

FT STRATEGIES

Supported by

Google News Initiative

ARTIFICIAL INTELLIGENCE & NEWS: BUILD OR BUY?

Lessons from the Inaugural AI Launchpad Programme

CONTENTS

<u>3</u>	Introduction
<u>6</u>	Where to start with AI
<u>8</u>	Build vs Buy
<u>11</u>	Publisher Case Studies
<u>16</u>	Conclusion
<u>18</u>	Final Thanks
<u>19</u>	Appendix

INTRODUCTION

INTRODUCTION

Amid an ongoing wave of innovation in the media industry driven by advancements in Artificial Intelligence (AI), FT Strategies designed and delivered the first-ever AI Launchpad programme in partnership with the Google News Initiative.

The programme took place from February to July 2024 and welcomed participating publishers from eight countries across the Europe, Middle East and Africa (EMEA) region.

The AI Launchpad focuses on three practical questions to help publishers leverage AI in their organisations:

- 1 Where do I start and what applications of AI should I experiment with first?
- 2 Should I 'Build or Buy' AI technology?
- 3 How do I embed AI in my business beyond initial experiments and proofs of concept?

About the AI Launchpad

We received over 100 applications for this programme, demonstrating the importance of this topic for news publishers in 2024. The eight publishers selected met the following criteria:

- Existing data and tech capabilities to run AI experiments

- Leadership commitment to support the adoption of AI within their organisation
- Significant digital audience bases (>500k monthly unique visitors)

The programme began with publishers identifying and prioritising possible AI use cases across each participant's news business. Each publisher selected one practical AI experiment to undertake. Some publishers opted to build new tools in-house with support from the Financial Times' data scientists, while others trialled tools from partner vendors sourced by FT Strategies. Publishers who built tools in-house created an initial proof-of-concept and upskilled their teams in the process. Those who used vendors designed experiments to test the return on investment of external tools.

Finally, each publisher participated in a series of in-depth onsite workshops hosted by FT Strategies. Individuals were encouraged to shift from existential deliberation about AI to practical usage and challenges.

The focus was to make AI part of ongoing product and editorial development, including governance and implementation across their organisation. This included recommendations on the next steps for their chosen experiment and advice for upskilling the wider company, including colleagues in the newsroom and product teams.

This report focuses on how a publisher should start their AI journey: should I 'Build or Buy' AI technology?

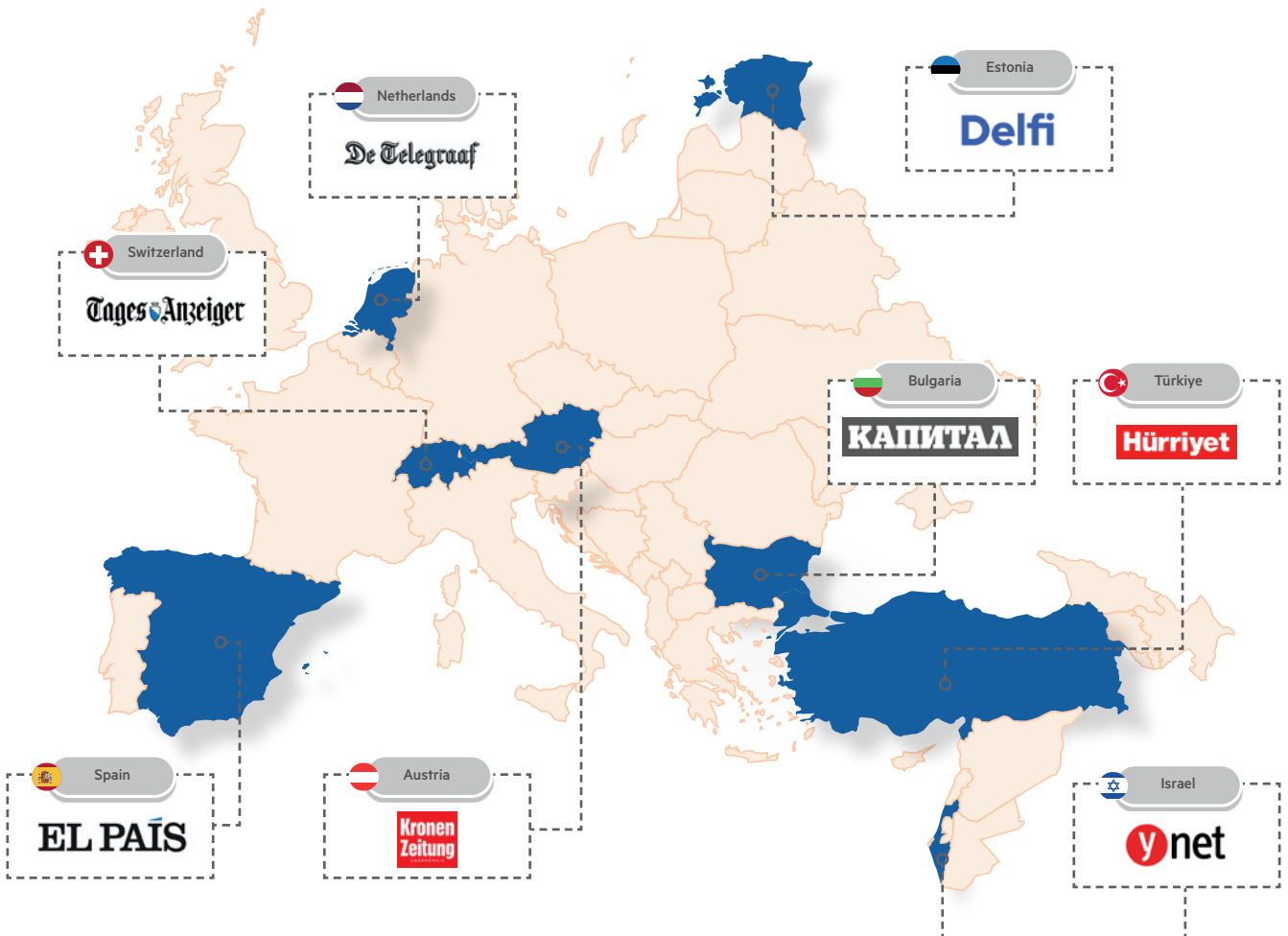
We will spotlight four participating news organisations and the lessons they learned from their experiments. A summary of the other publishers and their experiments can be found in the appendix.

We would like to thank the inaugural cohort of the AI Launchpad for their enthusiasm and active participation throughout the programme:

- **De Telegraaf**, Mediahuis (The Netherlands)
- **Delfi** (Estonia)
- **El País**, Prisa (Spain)
- **Hürriyet**, Demirören (Türkiye)
- **Capital**, Economedica (Bulgaria)
- **Kronen Zeitung** (Austria)
- **Tages Anzeiger**, Tamedia (Switzerland)
- **ynet** (Israel)

Cohort sharing was a particularly important aspect of this programme. Publishers found it helpful to compare approaches with the Financial Times and other participants in an open discussion format.

Before turning to the ‘Build or Buy’ dilemma, we begin this report by considering an important initial question: for organisations looking to leverage AI for the first time, where should they start?



WHERE TO START WITH AI

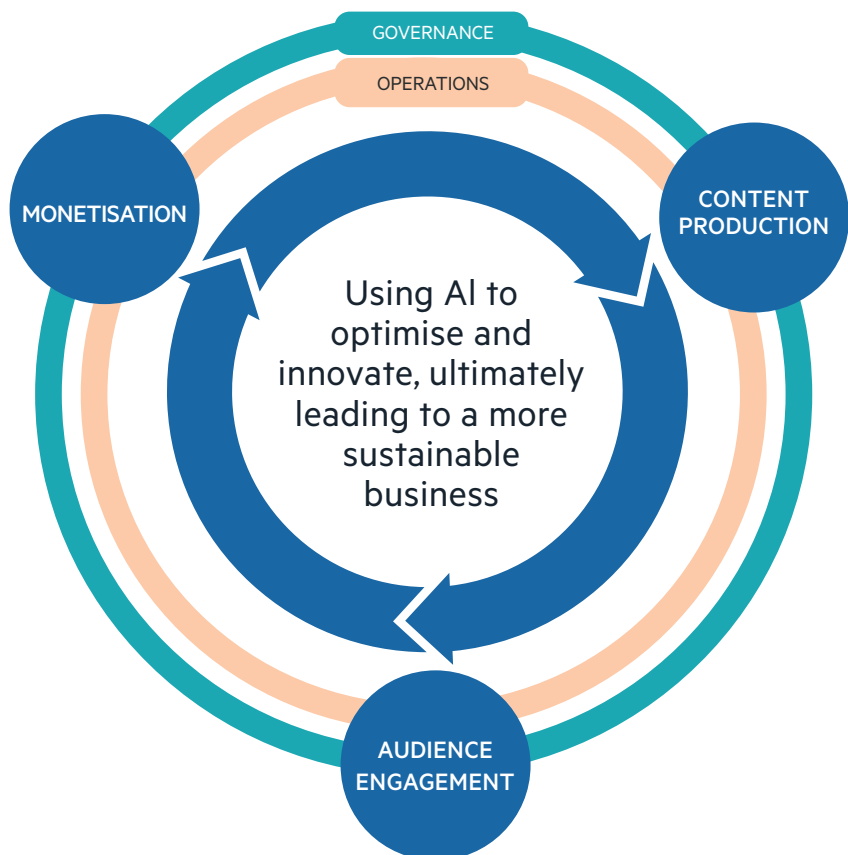
AI has been the central topic of most industry conferences and research papers for the past year, which can lead to a feeling of analysis paralysis. To help ease this anxiety, we developed a deliberately simple framework to help news publishers think about possible AI use cases across their business.

In the framework, we encourage media organisations to first think about possible AI applications in each of the major activities in their businesses which we describe as:

- 1** creating journalistic content (e.g. headline generation)
- 2** engaging audiences with that content (e.g. recommendation algorithms)
- 3** monetising audiences (e.g. churn modelling)

The framework also includes a ring for internal operations, referring to supporting processes necessary for a digital news business to function, as well as another ring for governance, referring to how publishers organise and oversee their AI initiatives.

FTS' framework for AI in a media business

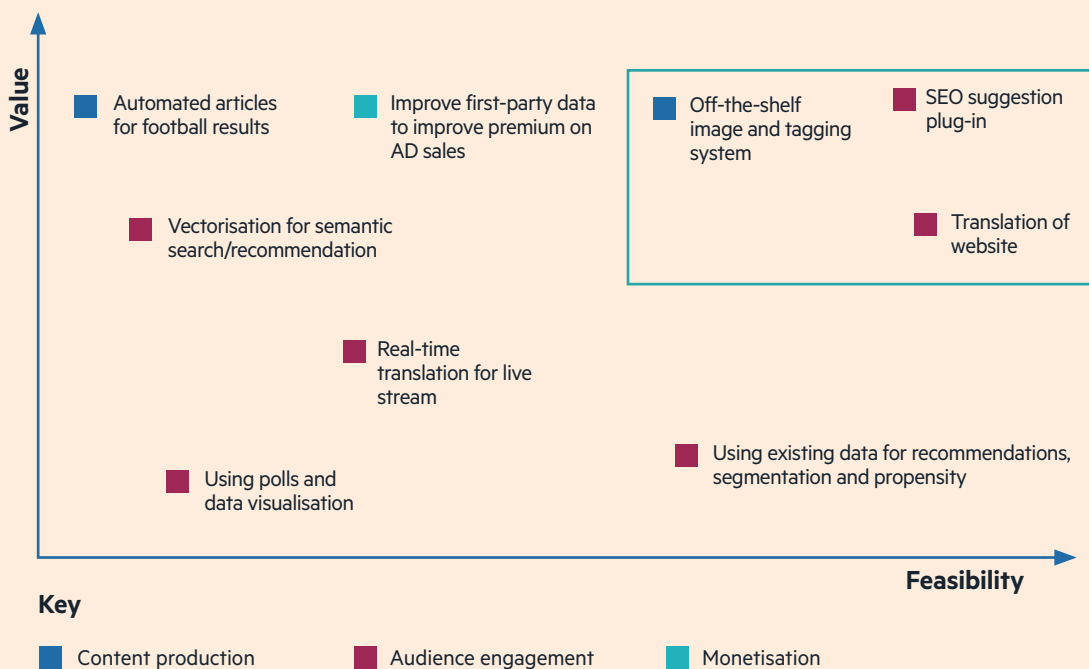


After identifying multiple ideas with each participant, the next step was to prioritise these AI use cases in terms of value and feasibility. For value, we asked participants to consider whether the solution would enable incremental revenue or cost savings, how quickly the use case could be deployed and what the potential impact on audiences might be.

For feasibility, publishers were encouraged to assess whether they had the data and talent required to build each solution. Plotting value against feasibility allowed publishers to zero in on an experiment that could generate a significant impact for the business while being realistically achievable.



An example of a value–feasibility analysis for one participant



BUILD OR BUY

During the feasibility analysis, a frequent question that came up was whether publishers should build a particular solution in-house or buy one instead.

The 'Build or Buy' dilemma is not a new one for media businesses, but it is particularly acute when it comes to AI. To address this, we presented a matrix to the programme participants that summarised the pros and cons of each approach.

BUILD TECHNOLOGY



PROS:

- Autonomy:** Control over the data collected, ensuring accuracy.
- Customisation:** Tailor the data processes to organisational needs.

CONS:

- Resource-intensive:** Creating and maintaining first-party data collection can be expensive, take lots of time, and be high effort.

BUY THIRD-PARTY TOOLS



PROS:

- Efficient:** Usually able to deploy more quickly than building.
- Easily accessible:** Products are available from a range.

CONS:

- Cost:** Can be expensive if you choose the wrong tool for your budget.
- External support:** Needed to implement and manage.

How the 'Build or Trial' pathways worked in the programme:

BUILD (IN-HOUSE)

Build an experiment for a use case that can be piloted through an external vendor partner

Best for: Publishers with very specific "jobs-to-be-done" who have the technical expertise and resources to roll out a bespoke experiment

FTS supports by: Providing weekly sessions with internal subject matter experts where publishers can ask questions about design and implementation

Requirements: Substantial and dedicated engineering resources from participating publishers

TRIAL (VENDOR)

Build an experiment for a use case that can be piloted through an external vendor partner

Best for: Publishers with limited or overstretched technical resources who are looking to build more confidence with AI tools or those interested in trialling new services

FTS supports by: Providing (free) access to existing vendors with clear use cases and team support

Requirements: Onboarding of tool and solution by vendors, potentially integration support

Many of these principles hold for any technology deployment, but there are also some special considerations. In particular, we would highlight the following considerations for building or buying with AI:



Reasons to BUILD



Control

Building in-house allows you to maintain control of your data, a big potential risk when using any external tool or application. This is particularly important in the context of AI and journalistic content: outsourcing a solution to a third-party vendor risks leakage of intellectual property, or breaches in data privacy if customer data is involved. For this reason, introducing third-party tools can often involve a lengthy compliance and procurement process at larger media organisations.

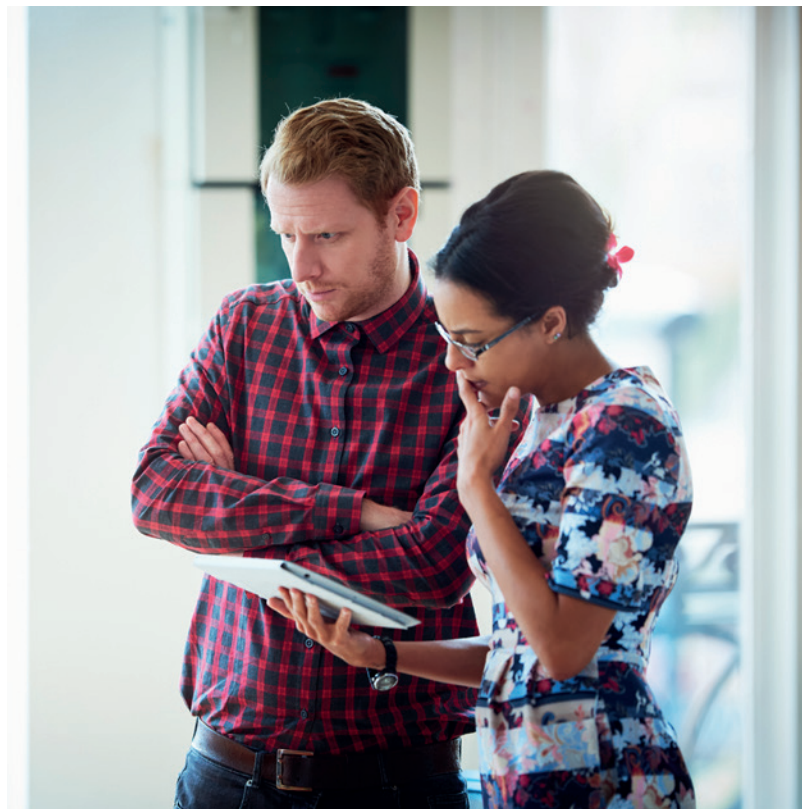
BUT: It is important to flag that when working with AI, even ‘build’ projects are subject to data risk as they may be building on top of an open-source model. For this reason, many publishers are paying for enterprise licences to ensure they are experimenting in a “walled garden” and there is no leakage of their intellectual property.



Customisation

The news publishing business requires highly specialised products, which may not be readily available off the shelf. A good example of this is content management systems, which many publishers build themselves. Customisation is highly relevant for GenAI experiments involving the newsroom as many journalists use a unique combination of tools that is often particular to their ways of working.

For this reason, many journalistic AI projects (e.g. automation of a style guide or assisted newsgathering) have to be bespoke.



Capability-building

When you build an AI solution yourself your team gets smarter in the process. Uli Köppen of Bayerischer Rundfunk calls this “the Trojan unicorn”, meaning that the hidden value of an AI experiment can be the learning that your colleagues gain.

Furthermore, many AI projects possess transferable value: for example, someone who has worked on embedding content can use that experience to develop both internal and external search capabilities for the newsroom or for readers.



Reasons to **BUY**

Cost

Buying a tool could be cheaper than building in-house, which is often a privilege reserved for well-resourced companies. Likewise, many technology companies are rushing to integrate more AI into their product suite. This begs the question: should publishers bother investing time and money in building something that a larger, better-resourced tech company is already working on? In-house projects require dedicated tech talent who can be expensive to hire. Furthermore, unlike traditional product development, AI products are rarely truly 'finished,' as models need to be retrained continuously and costs to maintain models can rise with adoption. This can make in-house projects more resource-intensive over time than they initially appear.

BUT: GenAI has to some degree democratised the product development process — what used to take large teams of developers can now be done more easily, leveraging free resources from websites like Hugging Face and 'Grow with Google'. It can also be hard to understand how expensive an off-the-shelf solution will become once it scales with usage-based pricing models.

Flexibility

Large Language Models (LLMs) are developing rapidly. There is a risk that you may build a solution on top of an LLM that becomes outdated in a few months, whereas you can always end a contract with a third-party tech vendor with minimal sunk cost. BloombergGPT is a cautionary tale of building too soon on a model that is now outdated. Of course, rapid innovation is also a challenge when buying tools: for example, you may invest in a product that later becomes part of a tech giant's offering, but, as mentioned, the switching cost is a lower risk.

Time to market

A plug-and-play solution is often faster than the research and development needed to build something similar yourself. As Melle Drenthe, part of Mediahuis' AI Launchpad team, put it, "I can tell a product is valuable because I always think 'Is this something I could build myself?' And, in this case of this tool, it would have taken me five years to build." By contrast, to get started with a text-to-speech tool, all you have to do is start a free or paid trial.



Hybrid options:

We realised in this programme that sometimes publishers 'buy to build.' Trialling or using an off-the-shelf solution can be a useful way to assess a use case before investing money in a 'build' project. Using a vendor can also be a way of validating an idea and then evaluating whether it makes sense to continue outsourcing or to use it as inspiration for in-house development.

The next section looks in more detail at the experience of participants in the programme of addressing the 'Build or Buy' dilemma.

PUBLISHER CASE STUDIES

КАПИТАЛ

Capital (Economedica) - Streamlining article summaries to increase editorial efficiency

**Publishers
who
BUILT**

The situation

Capital, part of Economedica, is the largest business-focused title in Bulgaria. It was the first Bulgarian title to launch a paywall in 2017 and leads the market with over 10,000 subscribers.

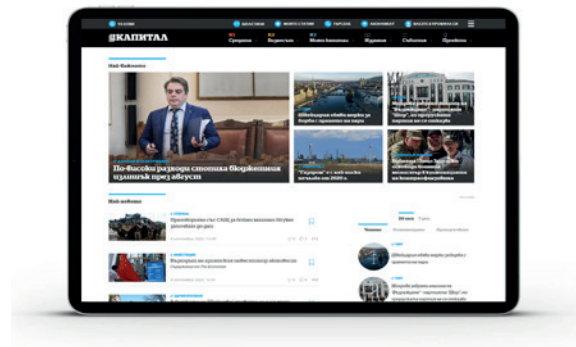
Capital is known for economic coverage, in-depth investigative journalism and analysis. Some of the longer, sophisticated articles can be a barrier to engagement. Recognising this, the editorial team manually writes bullet point summaries that appear at the top of articles over 2,000 words. This task is time-consuming, and writers would rather spend time focusing on the breadth and depth of their coverage. They also saw an opportunity to create paragraph summaries to be used in their newsletters.

AI Use Case

The Capital team turned to AI to generate these summaries. Beyond creating efficiency, they hoped that this could also lead to more reader engagement, encouraging more people to read the full articles.

Working closely with data scientists from the Financial Times, the team built a summarisation tool from scratch. They decided to build this in-house because one of their key goals for the programme was to upskill technically and grow in confidence with AI. Additionally, they realised that there are very few off-the-shelf solutions fit for purpose in Bulgarian.

As a first step, Capital chose to evaluate the quality of summaries generated by six different LLMs. Prior to this project, there was no published research comparing the accuracy of prompting in English and Bulgarian on popular LLMs, or with the Bulgarian BgGPT, a model developed with funding from the Bulgarian government. Capital found that existing LLMs effectively summarise articles in English, but are limited when Bulgarian material is inputted. The evaluation process involved blind-testing by over 20 journalists who evaluated the results of each LLM via detailed scoring. Tests for each LLM were conducted with prompting in both English and Bulgarian.



The tech team also ranked each LLM with a ROUGE score (a comparative score between a reference and an automatic summary) and a TF-IDF (a measure of the importance of a term within a document).

Results

Ultimately, the team found that Gemini 1.5 (prompted in Bulgarian) and Anthropic Haiku (prompted in English) were the LLMs that generated the highest quality summaries. As a next step, the team will be rolling out the tool to journalists, integrating it into their Content Management System, and capturing changes journalists make in the CMS as a means for retraining the model. In the long term, the business will continue to refine the prompt and understand what product angle could be possible for these summaries, such as using them in reader-facing newsletters or podcasts.

The decision to build in-house means that the engineering team now has a far improved technical understanding of AI. A conscious effort was made to build in a model-agnostic way so that future LLMs can be evaluated and potentially incorporated into the solution. This was also a valuable exercise as it brought together editorial and technical teams, proving to editorial that AI projects can be accessible and relevant to their work.

**The programme helped us
demystify AI and learn how
to evaluate opportunities**

GALYA PROKOPIEVA, CEO, ECONOMEDIA



Hürriyet - Implementing automation to report on earthquakes

The situation

Hürriyet is a major Turkish news publication, owned by the Demirören Group. The company currently makes most of its revenue via advertising but is considering launching a subscription. This would be a pioneering move as the news subscriptions market in Turkey is still nascent.

During the programme, the team decided to experiment with AI-assisted content creation, something that could help with two of their business goals: increasing website traffic and improving editorial efficiency. Hürriyet often writes about earthquakes due to high tectonic activity in the region. The team thought these would be good articles to work with since they tend to follow a similar format. They also saw value in automating parts of the process, given it can be hard for journalists to quickly create earthquake articles due to the volume of information and unpredictable timing of such events. It takes a considerable amount of time and manual labour to continuously monitor data and create articles from scratch each time.

AI Use Case

While the Hürriyet team has a flexible approach, often working with external tools, they saw this as an opportunity to build a custom solution themselves. Data was readily available thanks to AFAD (Disaster and Emergency Management Authority), a public database where essential data is first reported whenever there is an earthquake in the country. The team's more technical members also saw this as an exciting intellectual challenge.

With mentorship from FT Data Science, the team created a fixed template pulling AFAD data such as earthquake location and magnitude. They then engineered a system prompt to transform this data into structured articles, with input from editorial.



Publishers
who
BUILT

Next, the team experimented with different LLMs to see if GenAI could be used to write the more descriptive parts of each earthquake article. They asked members of editorial to score the outputs, comparing them with existing human-written articles in a 'blind test.' The team evaluated six LLMs in this way, scoring against criteria such as Turkish language quality, accuracy and detail.

Results

The first output of the project was the article template which can now be used as a guide for all further earthquake articles. For the generative aspects, the team identified ChatGPT 4o as the most successful model based in Turkish and was able to create several articles that were highly rated by editorial.

The next part of the project will be integrating the system prompt into the CMS which can then be used by journalists when they are writing earthquake articles. In the long term, the plan is to evolve the model to require less human intervention. Another ambition is to develop further article features such as a timeline of updates. As emergency news is a high-risk area for experimentation, due to possible model hallucination, the next step is to responsibly add in more GenAI aspects and see if the technology can be trusted to write articles instead of just pulling data and templated aspects.

We learned a lot from you guys because you showed us what kind of work has been done all around the world. We learned and saw a lot so we changed our strategies... this had a big impact on us

**MUSTAFA ARSLAN, GROUP DIRECTOR OF PRODUCT,
DEMİRÖREN**

EL PAÍS

El País - Using text-to-speech to create authentic audio experiences



Publishers who
TRIALED

The situation

El País is the flagship news publication of Prisa, a leading global media group with a market-leading position in Spain and entities in Latin America. In addition to newspapers, the company also owns several radio stations and popular podcasts such as 'Hoy en El País.'

The El País team saw an opportunity to see if AI could help their audio team create more content, and whether AI could actually create a quality audio experience. Audio is a popular format, particularly for young audiences, so this could potentially help engage this growth segment. Currently, journalists and audio teams work to turn articles into audio material in dedicated recording sessions in the studio. This is time-consuming and requires high-quality recording equipment.

AI Use Case

The team decided to test whether AI could be used to quickly and accurately create a new podcast. Specifically, they aimed to create a daily podcast that reads the Op-ed section of El País and a short daily audio roundup of key headlines.

Partnering with Wondercraft, a company specialising in AI-generated audio, the team created a proof of concept for automatic podcast generation, converting text-to-speech, and turning written articles into podcasts with a synthetic voice. The team decided to

go with an external vendor to fast-track the experimentation process. The 'buy' decision is in line with a wider company approach to experiment with external tools where possible, to keep pace with rapid AI advancement.

This project started with the challenge of creating an authentic voice. Wondercraft offers several ready-made voices in different Spanish accents, but the team felt that these options were overly artificial sounding, and often overly enthusiastic for news coverage. Using Wondercraft, they created two new synthetic voices based on real journalists' voices to achieve neutral accents and an authentic tone.

They then used these synthetic voices to create sample podcasts, starting with two opinion pieces and two morning news roundups. These samples were evaluated by a panel of editorial judges. Each of the four podcasts was ranked in terms of, for instance, quality of Spanish, style of voice, hallucinations, etc.

Results

The scores from these experiments showed positive feedback for most of the podcasts. The team was even able to share the podcasts with selected subscribers. This focus group was asked to listen to one using an entirely synthetic voice and one with a human-trained voice. Subscriber feedback for the wholly synthetic voice was negative whereas, the subscribers said they liked the voice that was based on those of

real Prisa journalists and reported that they did not mind that the voice had been synthetically created.

Success was also measured in terms of efficiency gains. These were somewhat lower than expected. Initially, the time to produce a podcast has only reduced by 25%. Some teething issues result in further editing. For example, the tool struggles to pronounce some names and words from other languages, such as presidents of other countries, and these then have to be fixed in post-production. That said, the experiment validated the idea that AI could help to produce net-new podcasts with the same-sized team.

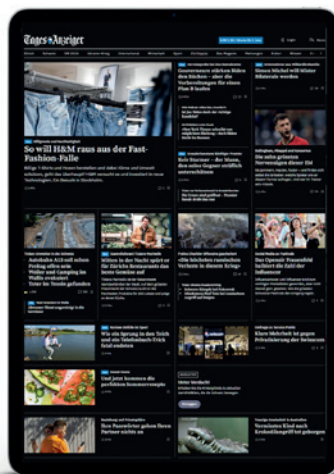
Now that the team has developed a voice they are satisfied with, the next step is deciding whether this can be used in a customer-facing product. The team is in the process of evaluating ethical AI guidelines and where this use case sits. They hope to continue trialling Wondercraft and refining results to ultimately launch an AI-enabled podcast. Finally, the experiment proved that their editorial team is open to further experimentation, something that was not taken for granted at the outset.

This programme is like a master's of nonstop learning.

ALVARO GONZALEZ RIPOLL VIGUERA,
CHIEF DIGITAL OFFICER,
EL PAÍS & HUFFINGTON POST

Tages-Anzeiger

Tages Anzeiger - Deploying user needs content analysis to inform editorial strategy



Publishers who
TRIALLED

The situation

Tages-Anzeiger is Tamedia's flagship title and a leading publication based in Zurich, Switzerland. The Tamedia team joined the programme to accelerate Tages-Anzeiger's digital transformation, boost newsroom efficiencies and increase reader loyalty.

To achieve this, they first wanted to understand content production and performance from a 'user needs' standpoint, providing newsrooms with insights to deploy resources more efficiently. The goal of the user needs framework is to classify content based on what the reader gained from the article, e.g. 'update me' or 'educate me'. Tamedia wanted to better understand their holistic content performance with the ultimate goal of producing more engaging content. Furthermore, Tamedia identified a company-wide priority of adopting a user needs mindset in their newsrooms.

Specifically, the aim was to attract fly-by users with a better content mix. With more engaging content on their site, these users would theoretically become loyal and be inclined to subscribe. Furthermore, Tamedia identified a company-wide priority of adopting a user needs mindset in their newsrooms.

AI Use Case

In collaboration with SmartOcto, a smart content analytics system designed for newsrooms, Tamedia applied the user needs classification to a representative sample of 5,000 of Tages-Anzeiger's articles from the past 7 months. They felt that trialling an external AI tool would be an efficient way to kick-start this process.

Initially, this involved categorising content into relevant 'user needs' groups and the compilation of a report by SmartOcto. Subsequently, how these user needs performed was analysed by connecting them to performance metrics such as page views. Finally, the Tamedia team conducted further research by segmenting the analysis through the lens of specific user cohorts (fly-by, loyal, heavy).

A tool could make things more efficient: AI tagging is instant whilst manual tagging means reading the article and can take up to 10 minutes. But, would the results be reliable? A further aspect of the experiment was understanding the accuracy of Smartocto's automatic tagging versus that of humans.

Results

The user needs analysis enabled the newsroom to understand content performance in a new way. The analysis showed that there was a content imbalance with context-driven ('understand') and fact-driven ('know') articles being produced in significantly higher quantities than emotion-driven ('feel') and action-driven ('do') articles, which ultimately performed better. On average, 'Do' and 'Feel' articles have 40% and 20% more views than 'Understand' articles. However, 'Feel' and 'Do' articles only accounted for 10% and 4% of all articles.

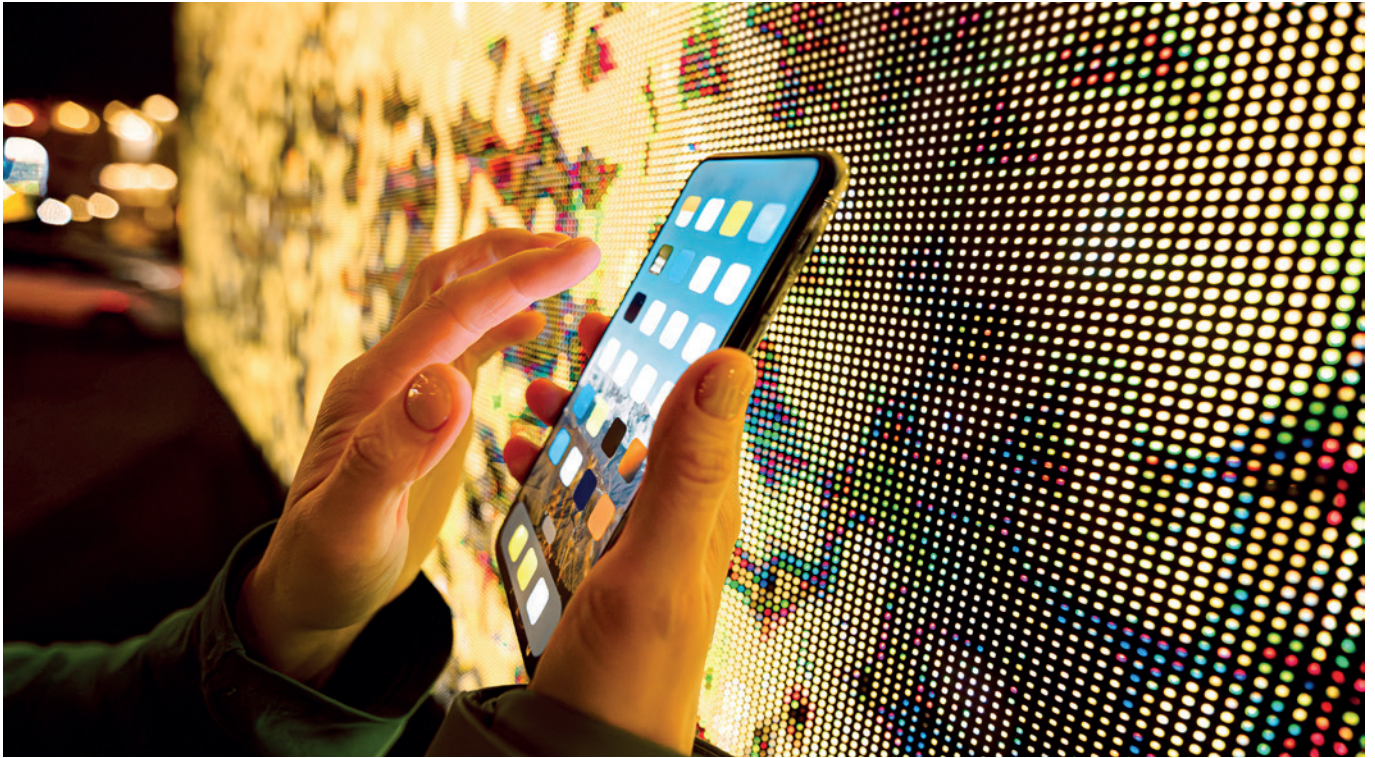
Tamedia found that Smartocto's AI and human labellers often took a different view of what user needs category an article falls into and were consistent with each other for 60% of articles. Specifically, Smartocto's AI labelling was more likely to label an article as 'understand', relative to the human labellers. This is not to say that one set of labelling is more or less accurate than the other but highlights the need to adopt a clear classification and apply it consistently.

Looking forward, the team would like to create a live user needs dashboard for use by the editorial team. Now that they have validated this approach, Tamedia is evaluating their options for how they can best operationalise user needs. The team is also hoping to integrate this framework into the CMS so that user needs are considered holistically from story ideation to daily monitoring of content mix.

The AI Launchpad programme has given us incredibly inspiring insights and useful information. It's been a really beneficial programme!

NADIA KOHLER, HEAD OF AI LAB, TAMEDIA

CONCLUSION



CONCLUSION

As the experience of these four publishers demonstrates, there is no one answer to the question of ‘Build or Buy.’ The right approach will depend on the individual circumstances of each company, where you are in your AI journey and the scale and expertise of your in-house technology and product teams. The best way to approach this problem is by asking yourself a series of questions as you consider AI deployment in your business.

To begin with, we recommend answering the following questions:

- 1 Do we have the **time** to build this ourselves? How urgent is this project?
- 2 Do we have the talent and **capability** to build this ourselves? If not, do we want to learn that capability or outsource it?
- 3 Is there a data or other **risk** of using a third-party vendor?
- 4 What will we do if/when technology improves in a few months? Are we building enough **flexibility** into our process?
- 5 What is our **budget**, and do we have a sense of how costs could increase as we scale the solution?

This is a fast-evolving space and, due to growing interest in how to deploy AI successfully in news businesses, FT Strategies and the Google News Initiative will be running further editions of the AI Launchpad in the region this year.

We look forward to sharing further lessons and reflections with you later this year, with a focus on practical AI implementation in news media businesses.

ABOUT FT STRATEGIES AND GNI

FT STRATEGIES

FT Strategies:

FT Strategies is a boutique consulting firm within the Financial Times, staffed by the experts who successfully transformed the FT's business model in the face of Disruption. FT Strategies powers customer growth by providing the expert insight, integrity and consulting excellence required to transform the future of business based on real-world, first-hand experience. FT Strategies has worked with over 700 organisations globally by helping them innovate using data and helping build sustainable digital businesses.

The Financial Times:

The Financial Times is one of the world's leading business news organisations, recognised locally and internationally for its authority, integrity and accuracy. The FT has a record-paying readership of more than 2.4 million, ninety percent of which is digital revenue. It is part of Nikkei Inc., which provides a broad range of information, news and services for the global business community.

Google News Initiative

Google News Initiative:

In collaboration with editors and journalists worldwide, the Google News Initiative (GNI) strives to foster a more sustainable, diverse, and innovative news landscape. Through a comprehensive suite of programs, digital tools, training, and resources, the GNI empowers journalistic endeavours, enabling journalists to effectively discover, verify, and narrate compelling stories while upholding the highest standards of quality journalism.

We would like to thank our partners at the Google News Initiative for making this programme possible as well as the inaugural participants featured in this report.

We would also like to thank our partner vendors who were part of these projects (Smartocto, Dataminr, Wondercraft and Bridged Media) as well as those who joined us to demo use cases during the programme (Pixels.ai, Viafoura, Miso.ai, Labelbox and InsurAds). These companies generously waived their licence fees to allow publishers to trial the basic functionality of their tools free of charge for the duration of the programme.

Contributors

Jim Egan

Principal, FT Strategies

Aliya Itzkowitz

Manager, FT Strategies

Liat Fainman-Adelman

Manager, FT Strategies

Isabella Croes

Senior Consultant, FT Strategies

Jack Briggs

Consultant, FT Strategies

Shak Ebrahimjee

Consultant, FT Strategies

Carter Bedford

Consulting Intern, FT Strategies

Caroline Barylak

Senior Graphic Designer

Sponsors

Sulina Connal

Managing Director for News and Books Partnerships for EMEA Google

Matt Cooke

Head of Google News Lab

Benedicte Autret

Head of News Partnerships - UK/IE/Northern Europe, Google

Martin Schierle

Manager, Technical Solutions, Google

Alexandra Stafford-Blythe

Program Manager, News Partnership Programs, Google

APPENDIX

De Telegraaf

Mediahuis - Harnessing AI for news discovery to expand coverage

The situation

De Telegraaf is a Dutch general news publication, owned by the Mediahuis group.

The company currently has an organisational focus on better understanding and driving reader engagement. The team decided to see if they could use AI to discover news stories that they and their competitors might otherwise have missed.

AI Use Case

De Telegraaf tested Dataminr, a vendor that leverages advanced AI and machine learning techniques to analyse vast amounts of publicly available data from sources such as social media, news articles, blogs and police radio. The team chose to experiment with a vendor due to the time and resource-consuming prospect of building such a tool from scratch.

Over 15 journalists from De Telegraaf's various news desks received two months of access to the Dataminr tool to see whether this could inspire the creation of unique pieces of content and/or improve news discovery efficiency. Independent creativity remained an important aspect of the newsgathering process throughout the experiment.

Journalists searched for international breaking news stories. Some reporters also made custom lists to track events such as the Euro 2024 football tournament.

Another aspect of the experiment was to see whether Dataminr was appreciably faster than current alerting systems used by the newsroom such as newswires.

Results

During the trial period, journalists engaged well with the tool. This had been a concern at the beginning as it can be hard to change workflows in the newsroom. Dataminr alerts inspired the creation of at least 10 articles during the trial period.

While Dataminr was not appreciably faster than other sources (one to two minutes), it did surface some stories that were unique. For example, one news story about an arrest in Germany involving a Dutch national was widely read.

A possible next step is to try Dataminr with other Mediahuis titles for whom breaking international news is important. The team will use audience engagement data to assess whether the tool could provide a positive return on investment.

It was a great learning path to generate all the do's and don'ts from all the different international media organisations. Very valuable.

MARIËLLE VERMEER, HEAD OF DATA & SOCIAL, DE TELEGRAAF





Kronen Zeitung - Automating headline generation to streamline newsroom operations

The situation

Kronen Zeitung is a tabloid-style publication and Austria's most widely read news publication. Given their reach, advertising is a key revenue driver. Consequently, traffic and page impressions have a significant business impact on their overall business strategy.

Headlines are a key factor in a reader's decision of whether or not they click on an article. Creating catchy, concise, and effective headlines and subheadlines is incredibly important, but takes up editors' valuable time.

AI Use Case

The Kronen Zeitung team developed a bespoke tool to generate headlines and subheadings based on the content of each article. They decided to build the tool in-house, as opposed to partnering with an external vendor, to build capabilities internally while ensuring a human is kept in the loop for all content production. Building in-house also allowed the team to create a custom solution where they could 'tune' the specific style and tone of voice directly.

With guidance from FT Data Science, the Kronen Zeitung team initially designed a prompt and tested it with different LLMs in order to decide which model summarised and produced a headline the best.

Once the model was selected, they worked with an FT Data Scientist to align on a prompt engineering strategy to ensure accuracy. For example, the Kronen Zeitung team consistently found the titles were too long at first, exceeding the limit of 50 characters. They experimented with different prompts that included both threats and rewards until they achieved their goal. Once satisfied with the tool, it was tested with editors, who accepted or declined the AI-generated headlines. Their goal was to obtain an 'acceptance rate' of 20–30%.

Results

From testing LLMs, the team found that OpenAI's models worked best in terms of speed and quality of the recommendation. GPT4.0 provided the most accurate results, with 90% accuracy, though significantly slower than the GPT3.5 model which was accurate 50% of the time. They also found that prompting in English and then translating to German was more accurate.

The team successfully rolled out the tool to 10 editors across various departments who continue to provide feedback on the tool. This has also proved to be a great opportunity to educate editors on prompt engineering.

If we didn't do this programme we wouldn't have progressed as fast, it's really helped us get things done. With you, we have a clear commitment which forces us to move ahead

MANUEL ROSTORFER,
CHIEF OPERATING OFFICE

The team plans on conducting further analysis until a sample size of 800 to 1,000 tests have been carried out. Once the tool has been rolled out to all departments, the AI-generated headlines will be monitored for page impressions to see if there's a notable increase versus human-generated articles. Editors will also continuously be asked to provide feedback on how the tool can be made more efficient.

Reflecting on the decision to build this themselves, the team was pleased with the progress they made. They also shared a couple of learnings for other publishers. For example, they noted that obtaining consistent feedback from editors was more challenging than anticipated as different teams have different ideas of what success is. They also emphasised the risk of hallucination as sometimes off-the-shelf solutions are better equipped to track these. For in-house projects, extra time needs to be built in to validate the model's outputs.



ynet: Enhancing SEO through AI-generated headlines and content summarisation

The situation

Launched in 2000, ynet is a leading digital news site in Israel, providing news to over 10 million readers each month. ynet sees AI as a tool that can unlock creativity and boost efficiencies in the newsroom by freeing up editors' time.

ynet's first goal was to improve search engine optimisation (SEO). Given their business model largely relies on traffic on their homepage, they saw an opportunity to increase referrals. Up until recently, SEO suggestions had been created manually by their Content Strategy Manager which was effective but took a substantial amount of time.

Their second goal was to leverage article summaries to help various editorial and social media teams quickly understand the article's context and background information. These summaries could also be used to produce social media content and WhatsApp messages more efficiently. They also wanted to test the quality of the summaries to see if they could be rolled out within audience-facing content directly.

AI Use Case

The ynet team partnered with Bridged Media to improve their editorial workflows by introducing automated SEO headlines and article summaries.

With their headlines, Bridged Media's AI agent was able to understand the context of an article and use an API (application programming interface) to find which parts of the article were trending. Those keywords were then used to suggest a headline. It also had an SEO schema markup (the language search engines use to understand a web page) which can be copied and pasted into the website backend, making the article more attractive to search engine rankings. ynet and Bridged Media also partnered to experiment with a summarisation agent. The summaries produced were configured as a short paragraph or a specified number of bullet points, and the teams worked together to fine-tune the quality and tone of the output through trial and error.

Language constraints have traditionally been a barrier to using external vendors. However, ynet was impressed by Bridged Media's Hebrew capabilities. ynet ultimately decided to work with an external vendor largely to get the experiment off the ground faster.



Both the programme and the experience we've had with the experiments were amazing - we were really surprised to see how much progress and how feasible it really is.

DROR AMIR, CHIEF PRODUCT OFFICER, YNET

Results

Overall, ynet deemed the project with Bridged Media a success. While more testing is required to assess the long-term impact on SEO, initial results showed that most topic segments that participated in the experiment saw an uplift in traffic compared to the previous month. For example, the Vacation, Food and Fashion sections saw the proportion of traffic coming from search increase by 3%, 3% and 15% respectively. That said, the Parents and Dating sections saw minimal uplift or a slight decline. Beyond uplift, ynet found that the headline generator saved a considerable amount of time for their Content Strategy Manager and promoted a culture of innovation and openness to using AI in the newsroom.

For the summariser, editors are seeing value in certain use cases, such as captions for the social media team. One editor who participated in the experiment shared that the tool was easy to use and sparked curiosity. However, they see room for improvement as the current process requires a lot of editing. ynet plans to continue working with Bridged Media to update and improve it over time and is also exploring a reader-enabled feature which would allow audiences to request article summaries.

Delfi

Delfi - Understanding audience engagement through smart content analytics

The situation

Delfi is the Baltic region's leading news portal, hosting some of the oldest and most reputable titles across Estonia, Latvia and Lithuania. One of their business objectives is to boost audience and retention engagement given their strong subscriptions business. Consequently, Delfi was interested in implementing a user needs approach to better serve their readers. They also wanted to promote an audience-centric mindset in the newsroom.

AI Use Case

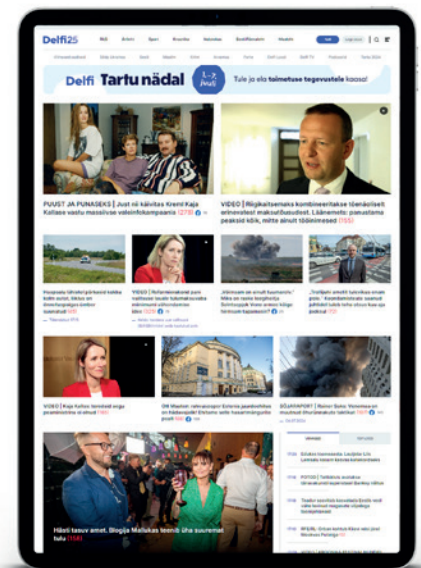
Delfi carried out two experiments. The first was around their content mix. They hypothesised that there was potential overproduction of articles in some user needs and wanted to validate this. The second experiment was around A/B testing. Delfi previously did not have the capabilities to A/B test features (such as headlines, leaders, and the position of articles) on the homepage, and was looking for a solution that had that functionality. The goal of A/B testing was to experiment with these different features on the homepage, in order to improve click-through and conversion rates. They also wanted to see if there were consistent patterns in how articles were displayed which their audiences found more engaging and led to them interacting with more content on the website.

Delfi decided to partner with SmartOcto, a smart content analytics system that applies user needs classification to historical and real-time content production data. Delfi applied Smartocto's tagging to 5,000 of their historical articles to understand whether they were over or under-producing articles with certain user needs. In the background, Delfi manually tagged 300 of those articles to cross-compare the accuracy of the model. They also implemented an extension that allowed them to A/B test the homepage.

Results

In the user needs experiment, Delfi found that 80% of manually tagged articles matched the automated tagging, giving Delfi confidence that the model could be used moving forward. When looking at the results, the analysis found that Delfi was overproducing fact-driven ('update me') and context ('educate me' and 'give me perspective') articles, with 70% of all of their content falling into these categories. Inversely, the under-produced emotion-driven stories tended to perform better, with higher scroll depth scores.

Moving forward, Delfi hopes to integrate the user needs tool directly into the CMS to allow editors to better understand the type of content they are producing in real time. They will also do a more comprehensive analysis of all archived content. Finally, they plan to train all editorial staff on the user needs methodology.



Your project has been really eye-opening and showed us that there are useful tools that media houses can use to grow their businesses

PIRET PÕLDIJA, HEAD OF BUSINESS DEVELOPMENT AND MEMBER OF THE MANAGEMENT BOARD, DELFI

