

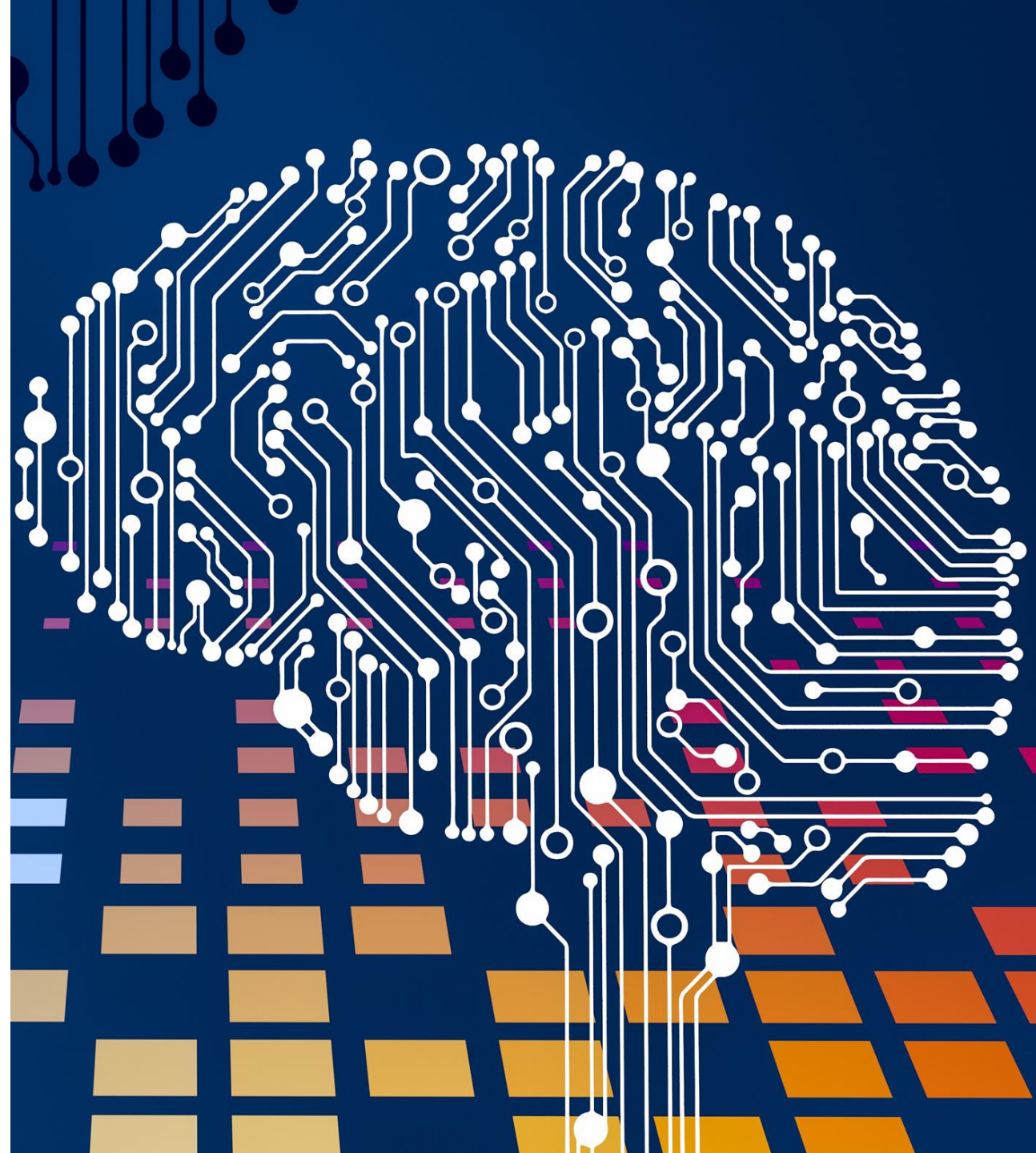


# The **ethical** challenges for using (generative) **AI in journalism**

Laurence Dierickx

**AIPCE Annual conference 2023**  
**December 15, 2023**

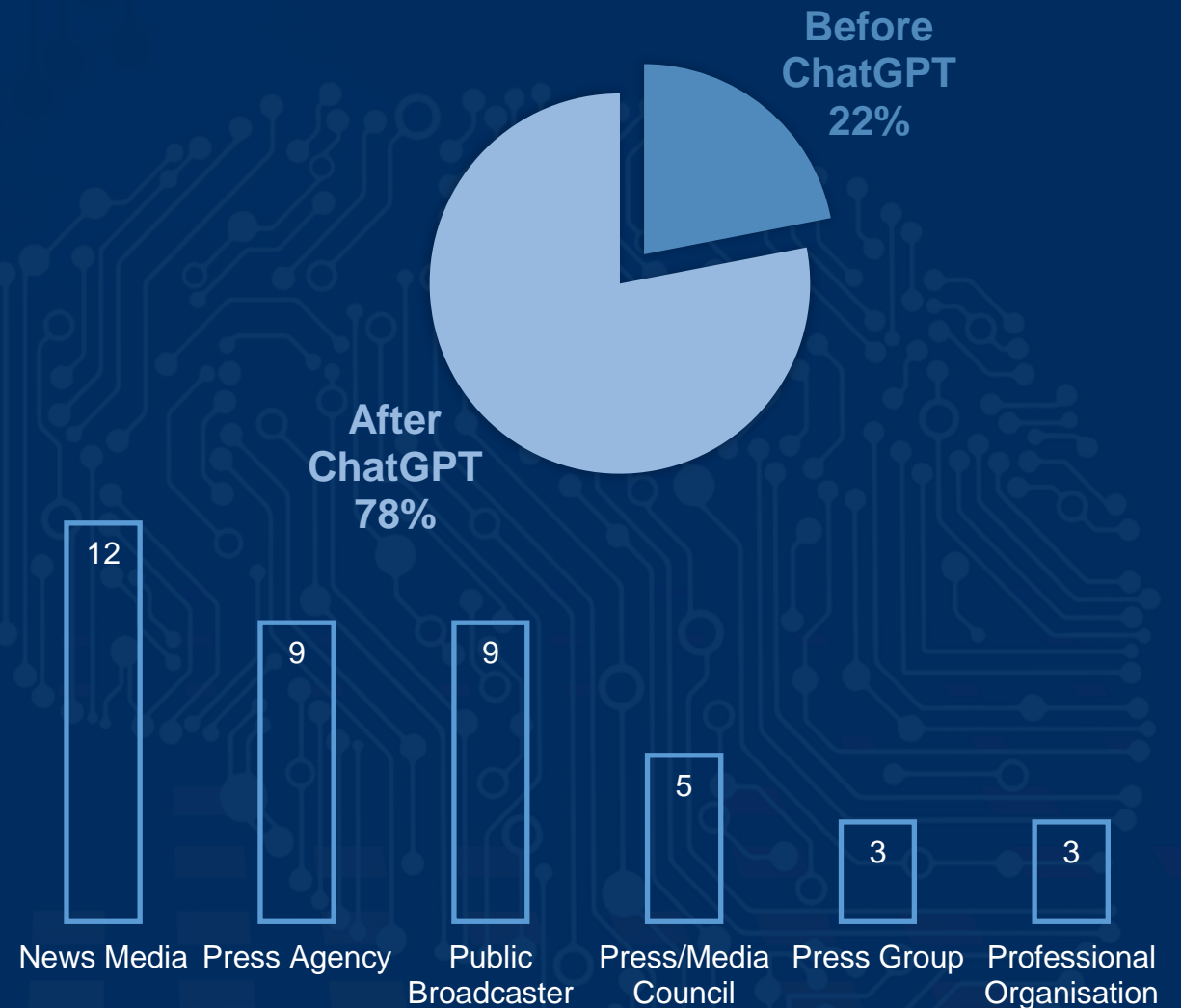
Université Libre de Bruxelles, Belgium  
Universitetet i Bergen, Norway



# How do European news organisations and self-regulation bodies frame ethical practices?

N = 41

- 15 guidelines
- 14 principles
- 7 positions
- 3 recommendations
- 2 ethical codes
- 12 countries (Western/Northern Europe)



*Dierickx, L. & Lindén C.G. (in review) Dealing with Biases and Hallucinations: The Ethical Uses of (Gen)AI Tools in the EU News Media Sector, University of Bergen/Nordic Observatory for Digital Media and Information Disorders, NORDIS*



# The era before ChatGPT

## Focus on news production

Transparency and accountability: “This paper was written by a machine” (clear identification)

## Focus on news recommendation and personalisation

Transparency, privacy and accountability: informed consent of the user, avoiding creating “echo chambers” (disputed concept in research)



## Acknowledging the importance of data in the processes

Data is the material condition for news automation (press agencies)

Importance of the principle of fairness (in data and computer science, relates to unbiased and balanced data)

Data quality important to avoid biases, developing data culture (BR, Germany)

Avoiding adverse effects of echo chambers and maintain diversity: data quality (BR), partial editorial curation (SRF, Switzerland)

## Acknowledging the importance of non-journalists in the processes

Human-centric approach (press agencies)

### **Guidelines for machine learning developers (BBC)**

The responsibility for AI lies with those who build the system

Consequentialist ethics (biases, echo chambers)

### **Bayerischer Rundfunk (BR)**

Responsible engineering

Developing interdisciplinary perspective



# Press councils' recommendations

*Alignment on public broadcasters and press agencies*

## **Finnish press council**

Transparency = credibility

Disclosing the non-human nature of the authored text

Providing information on how the users data are collected, for which purpose

## **Catalan press council**

Data quality

Data management and privacy

Process monitoring

Enhancement of the human factor

Proactive research and development

News detection

News discoveries

Social media  
analysis

Audio-visual  
search engines

News verification

Automated fact-  
checking

Data and text  
analysis

Text  
summarisation

Audio  
transcription

Machine  
translation

Text-to-speech  
applications

Data-to-text  
generation

Multimedia  
creation

News  
recommenders

Etc.



# The era after ChatGPT

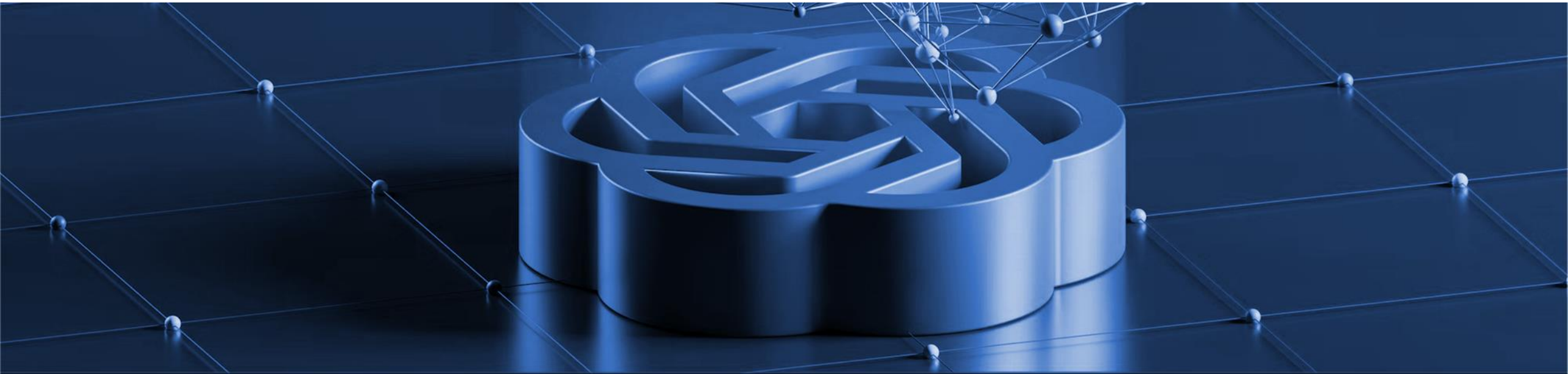
ChatGPT launched in November 2022

Widespread use of (Gen)AI in journalism

Highlighted that AI was not new in journalism

Made AI accessible (no programming required)

A field for experimentations



# Defining generative AI

A broad label describing any artificial intelligence that uses unsupervised learning algorithms to create new digital text, images, video, audio, or code. (The Guardian, UK)

Refers to technologies that can create new, unique content (such as text, images, audio, video, and code) based on the content used to train them. (SVT, Sweden/EBU)





## The common rules/statements

AI = part of the editorial process

Being open but experimenting carefully

### **Transparency**

Disclosure (generated content)

User agreement (personalisation)

### **Fairness**

Maintaining diversity

Requires trustworthy data

Watching out for biases in AI systems

### **Human-in-the-loop**

Human oversight and responsibility

Augmenting rather than supplementing

## The specific rules/statements

### **Privacy**

Preventing risks of leaking personal data and internal confidential information by not introducing them into the system (Ringier, Switzerland)

Journalists should not use their professional email addresses to create accounts in external services for security reasons (SVT, Sweden)

### **Environment**

We understand that the computing requirements of AI also burden the environment and we make choices to minimise acts. (Yle, Finland)

## **Telegraph journalists told use of ChatGPT will result in same sanctions as plagiarism**

Staff were warned against entering sensitive or proprietary data into ChatGPT.

Risks of contradictions linked to a proliferation of rules = confusion (trust)

Clear rules to avoid interventionism

Limitations: Practical ethical dilemmas

- De la même manière, nous pourrions comme pour les textes utiliser l'IAG comme un outil d'aide à la synthèse de vidéos, notamment pour faciliter la distribution de celles-ci sur différentes plateformes, et sous réserve d'une supervision humaine et éditoriale.

- All direct use of content created using generative AI models to the audience must be approved by the legally responsible publisher (RP) for the specific publication. The RP may consult the AI Council for advice. The content must always be reviewed by a human editor to ensure that it meets our journalistic requirements and editorial standards.

As a general rule, content generated by AI tools shall be labeled. Labeling is not required in cases where an AI tool is used only as an aid.

# Dealing with biases and hallucinations

- AI models can produce entirely false images and articles. They also replicate the existing societal perspectives, including historic biases' (*Financial Times*)
- They too often contain errors ("hallucinations") or biases ("bias"), and it is usually unknown what data the systems were trained with' (*De Volkskrant*)
- With computer-generated content, it is complex to guarantee the reliability of facts presented as true (ANP, The Netherlands)
- They also can generate 'false leads and boring ideas' (*Wired*)
- We should be aware that bias/prejudice may be inherent in the models, and editorial/legal considerations may be needed regarding objectivity and impartiality (SVT, Sweden)
- The sources used by AI are often obscure, making it problematic to use in editorial work (STT, Finland)

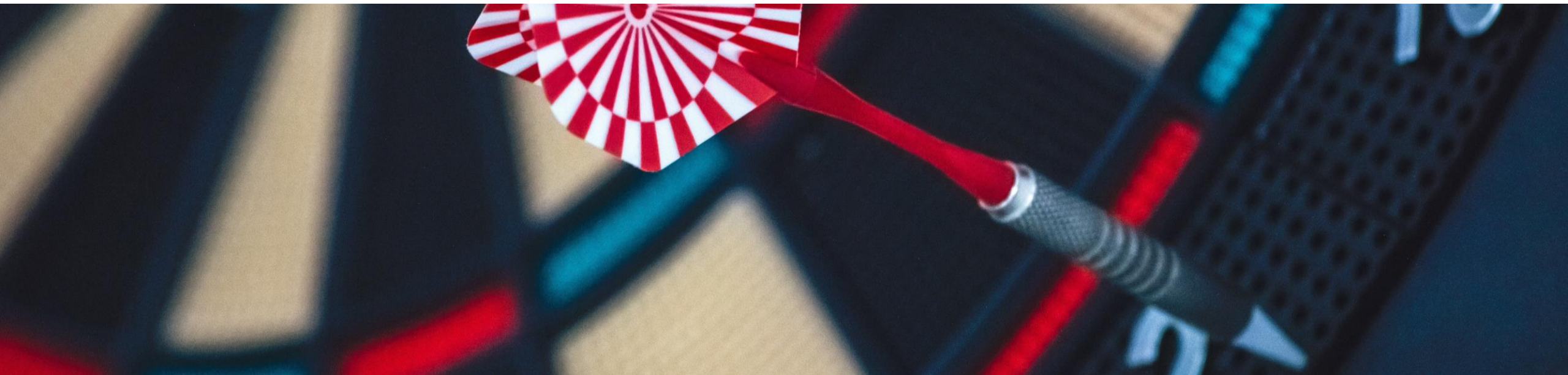




# Being honest, mitigating the risks

**Article 12.** The journalist is **transparent** about her/his approach and way of working. Clear disclosure news production / recommendation. Editors are responsible for editorial choices. Monitoring the application and implementation of ethical principles regarding the system developers (Raad voor de Journalistiek, Belgium-Flanders)

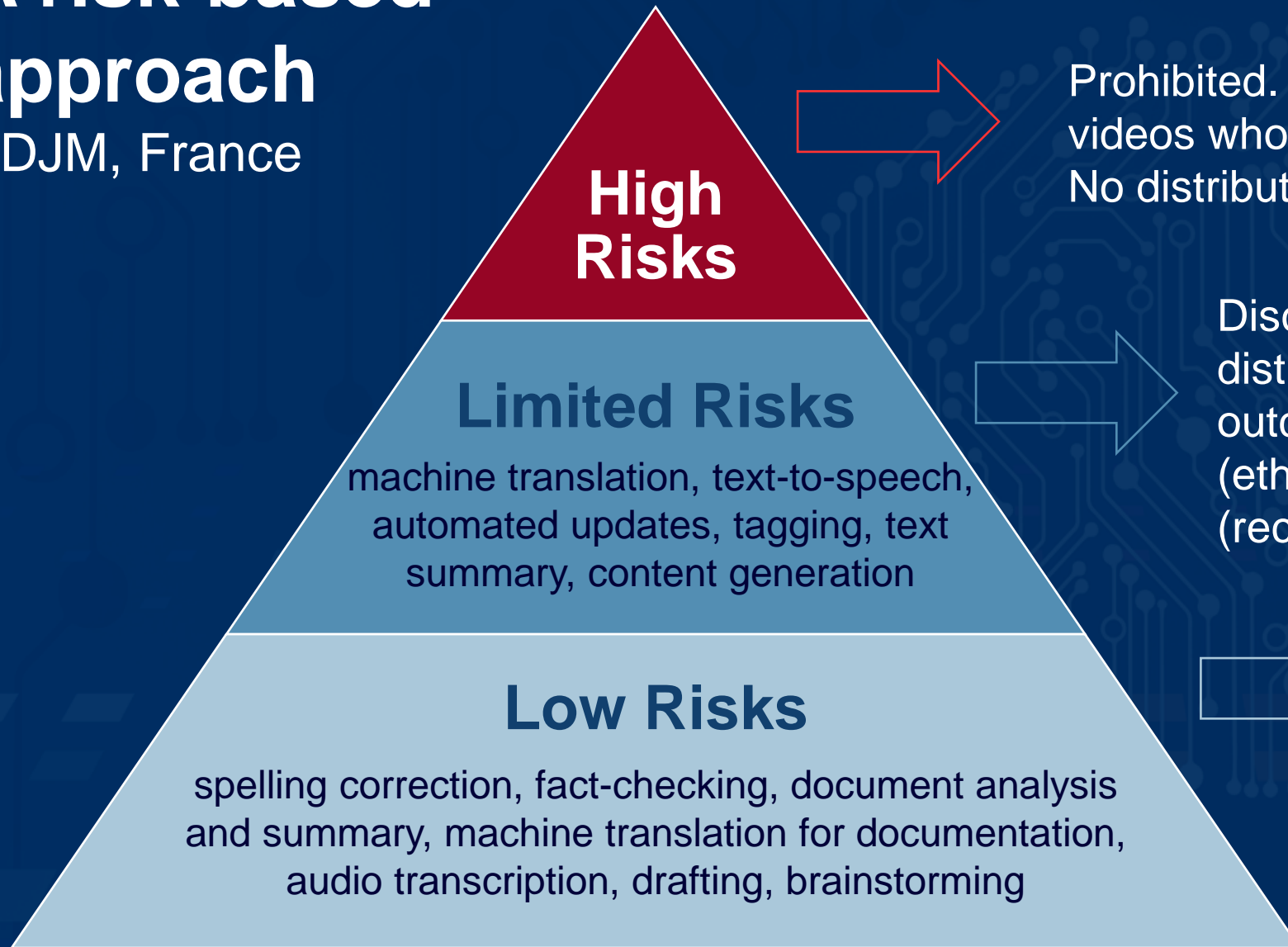
Ensuring **accuracy** of any AI generated content, AI is also used for creating and distributing false content, exercising human editorial oversight + **transparency** on the use of AI, including clear labelling on recommended content and disclosure. (Impress, UK)





# A risk-based approach

CDJM, France



**High Risks**

Prohibited. Realistic images, sounds or videos whose realism is likely to mislead. No distribution without human approval.

**Limited Risks**

machine translation, text-to-speech, automated updates, tagging, text summary, content generation

Disclosure for all type of content, distinguishing machine outcomes, human oversight (ethics), human editorial control (recommenders)

**Low Risks**

spelling correction, fact-checking, document analysis and summary, machine translation for documentation, audio transcription, drafting, brainstorming

No impact on information quality, no disclosure

# Connecting with ethical principles of journalism

- ✓ Accuracy (critical assessment and verification of sources and facts)
- ✓ Remaining cautious (sources might be unreliable, harmful or inaccurate )
- ✓ Respecting facts, no manipulation or distortion
- ✓ Fairness (avoiding biases and echo chambers, promoting diversity)
- ✓ Respect for (data) privacy
- ✓ Human responsibility and accountability

**Algorithms are parts of editorial processes.**

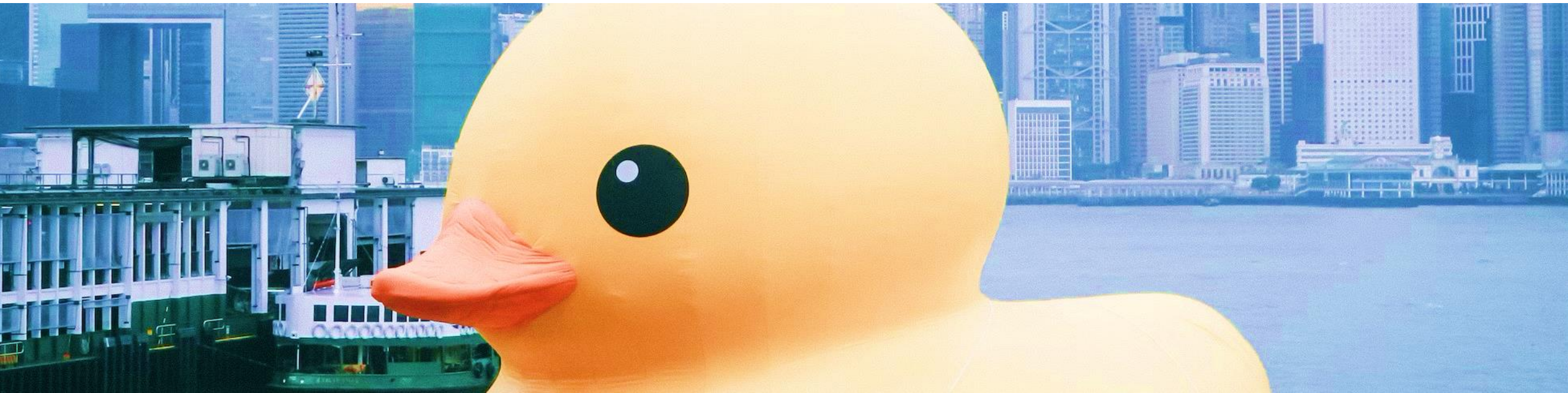
**Humans design them for human purposes and uses.**

# From flaws to pitfalls, remaining realistic

Disqualifying a technology because it does not meet journalism standards (bye bye, big tech)

Evaluating the quality of a given technology requires a high level of expertise, journalists often lack data and AI literacy

GenAI tools challenge the core ethical standards of journalism + the author's rights, the labour division and the overall relationship between the human and machine





# AI in/and journalism

Promoting interdisciplinarity and inclusive policies (computing and journalistic thinking, fine-tuning languages)

Strengthening the link with academia (research on echo chambers/selective exposure, uses of technology, limits of generative AI, automated fact-checking, etc.)

Developing robust data and AI literacy, better adapting initial training to evolving digital challenges, expertise is needed!





## AI ethics is interdisciplinary

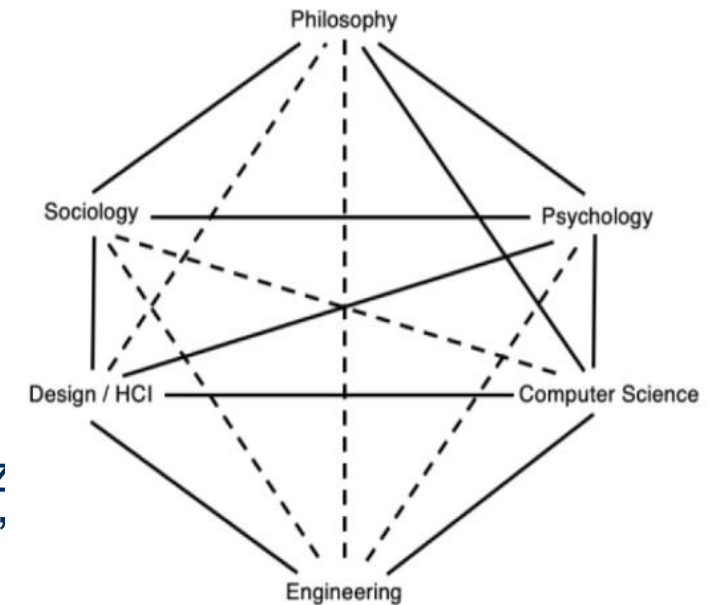
Could inspire ethics in journalism

(transparency and its limits, works on explainable AI, ...)

## Universal principles in data science ethics

“**Explain methods** for analysis and marketing to data disclosers.

**Maximizing transparency** at the point of data collection can minimize more significant **risks as data travels** through the data supply chain’



# ETHICS

# Thank you!

<http://ohmybox.info>

@ohmyshambles