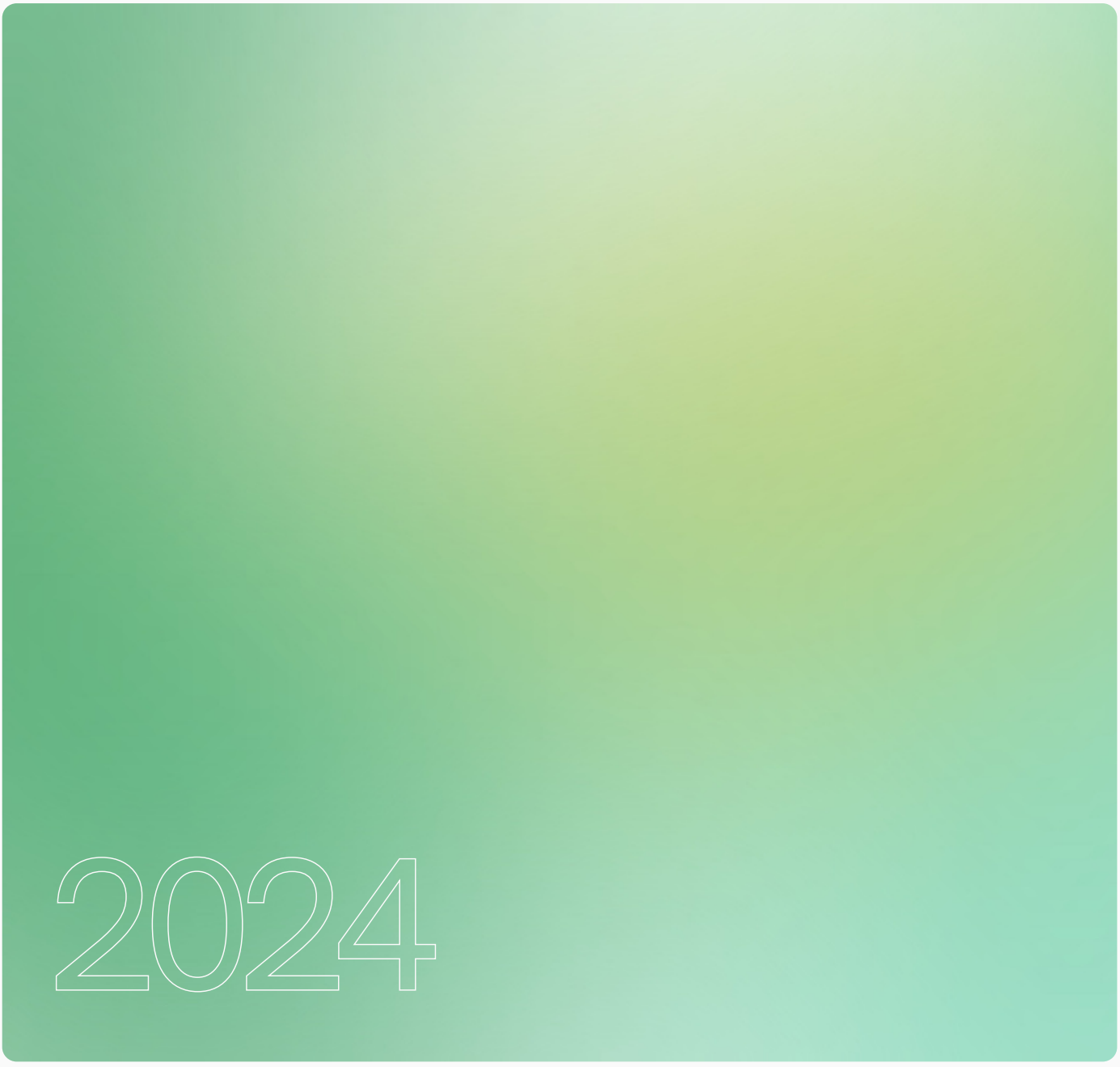
NORDIC AI JOURNALISM IN COLLABORATION WITH **UTGIVARNA**

:Utgivarna

AI transparency in journalism



Project team

Agnes Stenbom (project lead), Schibsted
Annie Lidesjö, NTM
Calle Sandstedt, Omni
Camilla Hillerö, NTM
Charlotta Friborg, SVT
Emil Hellerud, TV4
James Savage, The Local
Johan Möller, UR
Johan Silfversten Bergman, Svenska Dagbladet
Karolina Westerdahl, Expressen
Louise Sköld Lindell, Göteborgs Posten
Maria Kustvik, NTM
Martin Ekelund, TV4
Martin Jönsson, Bonnier News
Martin Schori, Aftonbladet
Mattias Pehrsson, Bonnier News
Mikaela Åstrand, SVT
Olle Zachrison, Sveriges Radio
Mattia Peretti (process design)

With support from the Board of Utgivarna

Members

Anne Lagercrantz (chair), SVT
Thomas Mattsson (vice chair), TU
Kerstin Neld (vice chair), Sveriges Tidskrifter
Viveka Hansson (vice chair), TV4
Hanna Stjärne, SVT
Cilla Benkö, SR
Christofer Ahlqvist, TU
Johan Taubert, TU
Jessica Wennberg, TU
James Savage, Sveriges Tidskrifter
Kalle Sandhammar, UR
Mathias Berg, TV4

Substitutes

Anna Careborg, SVT
Mimmi Karlsson-Bernfalk, TU
Anders Enström, TU
Åsa Junkka, TU
Cissi Elwin, Sveriges Tidskrifter
Unn Edberg, Sveriges Tidskrifter
Michael Österlund, SVT
Sofia Wadensjö Karén, SR
Margaretha Eriksson, UR
Fredrik Arefalk, TV4
Johanna Thulin Bratten, TV4
Åsa Rydgren, Utgivarna (co-opted)

Contents

This report summarises a unique project where executives from 13 Swedish media companies gathered for shared debate and learnings on how we should share information about our use of AI with the media audience.

Background

Advances in Artificial Intelligence (AI) have highlighted both potential and challenges for journalism. Among many other things, the development raises questions about transparency, which the Swedish media industry decided to explore together in this project.

EU regulations are currently in progress to ensure that AI systems are used responsibly and safely. The objective is often framed as a delicate balance between fostering innovation and growth while safeguarding the rights and well-being of EU citizens. By implementing regulations for AI usage, the goal is to address risks associated with discrimination, privacy infringements, and lack of transparency.

Transparency is a recurring theme in the EU's AI Act. While we defer interpretations of the legal text (which is still in development as of the authoring of this report) to other groups/forums, as practitioners, we can conclude that the proposed language on transparency for AI-generated content in editorial media offers an opportunity for the industry itself to advocate for a sustainable path forward. We can champion openness and accountability regarding AI usage within and beyond the Swedish media landscape.

In the autumn of 2023, Agnes Stenbom (Schibsted) and Olle Zachrison (Swedish Radio) took the initiative for an industry-wide learning process through the Nordic AI Journalism network. Under Utgivarna’s (the Swedish Publishers' Organisation) umbrella, a project was launched with the objective of providing the media audience with improved conditions to understand when, how, and why AI is deployed in Swedish editorial media.

About the project

During the winter of 2023/2024, senior representatives from 13 Swedish media companies gathered for a workshop series. Our joint learning process is summarised through recommendations on how AI transparency should take form in our industry.

Scope

Transparency is, just like AI, broad concept. The focus of this project has been transparency towards the media audience, with a specific focus on how Swedish media companies inform our users about the use of AI in our editorial media products.

Another relevant form of AI transparency concerns technical explainability, where questions about, for example, data bases and model weights are central. We look favorably on transparency requirements for various forms of AI models, but have not touched upon these issues in the project.

Questions explored

It is easy to support transparency as an ideal, but all the more difficult to concretise its nuances and relevance in practice. Through this project, we have first delineated the challenge (what/when), and then explored possible solutions (how/where). The following questions have formed the basis of our joint knowledge development:

About what and when should we be transparent about AI?
Is there a difference regarding the need for transparency about AI use for content creation and AI being used in other ways as a tool in the newsrooms?

How and where is transparency needed around AI usage?
Does it matter how and where information about AI is provided?

Participants & timeline

All participants received a personal invitation to the project. Some organizations were represented by multiple people due to scheduling conflicts and/or to include different perspectives from their own organization. The group met thorugh four digital workshops between December 2023 and February 2024. Utgivarnas' AI group was the steering group for the project, and their legal group provided feedback on draft recommendations. The recommendations were presented to Utgivarnas' board on 26/2 2024.

Basic assumptions

Four basic assumptions set the scene for our recommendations on AI transparency in Swedish editorial media. Our fundamental position is that the audience needs to be able to trust that content published by such media is always journalistically accounted for – regardless the tools involved in the production process.

01 **Transparency is a means of creating trust**

We believe journalism can earn the public's trust through transparency about our processes, including those involving AI. Transparency regarding editorial media's working methods and use of technology is thus not an end in itself, but a means to achieve credibility and trust.

02 **There are risks associated with excessive AI transparency**

Although transparency is often mentioned as something universally good, we believe that there are good reasons to discuss possible downsides and risks. With unreasonable conditions linked to our transparency around AI, we risk possible misunderstandings around journalistic principles and reduced trust in professional journalism.

03 **Editorial media needs to have our own strategy**

It is important to distinguish between the use of AI in editorial media compared to other types of digital platforms, not least such where user-generated content is shared. The editorial media's AI transparency should be based on our unique offer, where we can clarify that AI is one of many tools we use to deliver editorially quality-assured products.

04 **Clarity is needed about appropriate levels of transparency**

Inadequate guidelines around transparency about AI use can lead to arbitrary judgment, especially in large comapnies with many departments. Maintaining and earning the trust of media consumers requires a conscious strategy around AI transparency. To safeguard independence, it is important that the media takes responsibility for crafting such a transparency model.

Recommendations

We stand behind and share 7 recommendations on AI transparency in Swedish media. The recommendations are voluntary to follow, and should be adapted to the local editorial conditions. They mainly apply to news and magazine journalism.

FUNDAMENDAL

Our **fundamental recommendations** (*principles*) concerns overall values and guidelines.

PRACTICAL

Our **practical recommendations** are more concrete and suggest specific measures to fulfill the principles in practice.

FUNDAMENTAL

1.

AI with ‘significant journalistic impact’ requires transparency

Just as the Swedish press ethics provide a basis for journalistic work, there needs to be guidelines for when the use of AI requires transparency. We believe that *"AI with a significant journalistic impact must be handled with openness and clarity towards the audience*" is an appropriate such fundamental principle.

The principle is guiding rather than strictly defined. It is important that the definition of what ‘significant journalistic’ means (and thus, the decision about transparency) rests with the publisher, to ensure that it is adapted to the specific needs and circumstances of each publication. We do not find it suitable or desirable to define a fixed limit or type of AI-impact, but instead argue that the responsibility for evaluation should be placed in the editorial process.

The principle applies to different media and content types, thus covering both generative AI and other forms of AI tools and applications (e.g. when AI is used to conduct investigative journalism).

2.

Other internal AI-tooling does not require transparency

Provided that the above principle is adhered to, we argue that public information is not needed for internal applications where AI is used to support editorial or commercial work. For such use cases, internal AI policies apply.

Examples where AI can be used without actively sharing information with the audience:

- **Transcribing for internal use** does not affect the content or credibility of the published material. The focus should be on the finished result of the transcription rather than on the method used to achieve it.
- **Proof-reading software** is a common feature that does not affect the substance or authenticity of the journalism. The goal is to deliver correctly spelled texts to the audience, not which tools are used to achieve it.
- **SEO** work is part of the editorial and commercial process to make content more visible and accessible to a wider audience. It is the result of SEO strategies that are of interest to the audience, rather than the technical methods used to implement them.

An AI implementation whose impact on journalism we did not reach the same conclusion on in the different media companies is content recommendations. Some argued that such do not require any active information effort towards the audience, as they (in their standard form) only recommend content from the editorial database and thus do not change the journalism. Others argued that they have a significant impact on the journalistic *product*, and thus require transparency. The example strengthens our recommendation (1) that each newsroom should make its own journalistic assessment of what requires transparency.

FUNDAMENTAL

3.

AI transparency must be approached as an iterative theme

Many of the AI technologies now used in the media industry are at an early stage, during which it is necessary to be open and clear about their usage. This openness will need to be reassessed over time as technology and user habits and expectations evolve.

We are positive towards continued industry dialogue, and also see a great need for media companies to have an ongoing internal dialogue on the topic. We intend to follow up on our recommendations in six months, to evaluate their continued relevance.

PRACTICAL

4.

Be specific about the type of AI tool applied

To counter audience perceptions of AI being an autonomous force operating without editorial control, we encourage media companies to be specific when telling our audiences what type of AI we used. This can serve an educational function for both the media companies and the media audience, where we demystify the concept of AI and instead communicate about the technology as a tool.

We encourage the use of terms like “image generation” or “text analysis” to more specifically describe the task which AI technologies have been part of fulfulling.

Illustrative example

A editorial team has used AI to create a journalistic comic strip. The editorial team believes that this has a significant journalistic impact and should hence inform users that AI has been used. Therefore, in a caption, it states that "the comic strip is created with the support of an AI tool for image generation".

Some media companies may not want or are unable to unduly favor AI companies commercially by explicitly mentioning specific company names. However, it is important to note that the terms of use and naming of model names differ between different AI companies and models, and sometimes it may not be possible to avoid mentioning model/company names. Therefore, where journalistically relevant and copyright and/or contractually required, we believe that an explicit naming of the models (commercial or open source) used is appropriate. In addition to providing transparency to media consumers about our processes/tools, this can also make it clear to our own organizations whether we are utilizing a suitable range of AI tools/companies.

For most media aimed at the general public, we believe that technical details (such as model versions and model weights) are redundant and may even be counterproductive by making the consumer feel they do not understand the technical jargon. However, we see a relevant exception in method descriptions of editorial work, especially investigative journalism, where AI has a major journalsitic impact.

Illustrative example

An editorial team has utilized a language model for investigative journalism and concluded that AI usage has a significant journalistic impact by enabling data analysis that would not be possible with the editorial team's human capabilities.

As the model in question is a crucial component in the publication, the team shares information about the specific model they have used, and explains how method choices such as prompts and temperature may have influenced the investigation.

PRACTICAL

5.

Share information in connection with consumed content

Where AI transparency is needed, information should be shared in connection with the AI implementation. By sharing information alongside the consumed journalistic content, we provide media consumers with a chance to understand for themselves the extent to which AI is used with significant journalistic impact. The recommendation applies to both AI as a tool (see the above example regarding data analysis in investigative work) and for generative AI.

The recommendation is complicated by the fact that there isn't always a given definition of what "alongside" means, especially for audio media. Here, we argue that the fundamental recommendation (1) of "significant journalistic impact" is central to the assessment work, and particularly the editorial teams' responsibility to continuously evaluate the need for audience information. How information reaches the audience can and should vary between publications.

Illustrative example

An editorial team uses AI to generate an article text and believes that this has a significant journalistic impact as AI has made language and information selection. Therefore, the editorial team should inform users that AI has been used.

Even though it is stated in the editorial team's AI policy that they sometimes use AI to generate article texts, users must receive information about this specific usage alongside the relevant content.

PRACTICAL

6.

Harmonise the industry’s language around generative AI

We recommend a harmonised approach for describing AI within and between Swedish media companies. This is mainly motivated by media consumers more readily having an opportunity to understand and relate to AI in Swedish media if we communicate similarily about its usage.

As a first step, we recommend a harmonized language regarding generative AI. As a basis, we suggest the wording "created with the support of", to signal AI's actual impact on the content and remind the media consumer that there is an editorial process (and staff) behind the content. Alternative wordings that have been discussed but dismissed include "assisted by," "together with," "via," or "by." We believe these terms set incorrect expectations and may diminish the perceived role of humans in the editorial process.

We envision the following foundation for more harmonised formulations: **"[the content in definite form] is created with the support of an AI tool for [the task] ([the model/tool], [the version])."**

Illustrative example

An editorial team publishes an AI-generated illustration and considers it to have a significant journalistic impact. They inform their audience about it using one of the following formulations:

- The illustration is created with the support of an AI tool.
- The illustration is created with the support of an AI tool for image generation.
- The illustration is created with the support of an AI tool for image generation (DALL-E).
- The illustration is created with the support of an AI tool for image generation (DALL-E, v.2).

Depending on the specific media company's target audience or tone, a combination/variant of these levels may be appropriate. Just like in several of our other recommendations, this becomes an editorial decision.

Illustrative example

An editorial team publishes an AI-generated illustration and considers it to have a significant journalistic impact. They inform as follows:

- **The illustration is created with the support of the AI tool DALL-E.**

We recommend that individual media companies make their own assessments of any additional information that wish to add. We expect that many media companies will, perhaps on a case by case basis, choose to add a statement about editorial quality control.

Illustrative example

An editorial team publishes an AI-generated article and considers it to have a significant journalistic impact. They inform audiences using the following wording:

- **The article is created with the support of an AI tool for text generation and has been reviewed by our editorial team before publication.**

If Swedish media were to automate their content generation process (i.e., no longer have employees writing prompts, reviewing generated content, and manually publishing), the above formulations would naturally be reconsidered, as the content would no longer be created with the support of AI, but rather, created by AI.

We also believe that there is good reason to consider whether the term "AI" should remain in the formulation for long, but assess that it is relevant at present. Therefore, like everything else in these recommendations, the formulation needs to be continuously evaluated and updated where/when relevant.

PRACTICAL

7.

Avoid visual labels (icons) for AI in editorial media

We advise against standardised visual labeling (e.g., an icon) for AI in editorial media. We are motivated by three main arguments:

- An explicit visual differentiation may give the impression that the credibility of AI-generated/affected content overall differs from content created by human journalists or editors. The foundation must always be that if the content were not credible, it would not be published by our editorial media.

- Visual labeling may be challenging to implement consistently across different media platforms and content types. There can be many different ways to use AI in media production, and attempting to create a general visual label covering all these use cases can be complicated for media companies and difficult for our users to understand.

- It risks becoming a cumbersome process given where we are in AI development. New tools and use cases will emerge along the way, and a visual standard risks quickly becoming irrelevant.

This recommendation follows a discussion about whether it is appropriate for editorial media to use photorealistic AI-generated images, especially in news feeds. Recommendations regarding this are beyond the scope of this project , but we still want to emphasize that the recommendation would need to be reassessed for this specific case if any of the Swedish media were to start publishing photorealistic AI illustrations.

Summary of recommendations

FUNDAMENTAL RECOMMENDATIONS

1. AI with 'significant journalistic impact' requires transparency
2. Other internal AI-tooling does not require transparency
3. AI transparency must be approached as an iterative theme

PRACTICAL RECOMMENDATIONS

4. Be specific about the type of AI tool applied
5. Share information in connection with consumed content
6. Harmonise the industry's language around generative AI
7. Avoid visual labels (icons) for AI in editorial media

Final comments

The more we delve into the theme of user information about AI, the more questions arise.

Is it equally important to be transparent about the use of AI in entertainment as it is in news reporting? Should messages about AI be directed differently to build trust with different audiences, and can this be done automatically? Is it appropriate for news media to publish AI-generated images that could potentially be mistaken for documentary images?

The industry needs to continue exploring these themes to ensure that journalism remains credible and reliable in an era of rapid technological development, changing market conditions, and new user behaviors.

Last but not least, we believe that it is increasingly important to make visible and explain the human editorial processes. By making our editorial processes visible while also informing our audience more clearly about AI as a tool (see recommendations 4 and 7), we can give users a practical understanding of AI's role in the Swedish media landscape and clearly emphasize the norms and ethical principles of professional journalism. We can show our users that we take responsibility for the published content - regardless of the technical tools we are using right now.

With awareness of the potential pitfalls, we believe that the media industry has a very bright (and where relevant, transparent) future with AI tools!

www.nordicaijournalism.com/ai-transparency