AI4media

VRT Innovation

AI for News The Smart News Assistant

White paper - October 2022

Chaja Libot, Rik Bauwens, Paulien Lemay

Executive Summary

The consumption and production of news has changed a lot over the past years. The workload for news creators (journalists, editors, etc.) is ever increasing. The development of AI-tools could help creators in continuing to deliver trustworthy and relevant stories to a broad audience.

One of seven use cases in the AI4Media project aims to explore the new opportunities AI can offer for journalists and content creators. Apart from testing the technology provided by partners in the project, this use case also focuses on researching the challenges and needs of professional users.

This white paper is intended for AI researchers and technology developers with an interest in providing new research and technology related to supporting workflows of news professionals in an ever changing, challenging digital environment.

The white paper summarizes the results of the use case work and describes the challenges and user requirements of different tasks within the workflow of a news professional, ranging from the research and production to the publication phase. We see the use case as a 'Swiss Army Knife' for journalists, as it incorporates possible solutions and tools throughout their workflow.

We combined survey insights with user research done in VRT NWSlab. NWSlab is a cross-functional team consisting of user researchers, developers, journalists, editors and creative profiles at VRT NWS (the news department of the Flemish Public Broadcaster). Based on the survey results, one of the solutions, a comprehensive toolbox for fact-checking, was found to be the most relevant solution for the industry. This white paper also briefly analyses the importance of AI functions that are trustworthy and transparent for non-technical users.

Key messages

- There is a clear opportunity for AI tooling to facilitate mundane and burdensome tasks, giving more space to creativity and original investigative and informative work

- Because of the fragmented information landscape monitoring assistance is of interest to journalists.

- The fact that journalists are increasingly confronted with disinformation results in a need for an understandable, accessible and easy-to-use fact-checking tools.

Table of contents

2

Introduction

Validating the vision

In today's context, journalists are working in an extremely challenging environment. The rise of (social media) platforms has led to the need to publish as fast as possible to the extreme.

More people, especially young people, are using social media as a news source. They don't use a news app or website as the primary go-to place to catch up on the latest news. Instead, they resort to social media and private messaging services such as WhatsApp, Messenger and Telegram. As can be seen in Figure 1 below, especially people aged 12-34 years use social media as their primary news source.

Reuters' journalism, media and tech trends and predictions report of 2022 found that more publishers are putting effort into Instagram (+54), TikTok (+44) and YouTube, all networks that are popular with younger people.

Thus, news professionals **need to adapt their publication strategies and publish on these platforms to match media consumers' expectations.**

At the same time, j**ournalists need to ensure the published content is both relevant for their audience, and is a trustworthy source of information, avoiding errors and misinformation.** For example**, it's important to make sure correct information about COVID-19 is circulating among the population to make sure public health is not put at risk.**

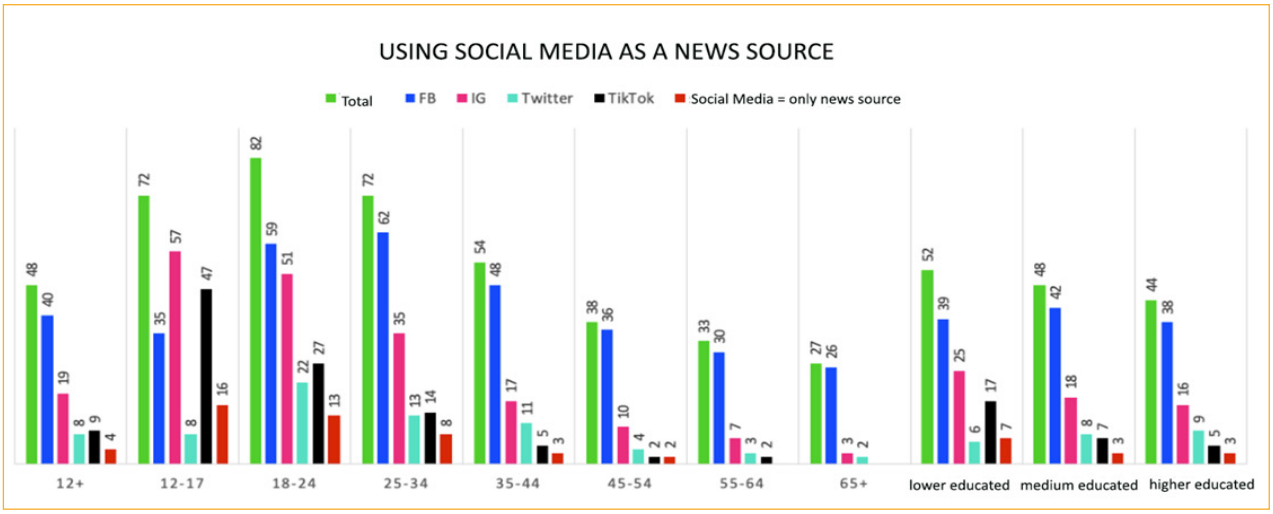

Figure 1 . Use of social media as a news source in percentage per age group - VRT Studiedienst 2021

Innovation development and research in the AI4Media project

AI4Media is a project funded under the European Union's Horizon 2020 research and innovation program. It aspires to become a Centre of Excellence engaging a wide network of researchers across Europe and beyond, focusing on delivering the next generation of core AI advances and training to serve the Media sector, while ensuring that the European values of ethical and trustworthy AI are embedded in future AI deployments.

AI4Media has designed seven use cases that demonstrate the use of AI in the media industry, including aspects of human-centric, ethical, and trustworthy AI. The AI4Media use cases (UC) are inspired by market needs, emerging opportunities, and a range of industry challenges, and they highlight how AI applies throughout the media and content value chain and how different types of media players aim to address user and business needs with novel AI solutions.

The Vlaamse Radio- en Televisieomroep (VRT) is the public broadcaster of the Flemish Community in Belgium. News and information are the core business of VRT as a public broadcaster, including research, creation, curation and publication of news content. "To inform" is one of the three key objectives of VRT's public mission, along with "to inspire" and "to connect".

The use case of VRT within AI4Media focuses on the concept of a Smart News Assistant, i.e. a tool that will support journalists in the creation of news stories by providing a variety of AI-enabled functionalities for story production and development, story curation and publication, and audience engagement.

The tools developed and validated in this use case will facilitate the digital transformation towards data-driven newsrooms. This is part of the aim to become a data-driven media organisation in which building cross-brand (i.e. beyond news) and personalised media journeys is one of the core objectives. As a public broadcaster, VRT also aims to be a frontrunner in the domain of countering disinformation and fake news. Moreover, it is well-placed to setup and drive collaborations in the Flemish media sector.

VRT's use case addresses the need for building a multi-functional and AI-driven toolbox, called Smart News Assistant. We are currently developing different solutions and tools, as part of the Smart News Assistant, together with news professionals working with us in NWSlab.

We do this together with journalists and media professionals, because we observed a big gap between technology-driven projects – technology-driven project - like AI4Media - and media professionals VRT NWS in the newsroom of VRT. That is the reason why **VRT NWSlab** initiative was created: **a cross functional team embedded in the newsroom that consists of developers, user researchers, journalists, editors and creatives.**

User research done within NWSlab gave us more insight into the workflow of news professionals. We learned how to translate users' needs into intelligent solutions in the form of multi-functional AI-enabled tools. This helps news professionals create relevant and trustworthy stories more efficiently. We focus on the different steps in the journalist's workflow.

We distributed the survey to journalists, editors, news strategists and creatives and received 49 responses, 32 internally from VRT and 17 via the Future Media Hubs. which is an international network of (public as well as commercial) media organisations that focuses on innovation through collaboration and knowledge sharing. Members include for example BBC, RTBF, NRK and YLE. Via VRT NWS, the news department of VRT, we received a total of **32 responses.** The 49 respondents work in the role of journalist, editor, news strategist or creative.

3

What are the challenges and needs?

In our discussions with journalists and during our observations of their current workflow for news research, production and publication, we discovered there is a need for support in the following tasks:

1 The amount of incoming content to work with is increasing overwhelmingly. There is a need for monitoring and assistance in curating what's worth investigating.

2 Fact checking is needed to avoid misinformation. Journalists are looking for ways to help them check information. A lot of tools and solutions do already exist. But currently, only experts, e.g. fact-checkers, are using them (because they tend to be complex), even though all journalists feel the need to check the information they plan to share.

3 Short videos are gaining ground, and publishers need to adopt these techniques (Journalism, media an tech trend and predictions 2022 – Reuters) Having automated (audiovisual) content suggestions could help the journalists make their content more (visually) appealing, which is something they currently don't have time for.

4 To meet high standards in diversity and inclusion of modern newsrooms, having diversity assessment can help in writing balanced stories.

5 The fragmented media landscape requires news creators to be present on an increasing amount of diverse (social) media platforms. Every one of these platforms has its own way of storytelling. It is not enough anymore to transform a story; it also needs to be translated to fit various media formats. Tools that can make a first edit or rough cut would be a welcome addition for journalists.

We defined the following domains within our Smart News Assistant use case for the AI4Media project (see figure 2 below). We see the use case as a 'Swiss Army Knife' for journalists, as it incorporates possible solutions and tools throughout their entire workflow.

Figure 2. Solutions plotted on workflow journalists

4

Validation challenges and needs

We discovered there is a need for support in several tasks of the current workflow of a journalist. In the following sections you can find insights of the survey conducted in February 2022, which helped us to validate possible AI driven solutions, as part of the Smart News Assistant.

Research phase - Monitoring dashboard

Information and news are everywhere on the internet. The amount of content to work with and to inform media consumers about is increasing. Journalists are getting their stories from a lot of different sources: news agencies (e.g. Reuters), internal communication channels (like iNews, Slack), other news brands, social media (e.g. Facebook, Instagram, Twitter) or via mailbox. A total of 28 (from 49) survey respondents indicated that they use at least 4 different sources.

Most respondents (69%) indicated that they would be "interested" or "very interested" in an interface or dashboard that presents a clear, organized overview of news items originating from any source (e.g. social media, news agencies, ...) (see Figure 2, nr 1). So, they can

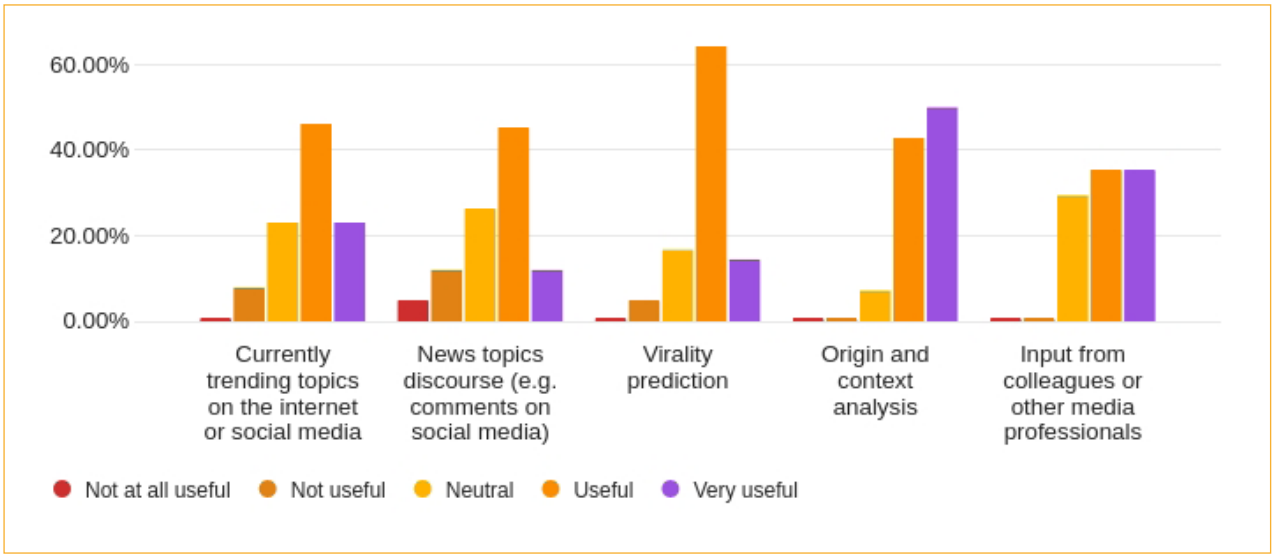

Figure 3 . Usefulness of monitoring dashboard functionalities according to survey respondents

decide which stories to research and focus on.

We presented different functionalities of such a dashboard in our survey (see Figure 3), and most people (92,9%) indicated that origin and context analysis of information is a useful functionality. The virality of a topic or story was also indicated as useful (78,6%). Lastly, comments on social media might be an interesting source of information for journalists, according to most respondents (71,4%).

Crawling social media takes a lot of time, but it is an interesting source via which you can get an indication of how relevant your story might be for your audience. Solutions that monitor social media and give an indication or notification to a journalist when something possibly interesting comes along are useful to decide which stories to focus on.

It's important to combine all sources in the dashboard, because currently journalists already use a lot of different tools.

Research phase - Fact checking

In the research phase, fact checking is relevant to avoid misinformation. Journalists are looking for ways to check information or verify claims that are spread on social media, e.g. images of the war in Ukraine that get manipulated or are used in a different context.

In our survey, a journalist explicitly mentioned she could use support in checking the authenticity of images: 'My colleague (mentioned specific name) is a great help and an expert in his field, but logically he is not always available.'

A lot of tools are already available but only experts are aware of these tools and understand how they work. A total of 80% of survey respondents indicated that they are "interested" or "very interested" in a comprehensive toolbox of fact checking tools to help check claims, images and videos (see figure below).

We asked respondents what this toolbox should consist of. They specified tools that could help in finding the origin of an image or video, checking images for manipulation, and giving an indication of the trustworthiness of a claim as most relevant.

When false information is spreading, a degree of virality in a monitoring dashboard (Figure 2, nr 1) might be an indicator to decide whether to create a fact-check on the topic or not.

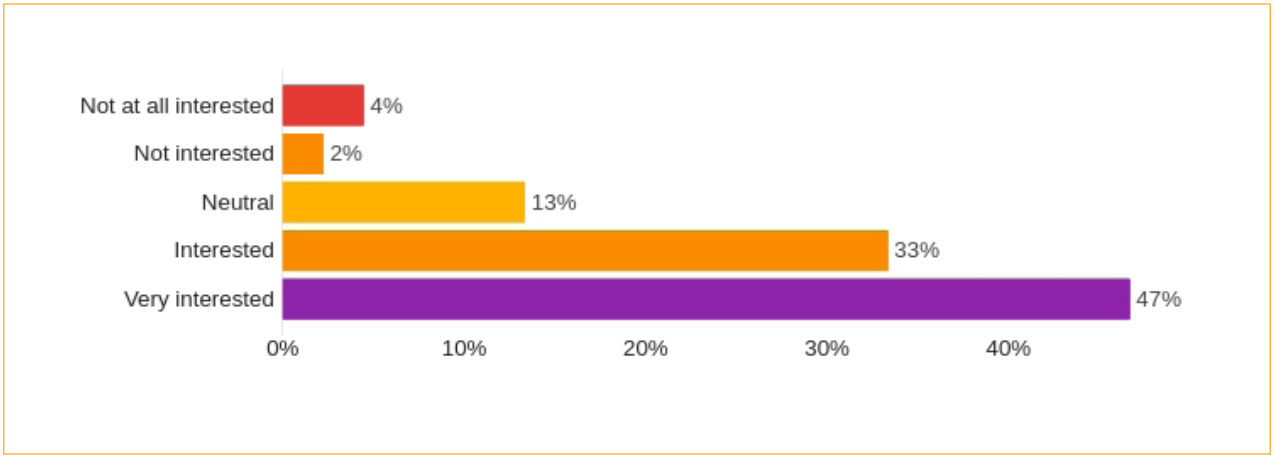

Figure 4 . Level of interest in fact-checking toolbox

Publication phase - Smart editor

Our survey showed that while creating a story, there are several opportunities to support news professionals in creating their story faster and making it more visually appealing. Respondents also indicated that tools helping them to avoid errors and report objectively would be a big win.

Multimedia

Currently, journalists use different ways to create their story, e.g. work in a Google Doc to combine research or work directly in the Content Management System of their app/website. While writing their story, they consume a lot of different sources to enrich the story (e.g. databases for stock images, biographies, previously written articles around the same topic, etc.).

Survey respondents indicate that they lose a lot of time trying to find the right assets for their stories.

"I lose a lot of time in finding the right images for my story."

"A big database that combines the best images that are already bought and/or adjusted by another journalist of VRT would be interesting."

A little more than half of the respondents (56%) indicated they are interested or very interested in a Smart Editor to receive automated suggestions while writing their story. These suggestions could either be media assets (like stock images), but also evergreen content, like a biography in the form of "an ID-kit of a country or a person" or short explainers like "I would like suggestions that give an answer to: what was this event about again?"

We also received some critical feedback, which is important to keep in mind. "Pure factual information can be helpful. Copy-pasting cannot be the intention. So preferably avoid ready-made content."

Diversity

Trust in news is falling. A recent Reuters news report polling news consumers across 40 countries found that 38 percent trust most news most of the time down from 42 percent in 2019.

There is a particular distrust in institutions that are considered elite institutions. Often the media does not look like the people it's covering. In Web Summit 2022 Marketing Trends Report Allesandra Galloni, editor-in-chief at Reuters, mentions it's really imperative for newsrooms to become like the world that they cover; to become more diverse in not only who they are, but in what they cover, and to expand their horizons in the sorts of stories they cover and how they cover them.

To consistently create journalism and media content that fairly represents our world, BBC initiated for example the 50:50 Equality Project, an initiative from the BBC committed to inspiring organizations across the globe. Solutions such as a diversity score could help meet those goals.

Journalists indicated that receiving a diversity score on gender, age and political stance in a smart editor before publishing stories was less interesting (47% indicated they are interested or very interested) compared to other solutions we presented. Nevertheless, it could still be relevant to gain more trust in news.

Production – story repackaging

Currently, news departments put a lot of effort into creating different formats for a story, e.g. an article on the website, an interview on radio and a video for social media. "I lose a lot of time in creating different formats of my story", one respondent pointed out. Some respondents indicated they could use some help with creating stories for Instagram, e.g. "I could use help in creating other story formats".

We asked respondents whether they would be interested in semi-automated repackaging their story source material in different derived formats so they can offer user-adapted or personalized versions.

Overall, they were interested (59% indicated they were 'interested' or 'very interested'). This might include a variant for Instagram Stories, a written summary, a bullet point summary or a video. In Figure 6 you can find how many people find which derived format interesting.

Semi-automation could be interesting and useful. But we need to keep in mind the user still wants control over what exactly will be published.

"You should be careful with solutions like this, the end responsibility should always be with the journalist, final editor!", one respondent suggests.

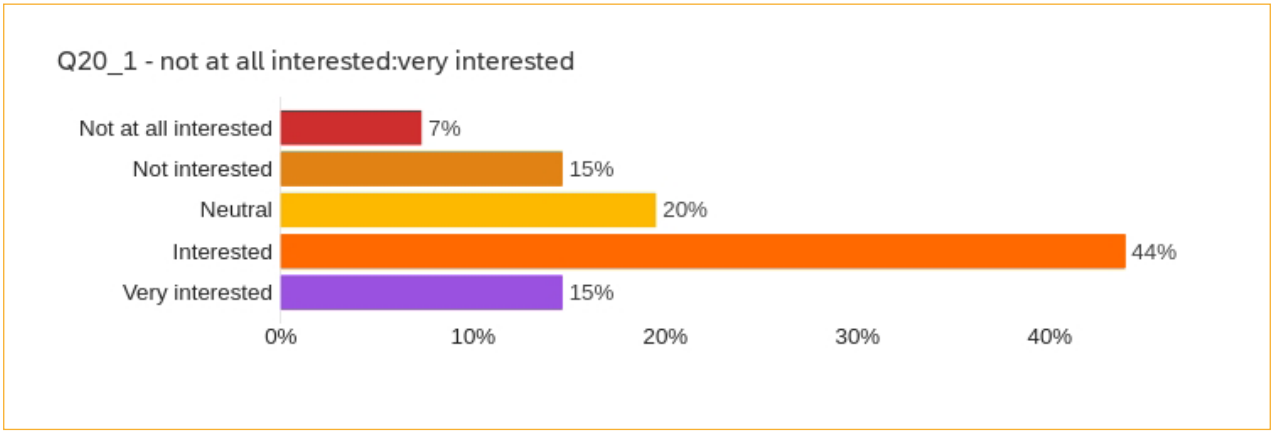

Figure 5. Level of interest in semi-automated repackaging story

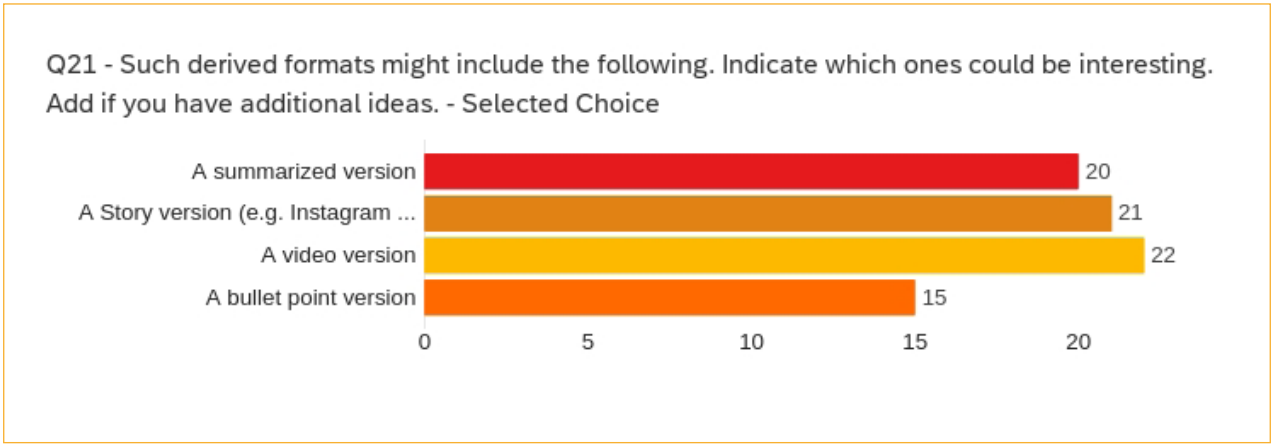

Figure 6. Which derived formats are interesting

Trustworthy AI

In the survey, we also covered a section on Trustworthy AI, a term that describes AI functions and systems that implement one or more of the features listed below - as applicable. Similar concepts are Ethical AI or Responsible AI. The aim is to adhere to ethical, legal and safety principles or non-technical methods, like:

- **preserving people's privacy (related to the use of personal data),**
- **complying with applicable regulations and legislation (e.g. GDPR or AI Act),**
- **being transparent and accountable by providing information on data sets, algorithms and models in use by AI tools,**
- **explaining how AI model predictions/outcomes were reached to the extent it is possible,**
- **listing measures taken to mitigate AI bias (increasing fairness), and**
- **detailing the level of technical robustness of used tools (safeguards against manipulation or attacks).**

Trustworthy AI was not a familiar concept for respondents: 67% never heard of Trustworthy AI before and only 22% indicated they have already come across Trustworthy AI features. Such features could include general transparency about AI techniques used, specific fact sheets on data sets or models, explanations about how predictions were reached, information on bias mitigation strategies or technical robustness as well as information related to legal compliance. It could also involve algorithmic solutions employed to increase/ensure fairness, explainability or robustness.

Respondents indicated the most needed Trustworthy AI feature is Privacy Protection (89%). As shown in figure 7, generally, there is a significant level of importance associated to all these Trustworthy AI aspects.

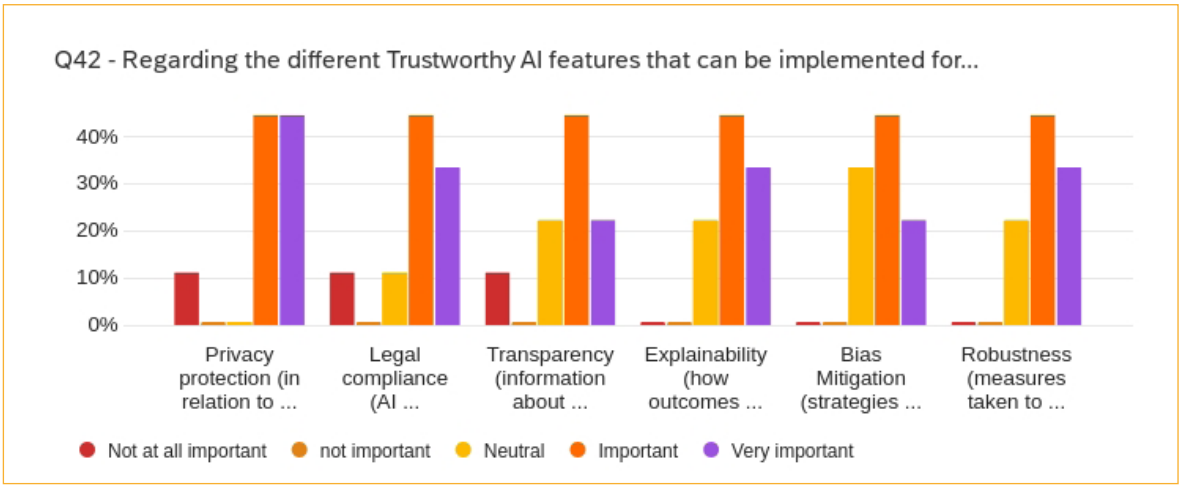

Figure 7. Level of importance Trustworthy AI features

5

Conclusions

The survey results indicate that journalists are mostly interested in using the fact-checking toolbox during their day-to-day workflow. This is the most popular solution followed by a dashboard for monitoring news from different sources in real-time.

The high level of interest in the fact-checking toolbox seems to originate from the fact that journalists are increasingly confronted with disinformation. Firstly, they have to verify the sources they use when writing a story. It is increasingly hard to check whether (online) sources can be trusted because a lot of them are anonymous or can't be verified directly. Secondly, they are confronted with the fact that more and more disinformation is spreading online. As journalists for a public broadcaster, it is also their task to perform a fact-check when a false story is getting a lot of public attention, and in some cases, even publish a counter-story with facts and/or counter arguments. The fact is that tools to help with this fact-checking are much needed and exist. Unfortunately, they are often too complex and are currently only used by experts. This, however, indicates a clear opportunity for AI researchers and technology experts who are building these tools. To make their tools more useful for all journalists in a newsroom, they should be simple, intuitive and (if possible) tailored to the workflow of a journalist.

The least interest seems to be about having diversity assessment of a news story. Although this is very important (especially for VRT as a public broadcaster), this makes some sense from a journalist's standpoint. Specifically, it could possibly cause even more work for the journalists themselves, e.g. when a re-edit would be required to include every point of view. We can conclude with two take-aways. Firstly, it might be more worthwhile to do an overall diversity assessment on a range of news stories instead of doing an individual assessment on every single story. Some stories might not include every point of view explicitly, but the overall news reporting (e.g. over a certain timerange) should. Secondly, if we do introduce such an assessment, it should not introduce extra work for the journalist, but rather give him/her insights/tips at the right time during the story creation process, e.g. in the research phase where collecting and verifying resources is most important.

The analysis of survey results shows that there is much opportunity for future research, AI solution development in the areas of news research, production and publication. Our survey provided a clear message to AI technology providers that simple, understandable and trustworthy AI-based solutions are highly welcome and needed to enhance, facilitate and modernise day-to-day journalistic workflows. An AI-driven News Assistant incorporating the components listed above can become an indispensable Swiss Army Knife for journalists and newsrooms, facilitating mundane and burdensome tasks and giving more space to creativity and original investigative and informative work. During the creation or adaptation of such tools, we would advocate for involving journalists as early in the research and development phase as possible, in order to produce useful tools which fit the current workflows and optimize them even further.

AI4media

Follow us @ai4mediaproject

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951911

info@ai4media.eu www.ai4media.eu