute website: <https://www.lenfestinstitute.org/reports/digital-pay-meter-playbook/>
- Milosavljević, M., & Vobič, I. (2019a). Human Still in the Loop. *Digital Journalism*, 1-19.
- Milosavljević, M., & Vobič, I. (2019b). "Our task is to demystify fears": Analysing newsroom management of automation in journalism. *Journalism*. <https://doi.org/10.1177/1464884919861598>
- Missaoui, S., Gutierrez-Lopez, M., MacFarlane, A., Makri, S., Porlezza, C., & Cooper, G. (2019). How to Blend Journalistic Expertise with Artificial Intelligence. Retrieved from City University London website: https://blogs.city.ac.uk/dminr/files/2019/02/Missaoui_etal-2aaaaou.pdf
- Mullin, B. (2016, April 27). Bloomberg EIC: Automation is "crucial to the future of journalism" - Poynter. Retrieved October 10, 2019, from Poynter website: <https://www.poynter.org/tech-tools/2016/bloomberg-eic-automation-is-crucial-to-the-future-of-journalism/>
- Neheli, N. B. (2019). Metrics and Analytics in the Newsroom: An Ethnographic Study Exploring how Audience Data are Changing Journalistic Practice. Bournemouth University.
- Nelson, J. L., & Tandoc, E. C., Jr. (2019). Doing "well" or doing "good": What audience analytics reveal about journalism's competing goals. *Journalism Studies*, 20(13), 1960-1976.
- Nerone, J. (2015). Journalism's Crisis of Hegemony. *Javnost - The Public*, 22(4), 313-327.
- Newman, N., Fletcher, R., Kalogeropoulos, A., & Nielsen, R. K. (2019, June 12). Reuters Institute Digital News Report 2019. Retrieved August 1, 2019, from Digital News Report website: <http://www.digitalnewsreport.org/survey/2019/overview-key-findings-2019/>
- Ng, A. (2018, December 13). AI Transformation Playbook: How to lead your company into the AI era. Retrieved October 13, 2019, from Landing AI website: <https://landing.ai/ai-transformation-playbook/>
- OECD. (2019). The Future of Work. OECD Employment Outlook 2019. Highlights. <https://doi.org/10.1787/9ee00155-en>
- Owen, D. (2019). The state of technology in global newsrooms. Retrieved from ICFJ website: <https://www.icfj.org/sites/default/files/2019-10/2019%20Final%20Report.pdf>
- Pallanich, J. (2019, May 16). Newsrooms Getting Smarter About AI - TV News Check. Retrieved October 10, 2019, from TV News Check website: <https://tvnewscheck.com/article/235067/newsrooms-getting-smarter-about-ai/>
- Pavlik, J. (2000). The Impact of Technology on Journalism. *Journalism Studies*, 1(2), 229-237.
- Peiser, J. (2019, February 5). The Rise of the Robot Reporter. *The New York Times*. Retrieved from <https://www.nytimes.com/2019/02/05/business/media/artificial-intelligence-journalism-robots.html>
- Petre, C. (2018). Engineering consent: How the design and marketing of newsroom analytics tools rationalize journalists' labor. *Digital Journalism*, 6(4), 509-527.
- Ransbotham, S., Khodabandeh, S., Fehling, R., LaFountain, B., & Kiron, D. (2019, October 15). Winning With AI. Retrieved October 21, 2019, from MIT Sloan Management Review website: <https://sloanreview.mit.edu/projects/winning-with-ai/>
- Robinson, J. G. (2019). The Audience in the Mind's Eye: How Journalists Imagine Their Readers. Retrieved from Tow Center for Digital Journalism, Columbia University website: <https://academiccommons.columbia.edu/doi/10.7916/d8-drvj-wj06>
- Sallinen, S. (2019, January 24). Utopia Analytics' global moderation tools clean up the Internet. Retrieved October 16, 2019, from Business Finland website: <https://www.businessfinland.fi/en/whats-new/cases/2019/utopia-analytics-global-moderation-tools-clean-up-the-internet/>
- Schulz, A., Levy, D. A. L., & Kleis Nielsen, R. (2019). Old, Educated, and Politically Diverse: The Audience of Public Service News. Retrieved from Reuters Institute for the Study of Journalism website: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-09/The_audience_of_public_service_news_FINAL.pdf

- Stray, J. (2019). Making Artificial Intelligence Work for Investigative Journalism. *Digital Journalism*, pp. 1–22.
<https://doi.org/10.1080/21670811.2019.1630289>
- The Broadcast Bridge. (2017, November 15). German Football League Adopts Automated Audio Mixing with Live Tracking - The Broadcast Bridge - Connecting IT to Broadcast. Retrieved October 16, 2019, from
<https://www.thebroadcastbridge.com/content/entry/9991/german-football-league-adopts-automated-audio-mixing-with-live-tracking>
- Van de Weghe, T. (2019, March 14). How to make your newsroom AI-ready. Retrieved October 13, 2019, from Medium website: <https://medium.com/jsk-class-of-2019/how-to-make-your-newsroom-ai-ready-74b4bc441b76>
- van Dijck, J., & Poell, T. (2015). Making Public Television Social? Public Service Broadcasting and the Challenges of Social Media. *Television & New Media*, 16(2), 148–164.
- Webb, A., Giraldo, E., Guzik, S., Palatucci, M., & Titlow, J. P. (2019). FTI Trend Report for Entertainment, Media & Technology (No. 3rd annual edition). Retrieved from Future Today Institute website:
<https://futuretodayinstitute.com/2020emt-report/>
- Wu, S., Tandoc, E. C., & Salmon, C. T. (2019). Journalism Reconfigured. *Journalism Studies*, Vol. 20, pp. 1440–1457.
<https://doi.org/10.1080/1461670x.2018.1521299>
- Zamith, R. (2018a). On Metrics-Driven Homepages. *Journalism Studies*, 19(8), 1116–1137.
- Zamith, R. (2018b). On metrics-driven homepages: Assessing the relationship between popularity and prominence. *Journalism Studies*, 19(8), 1116–1137.
- Zamith, R. (2018c). Quantified Audiences in News Production. *Digital Journalism*, 6(4), 418–435.
- Zremski, J. (2019, June 13). UB expert to Congress on “deepfake” deceptions: “It’s likely to get much worse’.
Retrieved October 14, 2019, from The Buffalo News website:
<https://buffalonews.com/2019/06/13/ub-expert-warns-congress-of-deepfake-deceptions-its-likely-to-get-much-worse/>

CONTACT INFORMATION

EUROPEAN BROADCASTING UNION
L'Ancienne-Route 17A
1218 Le Grand-Saconnex
Geneva, Switzerland
T +41 22 717 2220