



A Handbook
for Journalism Educators

Reporting on Artificial Intelligence

Published in 2023 by the United Nations Educational, Scientific and Cultural Organization 7, place de Fontenoy, 75352 Paris 07 SP, France

© UNESCO 2023

ISBN 978-92-3-100592-3

<https://doi.org/10.58338/HSMK8605>



This publication is available in Open Access under the Attribution-ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) license (<http://creativecommons.org/licenses/by-sa/3.0/igo/>). By using the content of this publication, the users accept to be bound by the terms of use of the UNESCO Open Access Repository (<https://www.unesco.org/en/open-access/cc-sa>).

Shutterstock images do not fall under the CC-BY-SA licence and may not be used or reproduced without the prior permission of the copyright holders. The designations employed and the presentation of material throughout this publication do not imply the expression of any opinion whatsoever on the part of UNESCO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

The ideas and opinions expressed in this publication are those of the authors; they are not necessarily those of UNESCO and do not commit the Organization.

Editor: Maarit Jaakkola

Contributing experts: Charlie Beckett, Edward Finn, Fredrik Heintz, Frederic Heymans, Suren Jayasuriya, Sayash Kapoor, Santosh Kumar Biswal, Arvind Narayanan, Agnes Stenbom and Jenny Wiik

Partial contributors: Jenny Bergenmar, Ammina Kothari, Bernhard Dotzler, Teemu Roos, Nicolas Kayser-Bril, Steve Woolgar.

UNESCO project management, review and editing:

Guy Berger, Director for Strategies and Policies in the field of Communication and Information, UNESCO

Vanessa Dreier, Junior Professional Officer, Communication and Information, UNESCO

Maksim Karliuk, Programme Specialist, UNESCO

External peer review:

Serkan Bulut, Associate Professor, Journalism Department, Çukurova University

Bruce Mutsvaivo, UNESCO Chair on Disinformation, Data and Democracy, Utrecht University

Manju Rose Mathews, Head of the Department, Media Studies, Christ Nagar College, Trivandrum

Kamilla Nigmatullina, Head of Digital Media Communications Department, Saint Petersburg University

UNESCO coordination:

Sophia Wellek, Media Consultant

Verónica del Carril, IPDC / Journalism Education Consultant

Graphic design: Aurélia Mazoyer

Cover design: Nico Bang Bang (Nicolas Chatain)

SHORT SUMMARY

Getting to grips with reporting on AI

The rise and control of artificial intelligence (AI) is impacting society as a whole. It follows that AI coverage must inform audiences about the implication of the technology itself, beyond journalism. For instance, reporting on the power dynamics in the changing relationship between companies, authorities, citizens and computer chips, and between data and algorithms. While many AI deployments serve public interest, journalists also need insight and expertise to alert about aspects like exclusions, unequal benefits, and violations of human rights.

As part of its journalism education series, UNESCO's International Programme for the Development of Communication (IPDC) has supported the World Journalism Education Council in commissioning this handbook. The aim is to inspire and empower so that journalism educators can help both journalism students and working journalists do justice to one of the major issues of our times.

The handbook covers:

- Understanding machine intelligence and identifying different types of AI
- Exploring AI's potential, as well as its strengths and weaknesses
- Imagining diverse futures with AI by recognising pervasive popular narratives that inform public consciousness
- Understanding journalism's role in mediating and shaping AI discourse
- Finding ways of reporting about AI in a nuanced, realistic and accountable manner
- Making connections to existing genres of journalism, ranging from general news reporting to data journalism

Strengthening journalism education is one of the key results sought by IPDC, a unique intergovernmental programme within the UN system that specializes in media development.



60%
of tech news
is dominated by industry products; the biggest category of sources is business-affiliated persons (30%)



"Since wars begin in the minds of men and women it is in the minds of men and women that the defences of peace must be constructed"



unesco

UNESCO Series on Journalism Education

A Handbook
for Journalism Educators

Reporting on Artificial Intelligence

Maarit Jaakkola, editor

Contents

Preface	6
Foreword: Addressing the Diversity of Artificial Intelligence	9
Foreword: Journalists Need to Know about New Technologies	13
Introduction: Pedagogies of Teaching about AI Technologies	15
Artificial Intelligence as a Curriculum Subject	20
The Focus of this Handbook	23
The Structure of this Handbook	25
Acknowledgements	26
<hr/>	
Module 1: Defining Artificial Intelligence	27
<i>Outline by Agnes Stenbom</i>	
<hr/>	
Module 2: Cultural Myths and Narratives about Artificial Intelligence	37
<i>Outline by Edward Finn and Suren Jayasuriya</i>	
<hr/>	
Module 3: Policy Frameworks and Recommendations for Artificial Intelligence	49
<i>Outline by Fredrik Heintz</i>	
<hr/>	
Module 4: Reporting in Algorithmic Cultures	61
<i>Outline by Frederic Heymans</i>	
<hr/>	
Module 5: Reporting on Malicious Uses of AI Technologies	69
<i>Outline by Santosh Kumar Biswal</i>	

Module 6:	
Patterns of Storytelling about Artificial Intelligence	81
<i>Outline by Maarit Jaakkola</i>	
Module 7:	
Artificial Intelligence in Journalism and Journalistic Practice	94
<i>Outline by Jenny Wiik</i>	
A checklist of eighteen pitfalls in AI journalism	109
<i>By Sayash Kapoor and Arvind Narayanan</i>	
Afterword: Educating AI-savvy Journalists	115
Glossary of Terms	120
Contributors	124
References	127
Table of figures	135

How to read this Handbook

In each module, you will find

- | | |
|---|--|
|  Summary |  Outline |
|  Questions for Reflection |  Suggested Assignments |
|  Module Aims |  Readings |
|  Learning Outcomes | |

Suggested Assignments will be broken down into

- | | |
|---|--|
|  READ AND REFLECT |  SEARCH AND SOURCE |
|  DISCUSS |  PRODUCE |
|  IDEATE | |

Preface

The first text below was written by GPT-3 at OpenAI (<https://openai.com/api/>). GPT-3 refers to Generative Pre-trained Transformer 3, which is a third-generation autoregressive language model using deep learning to produce human-like texts (see Floridi & Chiriatti, 2020). The instruction given for the robot was the following: "Write an introductory text about journalism and AI. Include pros and cons. Include democratic perspectives."

Journalism has long been a field that has been slow to change and adapt. However, with the rise of artificial intelligence (AI), journalism is starting to change. AI is starting to be used in a number of different ways in journalism, from help to write stories to identifying fake news. While there are some benefits to this, there are also some concerns.

One of the benefits of AI in journalism is that it can help writers create better stories. AI can help identify patterns and trends that human reporters may not be able to see. AI can also help gather data and information for stories. This can help reporters save time and energy that can be applied to writing a better story.

However, there are also some concerns about AI in journalism.

One worry is that AI could be used to create fake news. This is because AI can be used to generate realistic-sounding stories that are not actually true. This could lead to people being misinformed or believing false information. Another concern is that AI could be used to control what stories are being reported. This is because AI can be used to identify what stories are being shared and liked the most on social media and then prioritize those stories. This could lead to a biased and one-sided view of the news.

Overall, there are both benefits and concerns when it comes to AI in journalism. It is important to be aware of these pros and cons when considering how AI is changing the field of journalism.

The texts below are two (of several) options provided by ChatGPT, in December 2022. This is a beta service from OpenAI.com. The responses are to the question: "How should journalism teachers instruct student on how to cover AI?"

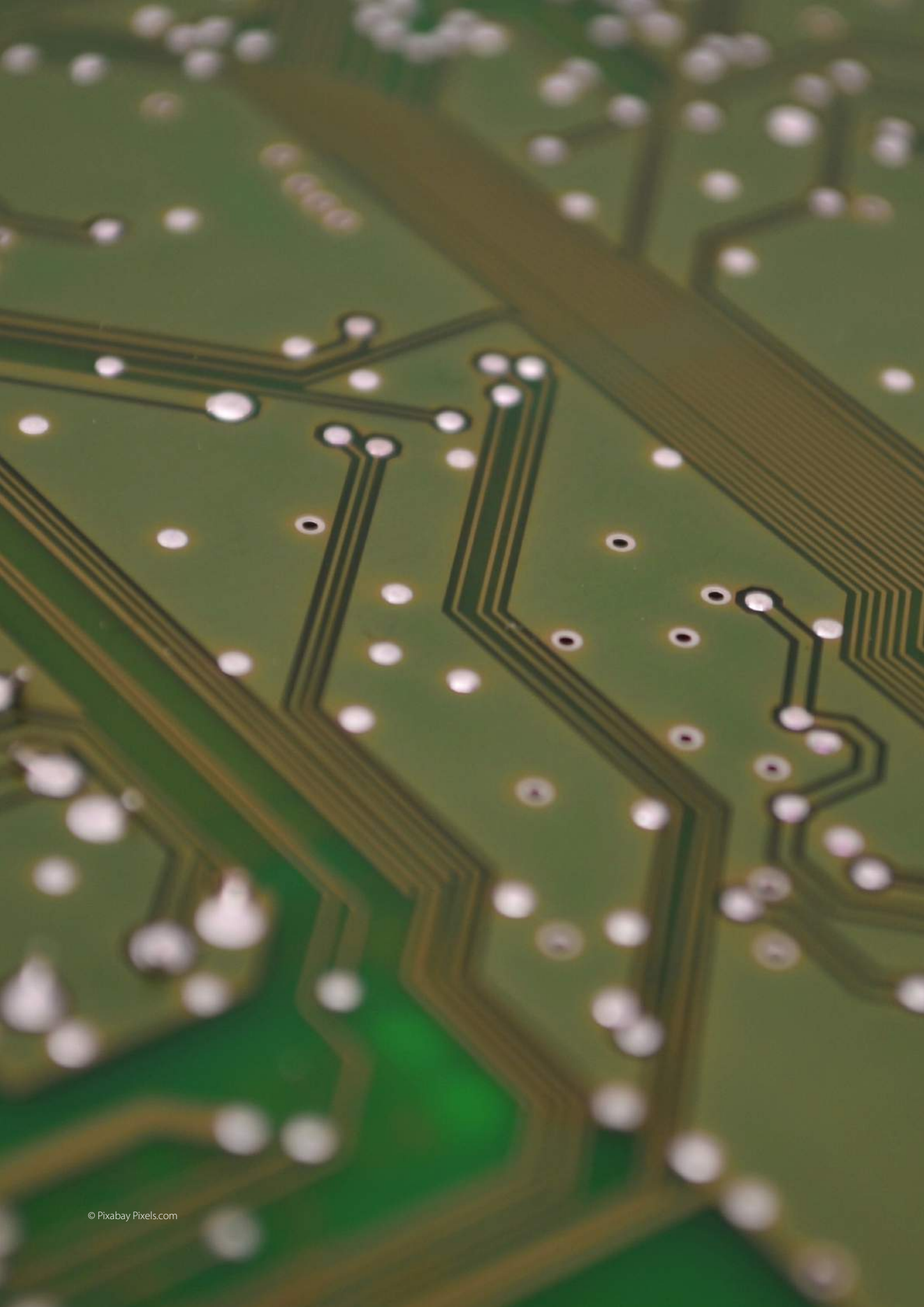
1. Start by explaining the basics of AI and its various applications in various fields such as healthcare, finance, transportation, etc.
2. Discuss the potential impact of AI on society, including both its benefits and drawbacks.
3. Emphasize the importance of researching and fact-checking when covering AI, as there is often a lot of misinformation and hype surrounding the technology.
4. Encourage students to seek out expert opinions and perspectives when covering AI, and to consider the ethical implications of the technology in their reporting.
5. Discuss how to approach interviews with AI experts and how to ask the right questions to elicit informative and engaging responses.
6. Encourage students to use a variety of sources and formats when covering AI, such as interviews, data visualization, and expert analysis.

7. Provide guidance on how to cover AI in a way that is accessible and understandable to a general audience, while still maintaining a high level of accuracy and depth.
8. Help students develop a critical eye when covering AI, encouraging them to question assumptions and think critically about the implications of the technology.
9. Provide opportunities for students to practice their AI journalism skills through in-class assignments and real-world reporting experiences.
10. Encourage ongoing learning and professional development in AI journalism, as the field is constantly evolving and changing.

Journalism teachers should instruct students on the following key points when covering AI:

1. The importance of staying up-to-date on the latest advancements and developments in AI.
2. The need to understand the ethical and societal implications of AI and its impact on various sectors and industries.
3. The importance of seeking out diverse perspectives and expert opinions on AI in order to provide a balanced and nuanced coverage of the topic.
4. The need to critically evaluate the claims and statements made by experts and industry leaders on AI, and to fact-check and verify information before publishing it.
5. The importance of being aware of the potential biases and limitations of AI technology, and to consider these factors when reporting on its use and application.
6. The need to communicate complex technical concepts and developments in AI in a clear and accessible manner for a general audience.
7. The importance of using a variety of sources and formats, such as interviews, data analysis, and visualizations, to provide a comprehensive and engaging coverage of AI.

These three examples of AI-generated texts contain points that intersect with the chapters that follow. But readers will nevertheless find a wealth of human-generated insights in this handbook, along with examples and creative exercises. For the present, journalism educators still have a unique and significant contribution to make.



Foreword:

Addressing the Diversity of Artificial Intelligence

Guy Berger,

former secretary of UNESCO's International Programme for the Development of Communication (2011-2022)

That many writers have contributed to this handbook helps explain the variety of views and definitions in its pages. Some chapters refer to “**responsible AI**”, others speak of the need for AI to be “**ethical**”. The contested phrase “Fourth Industrial Revolution” pops up. This diversity tells us what should be a starting point for teaching the coverage of Artificial Intelligence (AI) - namely, that there is a plurality of understandings at play, and that no one should take assumed semantics as if they were common currency.

Accordingly, at the very outset, journalism students should be acutely aware that the term AI itself means different things to many people. As journalists, their first job is to avoid conveying to audiences that there is a clearcut meaning of this quasi-magical combination of words.

Despite marketing impressions, the use of algorithms does not amount to AI (although AI assumes the deployment of algorithms). Data is essential to the AI package - but not all “data-driven” issues and developments are congruent with AI.

What the public, and hence media, should comprehend is that the diverse definitions and interpretations of AI tend to reflect different contexts and varied interests. This is especially when the term “AI” is used as “theatre” in order to hype the offers of consultants, companies and politicians.

! Critical journalistic scrutiny is essential for all sources (and statistics) concerning AI-related news – no matter whether sources are academics, regulators, big corporates or entrepreneurs. The same point applies to AI when reported as part of the operations of computer gaming companies, e-security services, vendors of Non-Fungible Tokens, crypto-currency exchanges, health care, agriculture, warfare, etc. There is a need to interrogate and deconstruct AI references, whatever the field of coverage.

Arguably, the most interesting thing about AI is often not the technicalities around the software, nor the hardware side. It is the **inter-linked human dimension** – including the vested interests always involved in defining and promoting AI, in developing software and control and access of hardware. And of course, the availability, ownership and conditionalities around data.

In point of fact, each micro-story – eg. like a medical breakthrough using AI – will be part of a bigger systemic human story that provides many rich angles to reporting the specifics. To grasp the extent, requires that we

recognize how the ensemble of technologies broadly known as AI are inextricably embedded in society and its power dynamics and challenges. In turn, this observation should alert those without specialization in AI technology to the journalistic possibility of the many human leads that they can follow when AI becomes evident as a component of their reportage on topics that seem, at first, separate from technological dimensions.

Attention to the human angle points, for example, to many strong stories about human labour. One set is about the labelling of data for AI in the Global South. Another set comprises stories about the (limited) demographics of the coders, whose outputs curate and interact with the data. Indeed, journalism can unearth the challenges of diversity and inclusion, along with brain drain. It can enlighten us about how these give specific shape to AI and how they impact on certain possibilities (eg. AI in self-driving cars) rather than others (eg. AI in managing public transportation).

The human angle includes people as objects of much AI – whether through facial recognition or AI-powered determination of mortgage loans or court sentencing. The **ubiquitous extraction (plus sale and use) of data from people to fuel AI** is another narrative that needs continuous monitoring – like the 2022 news about concerns of menstruation apps that disclose pregnancy conditions to third parties. And, as writers like Kate Crawford and Shoshana Zuboff have shown, data are not just a resource that is somehow there for the taking – which is a key observation for journalists to dig into in regard to AI stories.

It can be added that while “users” of AI may be active agents, they are also frequently at the same time raw materials for “commodities” being produced for sale to clients interested paying for access, not least advertisers and others with interests in manipulation rather than human autonomy.

The human questions around data include, but also go beyond, the issues of potential biases. They enjoin us to ask about individuals’ awareness and permission about the capture, use and ownership of their data assets, especially (though not only) when this concerns personal data. In this regard, the rise of AI has prompted interest in **alternative models of governing data** – for example, the idea of “data trusts” to facilitate legitimate sharing between data-holders and data-users – which is another story for journalists to find and tell.

While considering the human issues in AI, we should never of course leave out of reporting the environmental dimensions of AI. The “cloud” and “bitcoin”, for example, are firmly grounded in massive data centres consuming substantial energy with carbon emission consequences – a story that is relatively under reported.

In all this, the **centralization of power over AI development and deployment**, along with its network effects and path dependencies, cries out for **deeper media attention**. We’re talking here about a trend that is truly profound in terms of its significance for humanity as a whole in terms of human rights, democracy, environmental sustainability and inclusive development.

The paradox, in a nutshell, is that in reporting both AI stories - and AI within other stories - student journalists will need to see far beyond the technology itself. The entire chain of linkages needs attention... yet it also offers a cornucopia of stories to pursue, perhaps more than many other developments.

As a common thread in all this, at the back of all would-be reporters’ minds needs to be a consideration of how each and every AI development and deployment relates to human rights and related norms.

This means knowing how to ask the questions about the existence (or not) and quality (or not) of ongoing human rights due diligence by AI developers, their institutions and customers and clients. Journalistic interviews should never neglect to enquire about pre- and post- impact assessments and resulting mitigation steps.

Elaborating on this, we should recall that it is part of a journalists' general vocation to assess what human rights may be harmed. In the case of AI, rather than be blinded by the promises of progress, it is key to consider especially the rights to dignity, privacy, equality and justice. In this vein, news stories can put on the political agenda the issue of ensuring independent audits of AI effects - and of unintended effects in AI deployments. Reporters can also play a role in bringing whistle-blower efforts into the light, triggering changes in both companies and governments in relation to their use of AI and its multiple components.

Further journalistic questions for AI-linked stories can well include:

- What ethical protocols have been used concerning data creation, storage, sales and usage, and what testing of deployments have been done?
- What harm incident reporting and tracking is in place (or not), and what can be learnt for improvements?
- Does a system stack up against the global standard that "both technical and institutional designs should ensure auditability and traceability ..."? (The quote comes from clause 43 of the 2021 UNESCO Recommendation on the Ethics of Artificial Intelligence, agreed by 193 Member States).

Without informed and independent journalism on AI, society at large will have huge deficits in the **transparency and accountability** which are absolutely essential for the governance of such an impactful technology. On the other hand, it is quality reporting that can delve into what packages of state regulation and self-regulation are optimum, and what potential exists for alternatives like co-regulation and multi-stakeholder **regulation**.

Journalists can profitably apply the same critical mind-set to the integration of AI systems into their own workplaces and newsroom practices. And, of course, as in other fields, for all stories that touch on AI, transparency is also a must when journalists or their outputs are sponsored in one way or another.

All this dynamic character and the debate around AI makes for exciting opportunities for journalism educators, students and working journalists. In this context, UNESCO thanks the editor Maarit Jaakkola, the contributors and the World Journalism Education Council, under leadership of Verica Rupar, which has overseen the initiative resulting in this handbook.

Strengthening journalism education is one of the key results pursued by UNESCO's committee of 39 Member States who make up the International Programme for the Development of Communication (IPDC). Supported by IPDC funds, this handbook is part of the UNESCO series on journalism education.

Some publications in this series have met with such enthusiastic receptions that readers have offered voluntary translations. If any readers of this current handbook are interested to volunteer to help make this current resource available in other languages, you are warmly encouraged to do so. In addition, you may have ideas on addressing gaps in the work. Peer reviewers of this publication have urged that greater consideration be given to student voices, and to experiences from the Global South, in regard to teaching journalism about AI.

Anyone interested in translating and/or offering further research to supplement this handbook (perhaps for publication as a companion volume) is encouraged to contact the IPDC.Secretariat@unesco.org

Meanwhile, we commend your reading of this knowledge resource. You will be richly rewarded. Do let UNESCO know what you think. Your efforts are invaluable to helping journalism education maintain relevance in this fast-changing world.



FIRE ESCAPE

Foreword:

Journalists Need to Know about New Technologies

Charlie Beckett,

Professor of Practice in the Department of Media and Communications at LSE, and Director of Polis, an international journalism think-tank.

Go search online for “Artificial Intelligence” and the resulting images will be a shiny robot or glowing computer circuits. It looks exciting and futuristic but also quite frightening. Journalists feel the same.

! Are the robots going to take our jobs? Will algorithms write articles and edit the news?

The reality is both more worrying and more exciting.

After more than three years spent working with journalists around the world using Artificial Intelligence, I am convinced that these technologies can help us make our work more efficient, effective, and engaging. But AI is not going to “save” journalism or transform it completely. In practice it does quite boring, repetitive, simple tasks. It needs to be carefully programmed, adapted and managed to suit editorial demands.

AI is not so easy to use and it’s not as “intelligent” as you might expect. It brings with it **problems of bias and ownership of the technology**. It means news organisations will have to invest in new skills and ways of working. Many don’t have the resources to do that and might fall behind. But AI does already have the power to make a significant difference. The hope is that it can augment the human journalists, giving them more time and resources to use their human skills such as creativity, empathy and judgement to add value to their journalism.

Whether you take an optimistic or pessimistic view of this technology, it is vital that we all pay more attention to it. In the near future all aspects of our lives are going to be more data-driven and algorithmically-powered with forms of artificial intelligence shaping our societies. As journalists we need to understand what this technology is and what it can do. Partly because it will influence how news is made and consumed but also because we will need to do journalism to **inform citizens about its effect on our world**.

As this handbook will show, AI is difficult to define. In practice **pure AI does not exist**. There are no robots who think for themselves. In practice it is a mixture of machine learning, natural language processing, automation and data processing. You already use “AI” when you search online or plot your way home on a mapper app. The actual workings of AI might be complex but that’s no excuse not to understand its applications and potential. This handbook is a great resource for that.

At the Journalism AI project at LSE we tapped into a global demand for more information and debate around AI in journalism. We created training, information resources and a programme of collaborative innovation programmes with journalists around the world. They told us that AI can improve all aspects of their work. It can

help personalise content so that people get the news that is relevant to them in the way and at the time that suits them. It can help find stories that are hidden in data or to search for **new voices**. It can help create content, especially the more simple, functional, factual stories such as straight-forward financial, sports or weather reports. It can be used to fact-check at scale and even counter the human biases of journalists. At the moment the adaption of AI is being led by the big news organisations such as the BBC, Bloomberg or Wall Street Journal. But if you go to the Journalism AI website you'll also see great examples of smaller newsrooms developing clever uses to super-power their journalism.

! This is the third wave of technological change for our news media along with other sectors of our increasingly digital "metaverse".

First, we all went online, taking our traditional processes onto websites. Then about 15 years ago along came social media and our journalism became interactive, multi-format with a multitude of sources and platforms. Now this third wave of AI technologies promises to take over much of current journalism labour. And in addition, it offers the opportunities to develop new forms such as **Augmented Reality, audio or video on demand, hyper-personalisation and even greater diversity through automated translation, transcription and editing**.

Journalists also told us that there needs to be a debate about who produces and controls this technology. What **biases are inherent to the data sets or programmes** that are created? How do we use these tools and systems to improve journalism and not just create clickbait or misinformation? It is vital that journalism educators and students are part of this debate about how these technologies might shape the news media. It is also important that technologists, regulators, politicians and citizens have a voice in the debate. We live in a world where all our lives – from shopping to politics – is becoming more mediatized, more online, more social. It brings enormous convenience and exciting opportunities for expression, entertainment and education. But we are now also familiar with the problems that the digital world also contains such as misinformation and extreme speech.

As someone who spent more than two decades as a journalist in the analogue era, I am very aware of how much has changed. It excites me. Journalism, like many other professions, has become more competitive and in some places around the world, more dangerous. There is a huge responsibility on the next generation to deal with the ethical dilemmas that these new technologies bring. But they could also enjoy some new powers.



Introduction: Pedagogies of Teaching about AI Technologies

Maarit Jaakkola

Providing guidelines for future journalists on the topic of AI as a societal and, therefore, journalistic issue implies facing a major and complex work in progress. Even though AI-related technologies have existed for a while, it is not until now that many of the futuristic ideas of previous centuries are beginning to take shape. The AI-ization of society is a long, slow process that penetrates all sectors of life. When it comes to collecting advice and best practices for covering the topic in journalism, there are thus open questions and insecurities at stake that need to be taken into consideration.

Many of the aspects we talk about when referring to AI will possibly occur in the future, but we are not there yet. The technological development required to achieve “**superintelligence**” and “**singularity**” is a future process related to the development of computing – what it means exactly and how it can affect different sectors of life are questions whose answers cannot be fully and straightforwardly foreseen. The related public discussion and policy development have become more intense during recent years, but societies still have no shared, established and profoundly tested frameworks with regard to how to approach AI from democratic, or citizen and consumer, perspectives. Policies and the successive best practices are still very much in the making, not least in terms of using AI technologies in and for journalism. Simultaneously, more research in the fields of social sciences and the humanities is trying to find out how AI could be related to the existing strains of research, and new studies are constantly being published. Overarching works, such as *The Routledge Social Science Handbook on AI* (Elliott, 2021), were not published until very recently, and a number of them are still under preparation.

‡ *Consequently, we are to some extent lacking vocabulary to address the ontology and epistemology, and the implications and consequences, of AI as a societal phenomenon.*
‡ *What are the right questions for journalists to ask?*

What is clear is that **AI is an outcome of a socio-technological process** that poses a number of challenges for journalism and journalistic practice – both as a **topic of coverage and a method, technique and approach in journalism itself**. Journalism should deliver valid, fact-based information to citizens in its role as a public good that works as a resource for knowing things. This also applies to technologies, and the concept of AI, a buzzword that has become commonly used in everyday language, seems elusive to many. The term “intelligence” may evoke associations and visions that are far from our everyday practices, while, paradoxically enough, AI to a great extent is a profoundly practical phenomenon that is increasingly affecting our daily lives. Between intimidating dystopias and celebratory hype, it may be difficult to equip oneself with reasonable expectations.

Experts have suggested several tasks for journalists to undertake in the public sphere (see, e.g., Hansen et al., 2017, p. 7). It has been said that journalists have to find needles in haystacks. They should also be alert to identifying trends, as well as departures from the major trends. They should commit to examining an application of AI or computation as the subject of the story itself. Some scholars have even argued that journalists should be actively involved in the design of new technologies, integrate the journalistic perspective into such structures and partner with policymakers who shape the ethical frameworks for future AI. While many may disagree with the industry-active role of journalists, perhaps the most important task for journalism all around the world is to be involved in making sense of the technology-driven phenomenon and actively relating it to different contexts to be able to create, inform, balance and renew public discourses. With regard to journalism education, the earlier journalism educators can be part of discussions and take on an engaged role in shaping future journalistic thinking, the better the chances are that future journalists will be equipped with in-depth understandings of the characteristics of future societies.

To slightly exaggerate, there is no such thing as “reporting on AI”. Very seldom do journalists report on the technology itself, as a pure isolated entity, and very seldom do journalists thus need to know about the technical operations of neural networks or deep learning. Instead, more often journalists report on democratic processes, people’s behaviour or, for example, structural injustices – AI integrated in all these, and all these influenced by AI technologies. Because of the increasing centrality of AI to societal processes, journalists need to know about the central ethics, ethical discussions and regulations around AI, and realize the fears and hopes that AI may evoke in the popular imaginary. Journalists need to reflect upon AI in their own field, as well, and be prepared to approach AI-driven techniques and methods in an accountable way to maintain a sustainable relationship with their audiences and the thing that lies at the core of journalism: **trust**.

Before discussing AI as a curriculum topic, I would like to highlight two central observations that are relevant to journalistic thinking and practice resulting in the coverage of AI. First, AI is to some extent influenced by the very idea of being a **future projection**, and the reporting of it needs to be balanced against the extremes of the apocalypse and technology celebration. Second, AI is an umbrella term, and the information concerning it always needs to be synthesized from different sources representing different fields. Finally, there is still no consensus upon whether, to what extent and, above all, how subjects such as AI should be integrated into formal journalism curricula. What is clear is that facing a new phenomenon entrenched with **ethical challenges**, such as **human-computer interaction and human-like machine behaviour**, brings us to the very core of journalism: journalism needs to remain journalism by sticking to its core values, which make it distinct from policymaking, lobbying and advocacy, marketing and influencing. Next, I will address the aspects of future orientation and interdisciplinarity, respectively.

"Singularity will soon be here, and the machines will take over!"

"Intelligent machines will save the world!"

"Our everyday life will be similar to science fiction!"

AI needs balanced coverage.

The public AI discourse is shaped, like discourses on all emerging technologies, by uncertainty related to the realization of future projections. As discussed in Module 2: Cultural Myths and Narratives about Artificial Intelligence, this idea has been accompanied by a large body of literature, films, television programmes and other cultural products depicting the "robot race". Such an approach arouses interest in the topic and increases its value as a public matter because threats and alarms are always an entry point into the public's attention. At the same time, hybrids, like werewolves and vampires, have universally fascinated the human mind, arousing excitement filled with identification and otherness.

What may be forgotten in the emphasized novelty of AI, however, is that it is not a new topic. In 1987, when Brian B. Bloomfield published his seminal book *The Question of Artificial Intelligence*, intelligent machines were hotly debated at the onset of the revolution of information technology.

! *The increased penetration of information and communication technology (ICT) in society has resulted in the concept of an information society (Hofkirchner & Burgin, 2017) and, further, derivatives, such as a virtual society (Woolgar, 2003), a data-driven or data society and an automated society (Bloomfield, 1995).*

Individuals and organizations are expected to encounter massive amounts of data and evermore complex data structures, eventually resulting in computer systems that operate independently, to which the term AI refers. To achieve an "AI society" – or, in fact, scholars and policymakers often use the term a **"good AI society"** to highlight how it can be used for social good (Cath et al., 2018) – journalists are expected to mediate knowledge about these processes while they themselves are also increasingly required to become data analysts and curators with at least a basic understanding of computation as an essential tool for reporting.

Like many of the modern-day challenges that journalists need to face as topics in their reporting, such as climate change, the planetary resource crisis and pandemics, AI constitutes an abstract object with global implications for journalistic inquiry that is more inclined to report on concrete occurrences that can be clearly limited and identified. It is ubiquitous, yet it may be difficult to identify; it renders diverse implications that may be hard to

foresee. It requires expert knowledge to be profoundly understood, and developing full comprehension often demands triangulation, that is, integrating and compromising different perspectives. Last but not least, the very phenomenon of AI raises the question of whether journalism should remain as a mirror reflecting the ongoing processes and the entire project in the making or take on a more interventionist, solution-seeking position by exploring alternatives and suggesting answers instead of formulating questions.

Furthermore, the contemporary narratives concerning AI are pervasive and imagination-captivating, as they are very closely related to science fiction and fantasy, depicting utopian future horizons as potential catastrophes. Both theorists and practitioners of AI continue arguing that the **narrative of machines taking over is too dominant**. In many cases, the more subtle, down-to-earth everyday questions regarding machine learning and ubiquitous computing are overshadowed by the fascination with the fantasy of machines taking over humanity, while the algorithmic cultures shaping our everyday lives are dismissed as boring and uninteresting. Journalists should not be misled by the science-fiction fantasies, even though possessing an understanding of them is crucial. Journalists should be more curious about asking how intelligent systems affect us in versatile ways and looking for alternatives.

At the point of writing this handbook, there are still many unanswered questions and a lot of space for speculation when it comes to the uses and possibilities, risks, harms and threats of AI. This means that journalists face a specific situation where they need to combine different fields of knowledge and be able to deal with insecurities, while, at the same time, not evoke futile expectations or hopes or unnecessary fear or dystopias. This handbook intends to prepare educators for getting to grips with the society that we live in, which is increasingly penetrated and affected by technologies. Specifically, we want to answer the following questions:

! What kind of relationship should journalists, as public gatekeepers and opinion makers, establish with AI? When participating in the journalistic coverage of AI-related topics, what kinds of things should be taken into consideration? Which sources should be drawn on?

Teaching about AI as part of journalism education means adapting a flexible position with regard to both the society that is being covered and journalism. When covering AI, and perhaps with the help of such technology, journalists need to ask wide-reaching questions about the past, the present and the future. The ways in which things work now may not be how they will work in the future, and journalists may also be in the influential position of being able to impact the future. Typically, **journalism education** is about achieving a balance between the academia and industry (see, e.g., Zelizer, 2004); a recurring question is whether journalists' training should be based on the **status quo of the industry or reinventing ways of reporting**, thus providing a horizon beyond the current state of affairs. In the coverage of AI, this becomes especially prominent.

Our understanding of AI constitutes an interdisciplinary infrastructure, where expert sources need to be identified across different fields of knowledge. Academic research that can deliver an in-depth understanding of the phenomena involves theoretical development using very specific concepts to capture such aspects. Technological development that leans upon the innovative application of these technologies is not always very easy to access. The everyday experiences of laypersons and ordinary users are often not based on knowledge of these infrastructures and infrastructural development, as users do not really know how things work; they just happen to work. Our ways of talking about AI are based on circulated discourses. The task of journalism is to balance these knowledge fields, bridging different aspects to gain insights into what is of public concern. In order to perform such a balancing act, a journalist needs to gain knowledge about these fields and be able to filter and synthesize information that may sometimes be contradictory and complicated. Nevertheless, the challenge of creating a public discourse and understanding the phenomena that are increasingly powerful in our lives is a task for journalists in service of the public good. We need pedagogies to support this.

*"Will machines take our jobs?"
 "Can all the consequences of AI be anticipated?"
 "When will ideas become reality?"*

AI reporting needs to face and communicate insecurities.

AI constitutes an **interdisciplinary field of inquiry, or a "trans-domain field"** (Zhang & Pérez Tornero, 2021), where different areas of knowledge need to be monitored and drawn together to create a proper understanding of the technological, ethical, economic, societal and cultural aspects of AI systems in society. As a research field, AI draws on computer sciences, social sciences and also many areas of academic research that are interdisciplinary, such as human-computer interaction (design), game studies and policy inquiry. All these fields and subfields prioritize different aspects, placing, for example, the role of technology at the forefront in terms of the costs of examining the social effects. To address AI in public discourse in a lucid, understandable and sometimes didactic way is thus not always a simple task, and the journalists covering such topics need to consider how best to do it. Many of the concepts, such as machine learning, the internet of things, robotics and deep learning – or the very concept of AI – may not be familiar to general audiences.

Because of the interdisciplinarity of the topic, AI constitutes a **societal area of collaboration**. Even researchers and policymakers may interpret the results of machine learning and other automatic processes in contradictory ways.

! Computer systems, even if they are said to be "intellectual" and capable of imitating human behaviour and adopting human-like characteristics, do not operate on the basis of values and ethics, and they are unable to make related decisions that require the balancing of ideologies and cultures and may even need negotiation and compromise that may seem irrational in terms of logic.

Even if "the computer says no", humans need to interpret, contextualize and problematize the answer, and this responsibility should not be left to "intelligent" non-humans. In the extensive brain work related to this contextualization, journalists need to be partners of other actors governing AI technologies' future directions.

For journalists, there is – so far – no established genre or professional niche known as "AI journalism". Journalists of any genre may need to be able to question why and how disruptive technologies matter and how to deal with questions without **falling into technological determinism**. To some, AI might appear to be a topic that is covered by specialized forms of journalism, such as science journalism, health journalism or service journalism. As society becomes more and more digitalized, automatized, datafied and mediatized, the need to understand the technologies at work becomes more urgent. Journalists, ranging from those involved in political journalism to arts and cultural journalism, and involved in coverage spanning the international to the local level, regardless of the medium they are working in (newspapers, magazines, online news outlets, radio, television), need to understand how AI-powered technologies affect and penetrate all sectors of society.

One solution at such an early stage is to invite specialists in AI to visit journalism classes. In today's digital environment, video lectures are relatively easy to arrange, and organizing an international lineup of speakers is not an impossible task. Information is widely available online, and, for example, the European Union is monitoring the rapidly developing policies at the international level. Much of the translation work from the scientific and technical domains into the social domain and the democratic public sphere that journalism is

generating still needs to be carried out by journalism educators in collaboration with their students. Therefore, we can assume that if this handbook were to be written or revised ten years in the future, the guidelines would be different. Some aspects may develop into more complex sets of questions, while others may gain precision.

In the process of forming AI curricula, universities should remain independent and, without a doubt, partner with industry organizations, even if they themselves, to a large extent, hold the most recent insights into and information about technological development and bearing in mind that such institutions will typically lobby their own aims. Critical distances regarding what AI should be and in which directions it should be developed need to be established.

“How should the involvement of AI be reported in journalistic products?”

“To what extent should audiences know about the AI processes in journalism?”

“How should the non-preferred effects of AI reporting be mitigated? Are there unethical aspects of AI that should not be covered in journalism?”

Journalists have to reflect upon and be transparent about AI in their own reporting as well.

⇒ Artificial Intelligence as a Curriculum Subject

Computer-aided journalism has been an area of debate for many decades within journalism and journalism education. With regard to journalism education, AI is a relatively new knowledge area, which is typically considered within the category of new skills and competences, together with online journalism, mobile journalism and data journalism. Despite being debated and studied for decades in computer sciences, AI started to emerge as a topical area in journalism education through its integration with news organizations' work. The question of AI is attached to debates on new methods, such as the automated collection of big data, and new skill requirements, such as programming. For curriculum developers, this implies questioning which areas need to be excluded if new areas of learning are included in the curriculum.

Still, AI constitutes a concern for many stakeholders, and what is clear is that questions regarding journalism's role in democratic data-driven societies should be discussed in a proactive manner. It is vital to consider to what extent journalism and journalism curricula should be affected by technological development and policymaking. Traditionally, the relationship between journalism education and industry is ambivalent; while it is important to follow and reflect developments and match competences taught within journalism education with the outside world, journalism education should always be a step ahead, even regarding systems that cannot formulate objectives by themselves because these aims will be formulated by the humans who make these systems.

The question of **journalism curricula** not only incorporates the problem of **what to teach and how**, but, as Mark Deuze (2006) puts it, also **“structuring, rethinking, and building institutions, schools, or departments of journalism”**. How AI will finally be integrated into journalism schools' curricula will be addressed in the coming years, and this question is also, to some extent, dependent on future industry and policy development. Similar debates have been faced in flash journalism, documentary journalism and programming, and are still being faced, for example, in data, immersive and podcast journalism. These are “new competences” that are

negotiated and may eventually be merged with more traditional storytelling and fact-validating techniques and approaches.

Journalists, as generalists covering politics and other societal issues that lie at the core of democracy, do not need in-depth knowledge of all the societal areas they monitor. Some areas that are considered to be of high societal relevance and that require deeper knowledge in terms of their coverage become established as areas of specialization. This is to a great extent dependent on **newsrooms' organizational structures**: whether units or teams are established that include specialists. Science journalism is an established branch of journalism, encompassing specialists with deepened skills and scientific knowledge areas, above all, in natural sciences. While some areas of specialization, such as political, economic and cultural journalists, have a long and established position in newsrooms, others experience variation in their positionings, such as science, medicine, consumer and environmental journalists.

This handbook leans upon the generalists' approach, based on the belief that a generalist needs to gain some basic knowledge from different areas related to AI: the basic definitions of what we are talking about when referring to AI, the typical myths and discourses affecting public imaginaries and, thus, the hopes and fears regarding AI and the existing policy frameworks and ethical recommendations, as well as the best practices, including potential pitfalls and self-reflexivity on what AI is within journalism itself. Generalists' strengths lie in their ability to ask the right questions and present knowledge in an easily accessible way for general audiences and bridge the gap between specialized societal areas, such as science and industry. This task, which is a demanding one, is based on journalists' ability to establish themselves as authorities through strategic rhetorical acts that produce **interpretive authority**, which lies not in what journalists know – they do not “possess” the knowledge – but in how they represent their knowledge (Zelizer, 1993). Figure 1 depicts the major areas of inquiry for AI: science, ranging from computer science and the humanities, with a clear focus on the different subfields of computing, to applied science, as well as industry, education, policy and the users' lifeworld. Journalists need to acquire literacy skills in these areas in order to access them as sources by mastering a distinct vocabulary, a set of core concepts and theories and their traditions of rhetoric and discourse. A prerequisite for this area-specific literacy is the mapping of central actors and stakeholders, as well as infrastructures, and understanding how these cultural and political economies work. For example, concerning science, they need to know what scientific disciplines are at stake, how scientific knowledge production occurs, how scientific publications can be accessed and interpreted and how science works more generally. When it comes to policies, journalists need to identify the central policymakers and understand their work and be able to access and monitor their policy documents. For industry, journalists need to understand the commercial logics of AI production and recognize the interests of AI producers and their stakeholders, maintaining a regular dialogue with them.

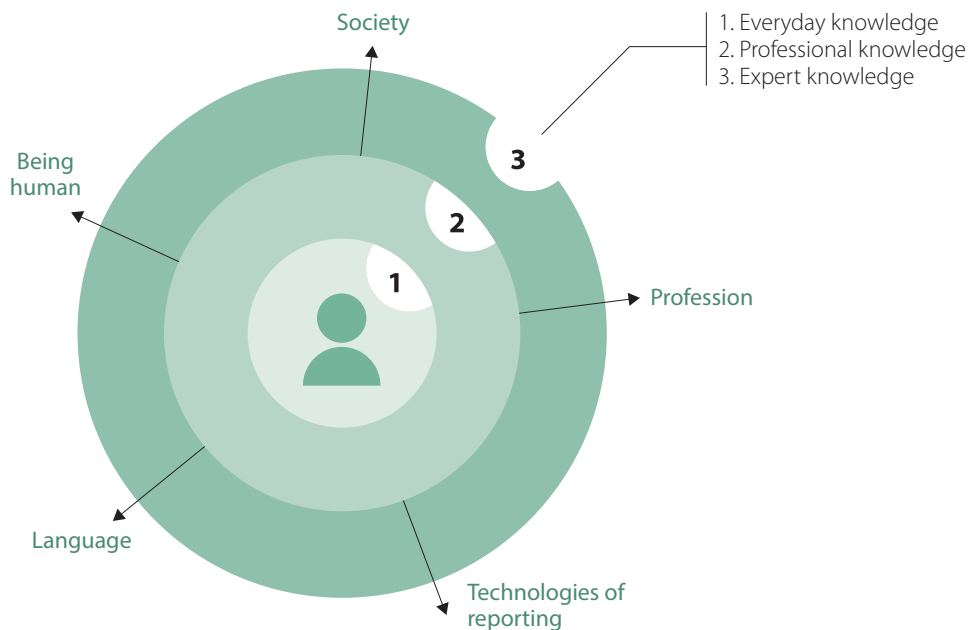
Preparing students to cover AI as a journalistic area does not significantly differ from other areas of coverage. Journalists need to employ their strategies as with any other subject, including the journalistic processes of contextualization, data verification, consequence and risk analysis, relativization, the identification of normativity and bias and the creation of critical distances to the objects of inquiry. Journalism on AI needs to be informative, fact-based and neutral, as well as independent, autonomous and representative, and, as with any other form of journalism, it should remain accountable, trustworthy and relevant to its audiences.

Journalism and journalism education can draw on existing experiences and lessons learned from reporting on other emerging technologies, such as nano- and biotechnology, science and other complex ethical challenges in society, ranging from segregation to cloning.

What kinds of competences are thus needed to cover AI? Figure 1 displays the competence areas that are of central importance for future journalists and that should thus be supported by journalism education pursuing

the quality coverage of AI and related emerging technologies. Sphere 1, the inner zone of the circle, represents an individual's lifeworld, sphere 2 is the journalism world and sphere 3 depicts the AI world, which respectively incorporate everyday lay knowledge, professional journalistic knowledge and expert knowledge related to AI. A journalist should draw on all these layers when using his or her competences, which are related to society, the profession, technologies, language and being a human.

Figure 1: Journalistic competence areas related to reporting on AI



Source: Original material by Authors.

When reporting on emerging technologies, such as AI, a journalism student constructs relationships between these core areas through the **lenses of everyday, professional and expert knowledge**:

- ▶ **Society:** notions of an automated and data-driven society in the modern-day democracy – citizens' experiences of AI, journalism's relationship to society, scientific knowledge of society
- ▶ **Profession:** notions of journalistic ideals, norms, practices and routines, "AI journalism" (AI as an object of inquiry) and "AI in journalism" (uses of AI for journalistic purposes) – personal professional experience, knowledge of journalistic professionalism, journalism research
- ▶ **Newsroom technologies:** the understanding and command of technologies related to journalistic production – the command of reporting tools, shared newsroom practices, best practices of technology uses
- ▶ **Language:** meaning-making, mastering expression, vocabulary and target-audience communication – command of meaning-making techniques, shared patterns for meaning-making in journalism and the newsroom, research on means of presentation
- ▶ **Humanness:** ontologies of being a human in a post-human era where human-machine interaction and hybridities between humans and computers increase – personal human experience, ontologies of humanness, philosophies of AI

A key pedagogical dimension when working on AI coverage for journalism educators is common notions and attitudes – you might talk about controlled development or a change in the cultural climate. How should

we relate to new and emerging phenomena? How can we observe their implications and, above all, the sociocultural concerns related to them, which affect democracies in a crucial way? How should we contextualize the emerging phenomena? How can we avoid painting dystopias or becoming advocates of biased interests? In pedagogical practice, this requires familiarizing oneself with different aspects of technologies and performing one's journalistic role in a proper way.

The disruptive character of AI technologies has as a consequence that while AI is transforming society as a whole, we are not only dealing with a certain type of coverage when discussing AI reporting. AI is expected to change the relationship between humans and machines in different sectors of society and life, including journalism, and this fundamental shift requires **reassessing our epistemological and ontological concepts of citizenship, society and human-machine interaction**.

In sum, the objectives that this handbook has set for understanding AI in and for journalism education include the following (cf. Long & Magerko, 2020):

- **Recognizing AI:** understanding what it is and distinguishing between technological artefacts that use and do not use AI
- Understanding **machine intelligence** and identifying different types of AI
- The **interdisciplinarity** of AI as a field and umbrella term and the diversity of technologies that enables it, ranging from cognitive systems to robotics and machine learning
- Exploring **AI's potential**, as well as its strengths and weaknesses
- Imagining future with AI by understanding and overcoming the **pervasive popular narratives** that inform the public debate
- Understanding **journalism's role** in mediating and shaping AI discourse
- Finding ways of reporting about AI in a balanced, realistic and accountable manner
- Making connections to existing genres of journalism, ranging from general news reporting to data journalism
- Increasing one's knowledge about the uses of AI in and for journalistic practice
- Enhancing **professional self-reflection** on the philosophical, ethical, social and cultural dimensions of AI technologies and journalism

⇒ The Focus of this Handbook

This handbook relates to previous handbooks on journalism education and journalistic practice published by UNESCO (Fengler et al., 2021; Hunter, 2011; Keith & Cozma, 2022; Impe, 2019; Ireton & Posetti, 2018; Muratova et al., 2021). UNESCO's series on journalism education is intended to function as a pedagogical resource in the form of a model curriculum, training manual or practice tutorial. Besides, UNESCO has published guidelines on AI for policymakers in the educational field (Miao et al., 2021). These handbooks highlight the challenges of policy implementation and collect best practices from a global perspective to inform those involved in practical work worldwide. In line with previous publications, this book is called a handbook, but to be more precise, rather than conforming to the academic tradition of handbooks, it is a pedagogical manual or a model curriculum outlining the pedagogy related to a relatively new yet deeply anchored societal phenomenon.

The handbook aims to develop educators' and learners' theoretical understanding of the phenomenon by relating this discussion to journalistic values as well as figuring out appropriate journalistic methods and practices, in addition to raising ethical concerns of public interest, which are also relevant for journalists. There are seven areas of focus:

- ▶ **Technology:** How can AI be defined and how can its development be traced?
- ▶ **Visions and development:** How do public visions and imaginaries affect our current and future understandings of AI?
- ▶ **Policy:** How should AI be regulated in nation states and internationally?
- ▶ **Infrastructure:** What are the central structures for producing, regulating, developing and consuming AI?
- ▶ **Everyday practices:** How is AI embedded in the social behaviour and cultural habits of citizens and consumers and the communities they are living in?
- ▶ **Misuses:** How can AI be used for maleficent purposes and how can one foresee and minimize the risks of such harmful uses?
- ▶ **Reporting practices:** What are the best practices and potential pitfalls when reporting on AI in journalistic coverage?
- ▶ **Journalism:** How is AI used in journalistic practice and how can it contribute to enhanced professionalism?

The handbook does not intend to make exhaustive and final claims regarding AI but rather support journalism educators in creating instructional designs and pertinent pedagogical practices that help journalism students to adopt a balanced perspective for approaching AI as a societal topic among many other areas of political, technological and cultural coverage. Therefore, this handbook should not only focus on the current perspective, or the state of things as they happen to be at the point of writing, but must instead pursue a more holistic perspective that will stand the test of time. One central aim is also to demystify a topic that may sound all too technical and futuristic to many and to provide a journalistic context to envision a phenomenon that has sociocultural consequences that reach well beyond the technical aspects.

! A crucial step is taken when we understand that no technologies are neutral: there are always a number of interests underlying the technological structures and the modus operandi of different platforms.

Furthermore, the development and uses of technologies are invested with emotions and attitudes. Technologies are approached based on **value choices**, and their uses are assessed in different ways by preferring and prioritizing some and ignoring and dismissing others.

This handbook is based on the following principles: the **diversity of voices** (because of the multi-stakeholder field, there are different perspectives and many unanswered questions, which makes it appropriate to approach the topic from different perspectives, giving voice to different experts); different **genres of journalism** (while AI can be integrated into all practices and remains an interesting and relevant question for all genres, it has so far been applied to a greater extent to some specific genres of journalism, such as political, science, technology and sports journalism, while other genres, such as arts and cultural journalism, still need to be explored in terms of the practical application of AI technologies within them; different genres also imply different sets of practices and traditions of reporting, together with different professional identities underscoring the professional values in slightly different ways, and that is why AI development in journalism should be sensitive towards genre differences in the journalistic field); and **research-based evidence** (the development should be based on recent proceedings in research, combining research evidence with the established ethical framework of journalism; journalists should be in direct contact with recent and rapidly developing fields of research, and

they should also monitor the research fields that inquire into journalism itself, such as journalism studies and studies on computational journalism).

⇒ The Structure of this Handbook

The handbook consists of chapter modules addressing technology, policy, infrastructure, everyday practices, misuses and journalism respectively. The structure of each chapter is the following:

- *Summary* provides a synopsis of the topic of each chapter.
- *Questions for Reflection* collect some preparatory pre-lecture questions for students to reflect in advance upon and small assignments that are typically limited tasks with observation of the one's environment. The objective of the pre-assignments is to stimulate the learners' own ideas with regard to the topic and support him or her to develop a personal relationship to the topic, to better take command and feel ownership of it. The questions and assignments can be carried out either individually or in small groups. Educators can choose which pre-assignments the learners should choose or ask the learners to select the preferred or most suitable assignments for an introduction. Pre-assignments suit well for being conducted at home before a lecture or a workshop, or in the classroom at the beginning of a common session.
- *Module Aims* highlights the pedagogical objectives to be pursued in each chapter.
- *Learning Outcomes* identifies the learning aims that the student is expected to reach learned having familiarized with the content of the chapter and conducted the assignments chosen by the educator.
- *Outline* marks the core of the chapter, presenting the key content. This part is co-written with an expert within the topic.
- *Suggested Assignments* collect learning activities that can be organized in different settings. Each assignment section contains assignments to be accomplished individually, in small groups, and within the entire group. The assignments include both assignments designed for lecture classes (assignments with theoretical approach) and training newsrooms focused on journalistic practice (assignments with practical approach or with a focus on journalistic content production).
- *Readings* identify literature in English that is of relevance to the chapter topic. Both academic books and journal articles as well as industry reports are included. Open access alternatives have been preferred, but it has, unfortunately, not always been possible to find such an alternative. Here, it is worth noticing that in a rapidly developing area of inquiry such as emerging and future technologies, the research becomes quickly outdated. Therefore, students should be encouraged to conduct literature searches on their own and educators should complement the readings with the newest studies available. Keywords mentioned at the end of the chapter summary can be used as search terms in academic literature databases (e.g., Google Scholar).

More specifically, the assignments of each chapter are designed to fit into the diversity of pedagogical approaches in journalism education and include the following categories, marked with respective symbols:



READ AND REFLECT: A theoretical assignment based on reading a scientific article and delivering a short reflection upon it. The purpose of these assignments is to make students familiar with acquiring information from scientific research.



DISCUSS: A preferably small-group assignment, typically related to a case, aimed at articulating opinions and views as well as arguments related to a topic. The purpose of these assignments is to enable the formation of opinions and promote reflection related to a specific topic.



IDEATE: A creative assignment that can be carried out either individually or in a dialogue with peers to find a journalistic angle on a topic, based on one concept that needs to be deconstructed. The purpose of these assignments is to teach students think journalistically around a topic.



SEARCH AND SOURCE: An interview assignment where the student needs to acquire information from a third-party source. The purpose of these assignments is to find appropriate methods and vocabulary to collect information.



PRODUCE: A journalistic assignment aimed at producing a journalistic presentation on a topic. The purpose of these assignments is to provide the student with a possibility to learn something new by acquainting with a topic through journalistic work.

The educators choose what types of assignments they want to pursue, according to the classroom conditions. It is worth noticing that assignments can always be altered or adapted to virtual environments. Integrating uses of digital tools is sometimes appropriate.



Acknowledgements

Guest writers have been invited to write the core part of each module, the “outline”. Otherwise, the pedagogical outline is written by the editor, based on experiences from journalism education in different countries.

To bring together knowledge and insights from different fields is necessarily a collective accomplishment, and several people have assisted me in compiling this handbook.

In the process of making this handbook, I have collaborated with and consulted both scholars and practitioners working on technological aspects and journalism. Thank you for separate contributions, spread across this book, to professor Teemu Roos from the University of Helsinki, Finland, professor emeritus Steve Woolgar from the University of Linköping, Sweden, professor Jenny Bergenmar from the University of Gothenburg, Sweden, professor Bernhard J. Dotzler from the University of Regensburg, Germany, Matthias Spielkamp and Nicolas Kayser-Bril from AlgorithmWatch, Germany, and professor Ammina Kothari from the University of Rhode Island, the United States.

In particular, I want to thank UNESCO and the World Journalism Education Council (WJEC) for the initiation of the handbook. In addition, I want to thank Guy Berger, Director for Freedom of Expression and Media Development at UNESCO, and Verica Rupar, the chair of the WJEC, for their generous support and valuable feedback along the way. I also want to thank all the proofreaders of this book, including Kristin Clay and Timo Lyyra, and Veronica Ines Del Carril for project management.



Module 1:

Defining Artificial Intelligence



Summary

This module defines the concept of artificial intelligence (AI) as an interdisciplinary umbrella term. It introduces the reader to some central concepts that are of relevance for understanding today's AI technologies in and beyond journalism, and highlights concrete examples of AI in practice. The module briefly outlines the origins and development of AI, and calls for a balanced understanding about the interplay of human and artificial intelligence in contemporary and future society.

Key concepts: *disruptive technology, applied AI, general AI, machine learning, deep learning*



Module Aims

The aims of this module include the following:

- ▶ Developing an overall understanding of what we are talking about when referring to AI.
- ▶ Learning the basic concepts that are essential for understanding AI.
- ▶ Discussing the relevance and implications of AI to public discourse, journalistic coverage and society in general.



Learning Outcomes

After the completed module, the student is expected to have learned the following:

- ▶ To define and distinguish between the different definitions of AI and AI technologies.
- ▶ To be aware of the central contexts of addressing AI in journalism.
- ▶ To identify AI in the technologically rich environment we are living in and have ideas on how to approach them journalistically.



Questions for Reflection

Here are some questions to reflect upon in advance:

1. What comes into your mind when you hear the term artificial intelligence (AI)? List your connotations freely and compare them with a peer. Did you come up with any similar ideas? How are these ideas possibly reflected in dominant public discourses on AI?
2. Envision the technological development of the future three decades in the following environments (alternatively, pick only one of them): home/family, school, healthcare. Which processes have been automated? How has automation affected people's behaviour, social interaction and experiences?
3. Whose voice is (or isn't) heard in mainstream debates around AI? Why do you think it is like this, and how might it affect our understanding of what AI is and can be?

Here are some preparatory tasks to do prior to entering the topic:

1. Talk to a person near you, a friend or a family member, and ask about their ideas about AI. What open questions and doubts do you perceive, and do you share these concerns?
2. Read through the daily news output in a newspaper or another news outlet and pick out articles covering technologies. Describe the style of presentation: which patterns are there?
3. Look for some concrete examples on how AI can be applied to the domains mentioned below by searching online for cases and examples.

Crisis prevention and management: prediction of disease outbreaks, migration crises, natural and man-made disasters, search and rescue	Economic empowerment: agricultural quality, initiatives of economic growth and inclusion, labour supply	Education: access to and completion of education, improving student achievement, teacher and administration productivity
Environment: animal, plant, land, air and water conservation, climate change and adaptation, energy efficiency and sustainability	Equality: Accessibility, anti-discrimination, inclusion of the marginalized, detecting bias for improving diversity	Health: improvement of treatment and healthcare services, mental wellbeing
Communication: detection of disinformation, combatting polarization	Infrastructure management: energy, real estate investments, transportation, urban planning, water and waste management	Security, justice, and peace: harm prevention, assistance of police forces and lawyers

See more: Fosso Wamba et al., 2021; Chui et al., 2018



Outline

By Agnes Stenbom

Artificial intelligence (AI) represents a collection of tools and technologies that are transforming operations and outcomes in diverse fields and sectors – from healthcare to transportation, agriculture, and art. AI is increasingly shaping the everyday lives of people across the globe, including the way we work, organize, and interact as human beings. It is an expanding area of research and policy concern. But what is AI?

AI was likely coined as a term at Dartmouth College during the summer of 1956, when mathematicians, computing scientists, and neuroscientists gathered for a workshop testing the hypothesis that “every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it” (McCarthy et al., 2006).

Today, the term represents no single item or technology but should be understood as an umbrella term covering various subfields and technologies. This umbrella – AI – does not have one universal definition but has been described in various ways by different people and in diverse disciplines. The following examples highlight how these definitions can vary:

- UNESCO’s Recommendation on the ethics of artificial intelligence (2021), endorsed by its 193 member states, notes that “AI systems are information-processing technologies that integrate models and algorithms that produce a capacity to learn and to perform cognitive tasks leading to outcomes such as prediction and decision-making in material and virtual environments. AI systems are designed to operate with varying degrees of autonomy by means of knowledge modelling and representation and by exploiting data and calculating correlations” (UNESCO, 2021). The European AI strategy states that “[AI] refers to systems that display intelligent behaviour by analysing their environment and taking action – with some degree of autonomy – to achieve specific goals” (European Commission, 2018).
- In the context of journalism, AI has been described as representing “a collection of ideas, technologies, and techniques that relate to a computer system’s capacity to perform tasks normally requiring human intelligence” (Brennen et al., 2018).

AI technologies have been depicted as *disruptive* because of their potential to change established practices and bring about transformation. Understood in contrast to sustaining technologies, disruptive technologies bring forward new ways of doing things – or even enable the doing of new activities. This does not mean that every AI innovation aims to overthrow the old ways of doing things but rather that the technologies have the potential to substantially alter products and practices.

Narrow and General AI

Within the broader references to AI, some unique technologies and ideals can be differentiated. First, distinguishing between the following concepts is important:

- Applied AI (also known as narrow AI), which represents computer systems that have specific intelligence allowing them to perform specialized tasks.
- General AI (also known as artificial general intelligence – AGI), which represents computer systems that are or are intended to be *generally* intelligent, just like human beings. AGI also relates to the concept of superintelligence, which has been frequently discussed by philosophers and science fiction writers alike.

While there are many cultural portrayals of general AI, including blockbuster films like the *Matrix*, current AI applications in practice are indeed “applied”. Still, it is important to remember that the term *can* represent the ideal of generally intelligent systems for some.

Notably, most contemporary AI use cases call for *hybrid* approaches in which humans are (beyond being AI system creators) needed as operators and/or evaluators. Such hybrid systems leverage the capabilities of algorithms, such as speed and scale, *and* draw on complementary human capabilities to perform a given task (Brynjolfsson & McAfee, 2014).

How various AI definitions resonate with people will be influenced by factors such as disciplinary background, previous technical experiences, and even personal beliefs. Personally, I approach AI as an umbrella term representing computer systems that aim to mimic human intelligence and at times surpass its practical constraints. I share this definition to highlight the perspective with which the paragraphs below are written and to encourage readers to seek clarification when others use the term.

Artificial Is Not Autonomous

When using the term *AI*, especially when referring to applied AI, we must remember that AI is not an autonomous force or concept that will magically appear among us. AI is, just as man-made tools have long been, powered by physical materials and governed by our choices. The systems are what we make them out to be, and they need our physical world to function – not least concerning the vast amount of energy that the systems consume. Hence, we must remember that there are people behind the technologies and that natural resources are also at play in the digital world.

In the 1990s, Herbert A. Simon (1995) noted how the term *artificial* had a negative air about it. While attitudes towards the general concept of an artificially created world may well have changed since then, society still worries about what modern-day artificial systems will entail. We must remember that artificial does not mean autonomous; the artificial part is created by us – humans.

The idea of *intelligence* is central in AI discourse, but just like AI itself, intelligence does not have one universal definition. What is an intelligent system? When should a tool be considered “smart”? Using the term *intelligence* can evoke expectations of non-human objects being human-like, which is not the case for most applied AI today. The influential AI-researchers Stuart J. Russell and Peter Norvig (2010) have said they prefer the term

rational over intelligent, signalling that it is indeed not the humanness – but the optimization of results – that is of interest.

AI is neither good nor bad, but AI is also not neutral. Human values steer AI system development, for example, by setting system goals and generating training data.

A Brief Look at the History of AI

People have long imagined machines with human abilities, including literary illustrations from ancient Greece, with Homer's *Iliad* describing "automotones" (living statues), self-propelled chairs, and other useful machines. Around 1495, close to four centuries before the first computer's creation, Leonardo da Vinci drew his "robotic knight", a humanoid mechanical creature with the ability to sit up, wave its arms, and move its head and jaw.

Fast-forwarding from da Vinci's robotic knight, the birth of what we today call AI has often been attributed to the Turing test in 1950. The test (which was originally called "the imitation game") was introduced by British mathematician Alan Turing and is a method of determining whether a machine demonstrates intelligence. In it, human evaluators observe an interaction between a human and a machine, and if they are unable to tell them apart (i.e. identify the machine as a machine), the machine is said to have passed the test. The practical emergence and study of machine intelligence thus dates back even before the previously mentioned Dartmouth workshop.

Many diverse innovations have come to shape the AI field since the first Turing test – from Joseph Weizenbaum of MIT's creation of one of the first "chatterbots" (now known as chatbots) ELIZA in 1966, to IBM's chess-playing expert system Deep Blue beating the then world-champion Garry Kasparov in a much-anticipated 1997 match.

However, AI has not enjoyed a linear story of technological success. There have been periods where developments have slowed and pessimism has coloured the field. The term **AI winter** has been used to describe periods where research funding is decreased and commercial investments are modest – effectively constituting a chain reaction started by a general sense of pessimism in the AI community.

Key AI Subfields

AI can be understood as its own domain in both research and practice; however, AI is fundamentally interlinked to the computer science and data science domains, among others. In short, this link could be described as computer science and data science constituting the overarching fields upon which the various AI subfields draw to create value through using data and computational methods.

Furthermore, AI is often discussed as being interlinked with the robotics field. While sometimes discussed synonymously, AI and robotics can be but are not necessarily related. Robots do not always leverage AI technologies, and AI technologies do not always need (in fact, they most often do not have) a physical form. AI in the manufacturing sector is, however, closely linked to the robotics domain, where a combination of the two fields aids in, for example, predictive maintenance, supply chain optimization, and quality assurance in the manufacturing of products of diverse sizes and types – from cars to computers.

Below, I outline two of the most fundamental subfields of the broad concept of AI:

- **Machine Learning**

While sometimes used interchangeably with AI, the term machine learning (ML) should be understood as a subfield of AI. ML refers to algorithms trained using examples instead of being coded via explicit instructions. ML can be understood as recipes learned from data; it is an approach to making repeated

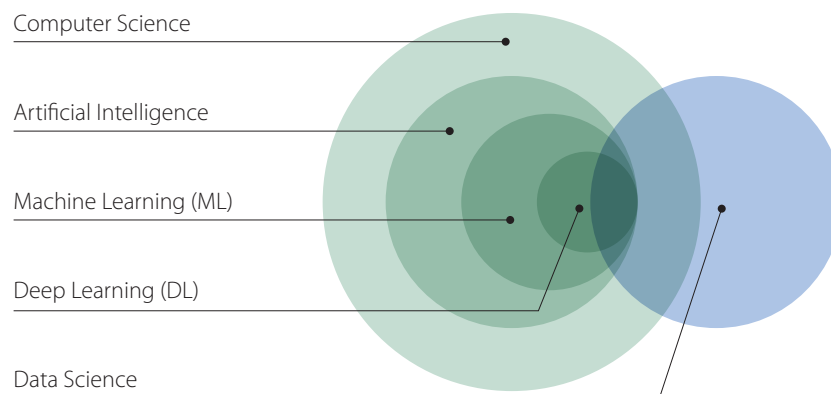
Defining Artificial Intelligence

decisions by algorithmically identifying patterns in historical data and using these to deal correctly (or in a desired way) when exposed to new data.

- **Deep Learning**

Deep learning (DL) is a family of ML methods that are based on multilayer neural networks – hence the term *deep*. Neural networks are computing systems set up to mimic the human brain and the biological neural networks that constitute it. DL approaches allow systems to learn from raw data.

Figure 2: Key AI Subfields



Source: Original material by Authors.

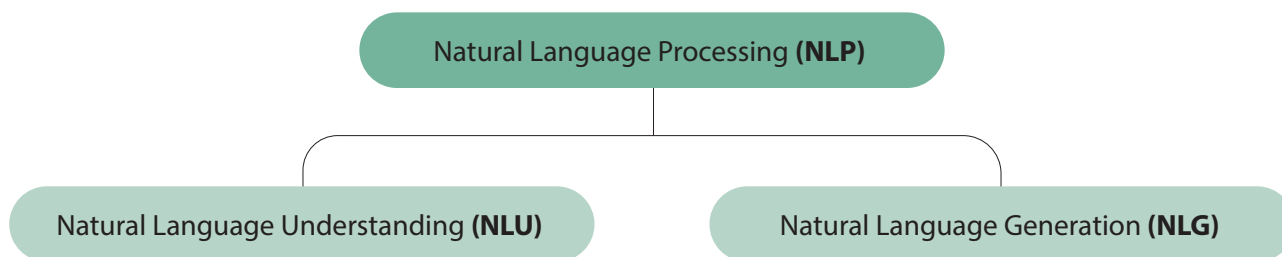
These concepts can be applied in a myriad of ways to find different value-creation opportunities in diverse fields and sectors. How AI is applied in mechanics or manufacturing thus differs greatly from how it is applied in, for example, journalism. Notably, the terms are likely to be interpreted and applied differently depending on a person's field, and their definitions (and importance) change over time.

In the following paragraphs, I outline three AI concepts that should be of particular interest to journalism students and practitioners, including brief examples of how they might be applied in the journalism field.

Natural Language Processing

Natural language processing – often referred to as NLP in short – represents various approaches aimed at analyzing and generating *natural language*. Natural language represents the language that humans use to communicate in both written and verbal forms. They are languages that have evolved through use and repetition and are hence different from constructed languages, such as those used to programme computers.

NLP has two prominent subfields – natural language understanding (NLU) and natural language generation (NLG). As the terms entail, these fields provide various approaches for understanding and generating natural language.

Figure 3: Natural Language Processing (NLP) Hierarchy

Source: Original material by Authors.

NLP can assist in processing vast amounts of texts, translating between languages, and generating new content. Everyday use cases can be found in translation or spell-check programmes (such as the one used to write this very text!), email spam filters, chatbots used in customer service, or virtual assistants like Amazon’s Alexa, Apple’s Siri, Microsoft’s Cortana, and Google’s Google Assistant.

NLP in journalism

A notable example of NLP in journalism is found in the global collaborative investigation by the International Consortium of Investigative Journalists (ICIJ) and 36 media partners, revealing how Isabel dos Santos, the wealthiest woman in Africa and the daughter of Angola’s former president, funnelled hundreds of millions of dollars in public money out of one of the poorest countries on the planet. The investigation – The Luanda Leaks – was based on 715,000 documents provided to the ICIJ by the Platform to Protect Whistleblowers in Africa. The documents included decades of emails, internal memos, contracts, consultant reports, tax returns, private audits, and videos of business meetings. To process them, the ICIJ collaborated with the AI studio at Quartz, which built a system that could “read” all the documents. As phrased by Quartz, this was done to “help journalists from Quartz, ICIJ and other partner organizations find the kinds of documents they expected in the cache of leaks – regardless of file format, spelling, transcription errors, phrasing, or even the language of the document” (Merrill, 2020).

Computer Vision

If NLP is understood as associated with human activities related to language (such as writing, listening, or reading), computer vision should be understood as related to seeing. In computer vision, computer systems are trained to analyze and classify/sort visual content.

A subfield to the subfield (AI is indeed an umbrella term!) is found in face recognition, which focuses on computer vision as applied to human faces. The risks of face recognition being used in invasive and controlling ways – including authoritarian regimes – are both real and relevant. Nonetheless, face recognition is used by millions of people voluntarily every day, as they use their own faces to unlock belongings such as smartphones.

Another everyday use case for computer vision is found in most smartphones, where it aids in classifying images according to their content. If people search for “cat” or “baby” in their image libraries, they will see the technology in practice.

Computer vision in journalism

Computer vision can be applied in various ways in journalism. A notable example of image classification and indexing can be found in media archives, where it helps provide structure to large databases. One example is from *The New York Times*, which used computer vision to create a digital archive of printed papers dating back to 1851. By categorizing the different components of printed articles, they pieced them back together into a structure that mimics the digital-first articles of today (Nordahl et al., 2021).

Like NLP, computer vision can also be used to carry out journalistic investigations. In Argentina, La Nacion used it to analyze election ballots to confirm the suitability of the electoral process and to avoid fraud. They used computer vision to analyze analogue voting ballots, and then validated the output through crowdsourced human reviews. The publisher had a similar project in 2015, in which they manually analyzed 16,000 ballots. By adding computer vision to their 2021 edition, they could review 100,000 ballots (Coelho et al., 2021).

Recommender Systems

Another important AI domain concept is recommender systems. At the most fundamental level, the term has been defined as “describing any system that produces individualized recommendations as output or has the effect of guiding the user in a personalized way to interesting or useful objects in a large space of possible options. Such systems have an obvious appeal in an environment where the amount of on-line information vastly outstrips any individual’s capability to survey it” (Burke, 2002, p. 331).

Modern recommender systems typically leverage machine learning concepts to suggest content, products, or actions of various types. A well-known example is the Amazon recommendation algorithm, which provides users with product suggestions based on, for example, their previous purchases and online behaviours.

In journalism, recommender systems are increasingly employed to personalize and curate news feeds. The design of such recommender systems varies, with public service broadcasters and commercial media optimizing their recommendations for different values, such as sharing public interest news, converting readers into paying subscribers, or maximizing time spent on news sites.

Recommender systems are also fundamental to modern social media, as they help curate relevant and engaging digital experiences that connect people across the world. Many media reports have accounted for the downsides of these systems, and criticism is not uncalled for. Facebook, YouTube, and Twitter have all been at the centre of a broad societal debate about AI-amplification of hate speech, threats, and misinformation, where recommender systems have been accused of sending users down metaphorical rabbit holes of radicalization.

Concluding Remarks

As noted in the introduction, AI is not one thing but a diverse umbrella term and toolbox that provides various opportunities for value creation in different settings and fields. This brief outline has aimed to provide readers with a basic understanding of this diversity and the often-ideological debates that surround it.

AI technologies bring with them vast opportunities but also major new societal and technical challenges. As discussed, AI can be disruptive and thus bring about significant change to both processes and people. Reporting on these changes and the sociotechnical effects they have is an important task for contemporary journalism.

Definitions of AI

"Artificial intelligence refers to machines or computer systems capable of learning to perform tasks that normally require human intelligence." (Bawack et al. 2019.)

"The field of artificial intelligence strives to understand and build intelligent entities." (Russell and Norvig 2010.)

"A system's ability to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation." (Haenlein and Kaplan 2019.)

"AI systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions." (Samoili et al., 2020.)

"Artificial Intelligence (AI) refers to any technology that enables machines to operate emulating human capabilities to sense, comprehend and act. AI is intelligence demonstrated by machines, unlike the natural intelligence displayed by humans and animals, which involves consciousness and emotionality. ... AI is general-purpose technology that can affect an entire economy like steam engine, electricity, or computer internet." AI for Africa blueprint (Sedola, 2021)



Suggested Assignments



READ AND REFLECT: Go to the leading digital news outlet in your country, use "Artificial intelligence" as a search phrase, and select one article that interests you. Analyse the article according to the categories below. Reflect what kind of a picture the article delivers on AI to its audience, and with what kind of consequences. How does the article succeed in delivering a lucid, understandable picture of AI?

Concepts AI	Application	Vocabulary	Style	Sources	Implications
How is AI defined? What kind of AI is in question?	What area of application does the article deal with?	What expressions and phrasings are used?	What is the intended audience?	What are the sources? What is lacking?	How is the article potentially interpreted? What kinds of effects may it have?



DISCUSS: What kind of general conceptions and misconceptions do people have about AI in your opinion? List your ideas and share them with a peer. Discuss where these ideas possibly come from, to which extent they are supported by public discourse and journalism, and in which topics or aspects people should receive more education.



IDEATE: Choose one of the following concepts of AI: a) green AI; b) friendly AI; c) game AI; d) educational AI; e) medical AI; f) AI in opinion mining; g) AI in law; h) AI surveillance; i) AI in music. Search for research articles addressing the topic to find out how it has been studied and touched

Defining Artificial Intelligence

upon so far. Come up with five different journalistically developed story ideas regarding the selected concept.



SEARCH AND SOURCE: Discover in historical articles by searching newspaper archives (typically available at public and university libraries) how the following inventions were initially reported when they were new: a) automobile (1886), b) smartphone (1997), c) Twitter (2006). What kind of expectations were endorsed and which of them turned out to be exaggerated and false, which turned out to be true?



PRODUCE: How to report on AI in an interesting but realistic way? Create a presentation on the basics of AI for one of the selected target groups: a) pre-school children, b) upper-secondary school/high school teachers, c) senior citizens. Select a video addressing the basics on AI and to be embedded into your article. Trying to employ the video's content, please cover the following questions in your presentation: What is AI and how does it matter for the target group? How is it manifested in the everyday life of the target group? What kind of ethical questions are relevant to the target group? Your presentation can be a blog post, an online article or a fact sheet, according to the publication channels, needs and preferences in your classroom. When giving feedback to each other's presentations, discuss how to avoid becoming an advocate for AI but remain critical and. In addition, how to problematize the issues in an appropriate way without being too alarmist?



Readings

- Beynon-Davies, P. (2022). *Data and society*. World Scientific Publishing. <https://doi.org/10.1142/12287>
- Boden, M. A. (2016). *AI: Its nature and future*. Oxford University Press.
- Crawford, K. (2021). *The atlas of AI: Power, politics, and planetary costs of artificial intelligence*. Yale University Press.
- David, M. (2005). *Science in society*. Palgrave Macmillan.
- Elliott, A. (Ed.) (2021). *The Routledge social science handbook on AI*. Routledge. <https://doi.org/10.4324/9780429198533>
- Ertel, W. (2018). *Introduction to artificial intelligence*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-40022-8>
- Frankish, K., & Ramsey, W. M. (Eds.) (2014). *Cambridge handbook of artificial intelligence*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139046855>
- Mainzer, K. (2019). *Artificial intelligence: When do machines take over?* Springer Verlag. <https://doi.org/10.1007/978-3-662-59717-0>
- Mitchell, M. (2020). *Artificial intelligence: A guide for thinking humans*. Penguin Books.
- Mueller, J. P. & Massaron, L. (2018). *Artificial intelligence for dummies*. John Wiley & Sons.
- Neapolitan, R. E. & Jiang, X. (2018). *Artificial intelligence: With an introduction to machine learning*. 2nd ed. CRC Press. <https://doi.org/10.1201/b22400>
- Poole, D. L. & Mackworth, A.K. (2017). *Artificial intelligence: Foundations of computational agents*. 2nd ed. Cambridge University Press.
- Samoli, S., Lopez Cobo, M., Gomez Gutierrez, E., De Prato, G., Martinez-Plumed, F. & Delipetrev, B. (2020). *AI WATCH: Defining Artificial Intelligence*. Publications Office of the European Union.



Module 2:

Cultural Myths and Narratives about Artificial Intelligence



Summary

This module focuses on the central role of cultural myths and common narratives when it comes to understanding AI in society. The chapter takes a look at the expectations humans have formed surrounding the technology, and the relationships we have forged with AI systems increasingly being deployed in our workplaces, schools, entertainment venues, and homes. The cultural narratives surrounding AI affect the public discourse and thus journalistic coverage. The chapter concludes on the necessity of reading the algorithm, developing a toolkit for journalists and other informed members of society to use as a lens to analyze AI in our daily lives.

Key concepts: *cultural myths, science fiction, speculative futures, rationality*



Module Aims

The aims of this module include the following:

- ▶ Understanding the founding myths and common rules that have been formed around artificial intelligence.
- ▶ Increasing the learners' understandings to critically analyze the relationship of AI to both ourselves and society.
- ▶ Increasing the learners' knowledge on algorithmic literacy and develop tools for holistically examining AI systems.



Learning Outcomes

After the completed module, the student is expected to have learned the following:

- ▶ To understand how the public imaginary and discourses have affected and continue to affect the common understandings of AI.
- ▶ To be able to identify some pervasive narratives of the discourse related to AI and think of alternatives.
- ▶ To critically analyze and "read" an algorithm leveraging both technical and cultural knowledge and tools.



Questions for Reflection

Here are some questions to reflect upon in advance:

1. What kind of AI was described in your childhood television programmes and films? Try to recall as many examples as you can, to share with your peers. When in the classroom, discuss: Did you recall the same examples? How many of these stories appear to have been realized in modern day technology?
2. Download a modern dating app such as Bumble, Hinge, Coffee Meets Bagel (or borrow a friend's if you don't want to use your own phone). Notice how the choice of your profile and your picture affects the "matches" you are presented with. Try to change key characteristics of your profile (e.g. education level, hobbies, pictures used) and see if you can spot any patterns with the recommendation algorithm. Can you describe, in your own words, how the "matching" algorithm of the AI works? Alternatively, you can try to recall what kind of patterns there have been in your Instagram feed or in another app with personalized content. How, and on what basis, has the algorithm tried to anticipate your interests?
3. Pick your favourite AI film situation (e.g. the robots become sentient in *Terminator*, the hackers uncover a plot of the machines in the *Matrix*, etc.). Write a quick news blurb in the "inverted pyramid" or bottom line up front (BLUF) style of communicating information as if this was the news of the day to be reported on.

Here are some preparatory tasks to do prior to entering the topic:

1. Watch or get familiar through second-hand sources with one of the following classic films and describe how they have contributed to differing cultural myths and narratives about AI:
 - a) *The Matrix trilogy* (1999–2003 by Lana and Lilly Wachowski: *The Matrix* 1999; *The Matrix Reloaded* 2003 and *The Matrix Revolutions* 2003) as well as *The Matrix: Resurrections* (2021 by Lana Wachowski),

- b) *A.I. Artificial Intelligence* (2011 by Steven Spielberg),
- c) *Blade Runner 2049* (2017 by Denis Villeneuve; see also the 1982 film by Ridley Scott, based on the novel *Do Androids Dream of Electric Sheep?* by Philip K. Dick),
- d) *Independence Day* (1996 by Roland Emmerich; see also *Independence Day: Resurgence* [2016] by the same director),
- e) *The Hitchhiker's Guide to the Galaxy* (2005 by Garth Jennings, based on a novel by Douglas Adams),
- f) *2001: A Space Odyssey* (1968 by Stanley Kubrick, based on the short story *The Sentinel* [1951] by Arthur C. Clarke),
- g) *Je Suis Auto* (2020 by Johannes Grenzfurthner and Juliana Neuhuber),
- h) *Coded Bias* (2020 by Shalini Kantayya),
- i) *2067* (2020, Seth Larney),
- j) *2.0* (by S. Shankar),
- k) *The Invisible Boy* (1957 by Herman Hoffman),
- l) *Metropolis* (1927 by Fritz Lang).

2. Works of popular culture often do not only describe technologies but also teach us something essential about them, or, more specifically, about the relation between machines and humans. What are the “lessons learned” of technologies in the following works of popular culture? Categorize the films according to whether they present utopias or dystopias, and in what ways. Please choose one genre and take a look at their notions of AI based on plot descriptions available, for example, on the Internet Movie Database (IMDb):

- 1) Sci-fi: a) *Archive* (2020 by Gavin Rothery), b) *Upgrade* (2018 by Leigh Whannell), c) *Zoe* (2018 by Drake Doremus), d) *Morgan* (2016 by Luke Scott), e) *Chappie* (2015 by Neill Blomkamp), f) *Transcendence* (2014 by Wally Pfister), g) *Autómata* (2014 by Gabe Ibáñez), h) *Ex Machina* (2014 by Alex Garland), i) *WALL-E* (2008 by Andrew Stanton), j) *I, robot* (2004 by Alex Proyas)
- 2) Comedy: a) *The Mitchells vs the machines* (2021 by Michael Rianda and Jeff Rowe), b) *Free guy* (2021 by Shawn Levy), c) *Ron's gone wrong* (2021 by Jean-Philippe Vine and Sarah Smith), d) *Meet the Robinsons* (2007 by Stephen Anderson), e) *Maniac* (2018 by Cary Joji Fukunaga and Patrick Somerville), f) *Austin Powers: International man of mystery* (1997 by Jay Roach)
- 3) Action and drama: a) *Tau* (2018 by Federico D'Alessandro), b) *Extinction* (2018 by Ben Young), c) *Ghost in the shell* (2017 by Rupert Sanders) or *Kōkaku kidōtai* (1995 by Mamoru Oshii), d) *Avengers: Age of Ultron* (2015 by Joss Whedon), e) *Tron: Legacy* (2010 by Joseph Kosinski), f) *Johnny Mnemonic* (1995 by Robert Longo), g) *RoboCop* (1987 by Paul Verhoeven), h) *WarGames* (1983 by John Badham)
- 4) Romance, family and lifestyle: a) *Superintelligence* (2020 by Ben Falcone), b) *Smart house* (1999 by LeVar Burton), c) *Jexi* (2019 by Jon Lucas and Scott Moore), d) *I am mother* (2019 by Grant Sputore), e) *Tomorrowland* (2015 by Brad Bird), f) *Her* (2013 by Spike Jonze), g) *Pixel Perfect* (2004 by Mark A.Z. Dippé), h) *Simone* (2002 by Andrew Niccol), i) *New Rose Hotel* (1998 by Abel Ferrara)



Outline

By Edward Finn and Suren Jayasuriya

Myths about artificial intelligence long predate the emergence of modern computers in the mid-20th century. A number of cultural traditions include stories about automatons or other beings that were designed and brought to life by humans, including the stories of Pygmalion and Hephaestus's servants from Greek mythology and the automatons of Indian sacred texts. These stories are still with us, reinvented and updated for the modern age. One of the most famous modern iterations is the Frankenstein myth, where a rogue medical student uses his scientific knowledge to create an artificial being and then fails to take responsibility for his creation. These narratives all build on a core set of questions that are central to the human condition: what responsibilities do we have as creators, parents, and makers in this world? How do we define humanity and personhood? What is the meaning of life, and what are the consequences of creating new life?

As technological progress continues to accelerate, the idea of "playing god" has gone from thought experiment to reality. Humanity has now played a major role in altering nearly every biological system on the planet, and we are creating and destroying life forms at a breathtaking clip. This is the context in which contemporary myths about AI circulate. We tend to reiterate the same few narratives with certain variations and crossovers. You can find at least one of the following archetypal narratives embedded in just about every discussion we have about AI today, and certainly in the stories we tell about the future of AI.

The robot apocalypse. Our greatest collective fear is that we will create AI that will replace humanity. Thinking machines that are more intelligent and more capable than humans might quickly surpass humanity, as envisioned by narratives like the *Terminator* film series, *I, Robot*, and many more. The question of AI as an existential threat to humanity preoccupies Western philosophers as well as filmmakers, including notable authors like Nick Bostrom and Max Tegmark. These stories also play on humanity's deep-seated anxieties around power, (eg. master and slaves, such as Karel Capek's famous play *R.U.R.*, which coined the word "robot" as a play on the word for serf. If we design AI as slaves to humanity, is it inevitable that these slaves will rise up against their masters?

The robot girlfriend. If AI does not overthrow us, it may seduce us (notably, men) instead. If we create sufficiently advanced AI, we may find these new beings irresistible: more perfect than mere humans. In some versions of this story, like the series *Westworld* or the *Blade Runner* films, the anxiety is that AI might "pass" as human, deceptively blurring or breaking the boundary between human and non-human. In others, AI may be seductive even without being fully human, such as in the Spike Jonze film *Her*. Unlike the ancient myths of Pygmalion or Adam and Eve, in Jonze's narrative the love story goes awry, and humans are rejected by AI because we are simply too boring for such advanced intelligences. More recent fiction has explored positive human-AI relationships, such as Annalee Newitz's novel *Autonomous*.

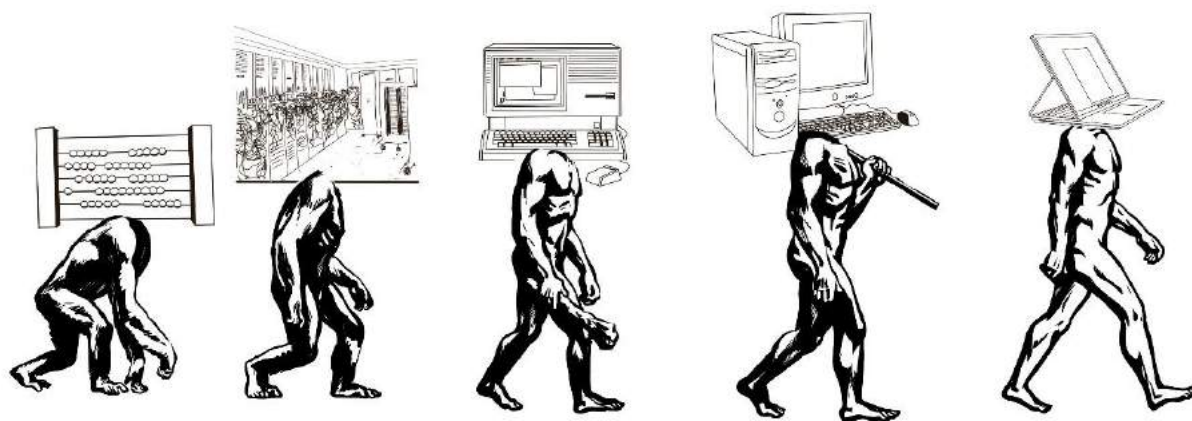
The God machine. This brings us to the final archetype of AI story, the super-intelligence. Many storytellers have explored the idea of a superhuman AI that eclipses our understanding; some people speculate that we will soon face a moment when such acceleration in AI capability becomes inevitable, which they term "the Singularity." In the *Matrix* films, this godlike machine imprisons humanity in a simulation of reality, and a school of philosophical debate argues that we may indeed be living in such a simulated world. In other narratives, like Douglas Adams' *The Hitchhiker's Guide to the Galaxy* and Iain Banks' culture novels, these machines co-exist with humanity, providing leadership, stewardship, and oracular guidance.

Notable in much of this science fiction and speculative fiction surrounding the above common narratives is the prevalence of Western male writers. This skew reflects and reiterates broader cultural biases that made it more difficult or impossible for women and members of other groups to write and publish their work; current

AI technology trends similarly reflect and amplify bias against women and other groups (Collet et al., 2022). However, Mary Shelley should be credited for imagining Frankenstein as a gendered human artifact, while Thea von Harbou wrote *Metropolis* (later made into a film by her partner Fritz Lang). Visions of AI from authors like Marge Piercy, Jeanette Winterson and Anne Leckie bring key feminist perspectives on gender inequality and sexuality to these futures but are less widely read than canonical male writers. Japan has developed its own rich imaginary around robots and AI, starting with post-war works like *Tetsuwan-Atomu* (more widely known as *Astro Boy*) and more recent works like the globally influential *Kōkaku Kidōtai* (*Ghost in the Shell*).

Many cultures share mythological narratives of androids and automatons, notably including the Ramayana and Mahabharata epics (Mayor, 2018). The politics of representation continue to be haunt real-world implementations of AI, which still frequently default to submissive female avatars and personae. Since so many AI narratives explore the theme of slavery, the question of race and AI is also important. African-American artists have reclaimed and reinvented these themes in works like *Metropolis* (a concept album by Janelle Monáe, 2007) and *Frankenstein (Destroyed)*, a comic book series by Victor LaValle). Elsewhere in science fiction the films *District 9* and *Black Panther* present different takes on “whiteness,” and there is growing interest in Afrofuturism and Africanfuturism as alternative, non-eurocentric models for imagining positive futures. Nevertheless, it is not insignificant that dominant discourse continues to draw on white male writers, who tend to be read by white male technologists, creating a sometimes blinkered feedback loop that has significantly influenced the development of real-world technologies, from the *Star Trek* computer which directly influenced Google’s goals for its search engine to Neal Stephenson’s vision of the metaverse, which continues to inspire a range of virtual reality technologies that intersect with AI (Finn, 2017b).

Figure 4: Evolution of HAL



Source: Original design by Nina Miller.

“Rules” of AI

Along with the myths and tropes of AI introduced in the previous section, there is an expectation of rationality and logical behaviour that underlies stories of human interaction with intelligent machines. This expectation manifests itself in intricate webs of rules, norms, and conventions that govern how we discuss, delineate, and even imagine artificial intelligence in society. The “rules” of AI typically touch upon important philosophical questions surrounding the morality and ethics of human-machine interaction as well as the nature and

characteristics of intelligence. Paralleling the development of AI systems, these societal and cultural rules started out with more explicit, logically-derived conditions to govern intelligent behaviour and its ethics, but soon shifted to being distributed and emergent with unexpected consequences and grey moral quandaries. Humans are story-telling animals, and our stories of AI reflect the tension between defining rules and norms to constrain the chaotic potential of machine intelligence within a rational and ethical framework while still dreaming and envisioning new possibilities that this technology fuels.

The science fiction writer Isaac Asimov in the 1940s and 1950s captured the early zeitgeist for determining a rationality or logic behind intelligent robots. His three Laws of Robotics were formulated in his science fiction as explicitly programmed into robots which protect or shield humans who interact with them.

First Law	A robot may not injure a human being or, through inaction, allow a human being to come to harm.
Second Law	A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
Third Law	A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

Asimov's Laws recur often in his writings, and many stories feature robots with programming that has altered or modified one or more of these laws leading to new, exciting, and dangerous behaviour. For example in "Little Lost Robot", the robopsychologist Dr. Susan Calvin notes that removing the "inaction clause" of the First Law could lead to a robot dropping a heavy weight on top of a human below it with the intention of catching the weight, but the second the weight is released also absolves the robot of any imperative to catch the weight and save the human.

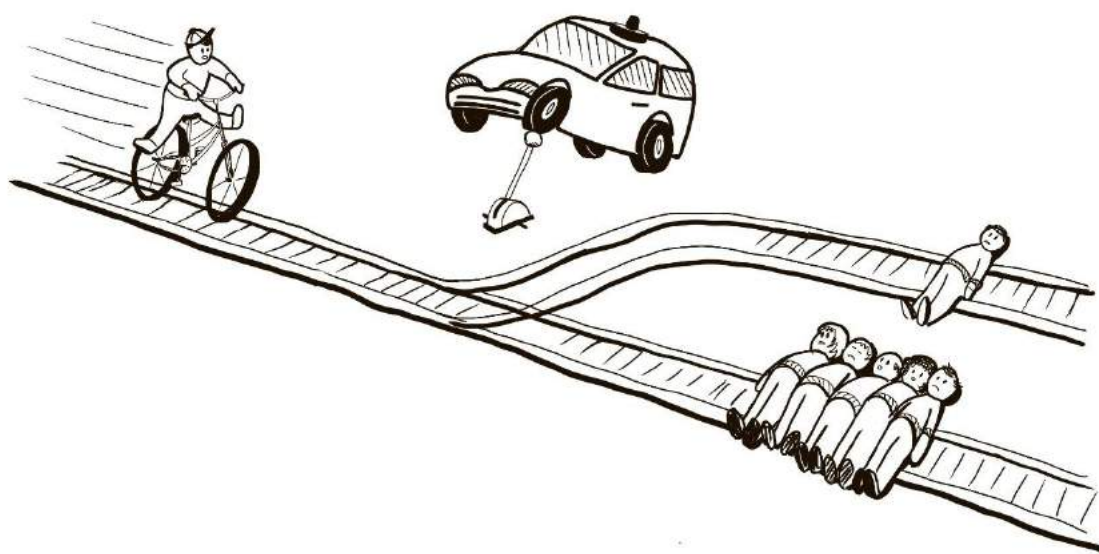
These "Laws" provide insight into our need to safeguard ourselves from our own Frankensteinian machines, and the desire to prevent harmful consequences through blind robotic obedience to this set of edicts. At the same time, these rules are very limited and do not prescribe an extensive list of norms, codes and regulations for the robots—in contrast to the myriad bizarre limitations placed on human beings (e.g. Asimov never wrote Law 4021 governing how robots should file their own taxes!). His stories have captured our imagination through an understanding our desire to find a rational basis to safely design intelligent machines.

Our fascination with rules and AI also extends into encoding ethical and moral behaviour into our machines and algorithms, and reflect our own difficulties with prescribing normative statements for ethics. In the debate over self-driving cars, a common philosophical thought experiment is the trolley problem. The premise of the trolley problem is a forced choice: an observer sees a runaway train headed down a track towards a group of people (typically five) tied down and in its path, and can flip a switch to steer the trolley down an alternative track which will then kill one person. While the majority of people polled would flip the switch, multiple variations of trolley problem have fascinated philosophers and psychologists including when the person on the alternative track is a baby or your grandmother (Note: for deliciously satirical takes on the problem, pick up a copy of Thomas Catcart's *The Trolley Problem, or, Would You Throw this Fat Guy Off a Bridge*). For AI technology, the trolley problem has been posited (perhaps overzealously and over-prominently) as a classic ethical dilemma for how we should design our self-driving and other autonomous vehicles and as a litmus test for whether we would be comfortable with a robot making such decisions.

As these examples indicate, AI is a new arena in which to grapple with ethical dilemmas we have never adequately or consistently resolved in society at large. In our stories and culture, the perceived power of AI has always been entrained by rules and norms governing its behaviour. These rules satisfy our psychological need for safety, particularly when confronted with the alienness of intelligent technology, but also provide

rich, imaginative content when we visualize the cracks, deviations and loopholes which can challenge our ethical and moral boundaries. Aristotelian philosophy classically posited humans are rational animals, and we have come to expect intelligent machines to follow a similar, if not warped, set of logic and reason-obeying behaviour. However this premise may be antiquated as we increasingly design autonomous systems that interact with human society and culture, (e.g. chatbots, news filters, self-driving cars), we are quickly finding the boundaries and failure points of rational behaviour and ethical consistency. As AI technology proliferates throughout society, these rules and expectations will change and morph to reflect the ever-changing nature of how AI can touch our lives.

Figure 5: The Trolley Problem



Source: Original design by Nina Miller.

AI and the Self

Stories about AI collaborating with humans often explore the boundaries of self and identity. One rich tradition of such narratives is the cyborg, or the hybrid of human and machine. The enhancement and extension of the human body with exoskeletons and new sensory organs has been spectacularly portrayed in movies like *Ironman* and *Avatar*, but these stories build on more quotidian modifications such as eyeglasses and artificial limbs. A second, more open-ended, thread is the role of computation in what some philosophers call the “extended mind.” Humans have always relied on technologies to help us think, from pencil and paper to language itself. As we do more and more of our thinking in the company of computational machines (how many such devices are you touching or in close proximity to as you read this?), the extended mind is quickly becoming a collaboration with algorithms and autonomous systems that filter, remember and recommend, operating in a continuous feedback loop with the operations of our organic brains. This kind of AI-supported collective or enhanced thinking has been explored in a variety of science fiction, such as Anne Leckie’s *Ancillary Justice* series and Arkady Martine’s *Teixcalaan* novels, which both imagine thousands of individual humans collectively sharing experiences through AI platforms. In ways large and small, this kind of collective cognition is already here, through the filters, suggestions, and feedback loops that nudge us to engage and consume based on the cognitive choices of millions of strangers.

The ubiquity of such computational influences on contemporary cultural life for most digitally connected people poses a broader question: if we are all influenced by filters and algorithms, what are the consequences for human imagination? Art and culture depend on serendipity, and serendipity on demand is now mass-produced by algorithms. What would AI imagination look like? Several AI stories have explored this question: the most cynical or suspicious responses, like Stanley Kubrick's masterpiece *2001: A Space Odyssey*, argue that AI imagination would be both instrumental and diabolically inhuman. In that story the AI HAL 9000 infamously tries to murder the human crew of its spaceship, deeming them superfluous to its scientific mission. More optimistic portrayals, like the android crewmember Data in the TV series *Star Trek: The Next Generation*, suggest that AI imagination might intersect in surprising ways with our own. Another approach might be to ask how AI could harness imagination into creative pursuits like art and music. New technologies are continuing to trouble this boundary: tools like DALL-E and Imagen can create breathtaking works of art (from abstract painting to photorealistic images) based on short verbal prompts, and human creative processes are increasingly dependent on various AI technologies, from auto-tune for musicians and image processing for photographers to the search engines and social media platforms that many artists mine for inspiration (Finn, 2017a).

In keeping with the underlying questions that drive our mythos around AI, stories about intelligent machines inevitably turn back to how this changes identity and a sense of self. We have already outsourced much of our memory and relational cognition work to machines, from the everyday (phone numbers, birthdays, maps) to the intimate and momentous (dating, investing, job-hunting). The question is, are we collaborating or competing with these systems? The scene in the film *Elysium* where the protagonist attempts to convince an AI police officer of his innocence offers a dystopian view of humans battling an alien AI imaginary. But these narratives are like the classic American folk story of John Henry, the railroad worker who competes against a steam engine and wins, only to die of over-exertion: competing against AI on its own terms will always lead to defeat. Instead, we must find ways to collaborate, and many AI narratives imagine optimistic futures in which we effectively team up with AI. From the classic TV show *KITT* to helpful side characters in movies and games like *Interstellar* and *Halo*, humans can flourish if we find ways to govern and work with AI rather than against it.

AI and Society

It is human nature to seek the extremes or boundaries of possibility: utopia and dystopia, disaster or total victory. The promoters of AI have always over-promised in their enthusiasm and ambition, while the detractors have always feared the very worst. We can draw a few important lessons from the mythological foundations of AI in culture.

First, the stories that we tell about AI inform the real technological systems that we design and build. All AI systems are ultimately envisioned, planned and supported by humans who embed their assumptions and narratives into the black boxes and machines they build. Companies designing new tools and platforms for virtual reality often hand out copies of science fiction novels to new hires, and thousands of engineers have cited Asimov's three laws in their research papers. So if we want to influence the future development of AI, we should start telling stories about AI societies we might actually want to live in. Cautionary tales are important, and we will never relinquish them, but we also need to imagine, and debate, what equitable, empowering, and inclusive AI might look like.

Second, we need to embrace the fact, as so many science fiction writers have shown, that AI represents a unique opportunity to explicitly design and experiment with ethical systems. For the first time in human history, we are building tools that autonomously observe, react, and interact with us. If we so choose, we can govern the design of these tools not just to be stronger or faster or smarter than us, but to embody our highest aspirations

for wisdom, empathy, and justice. Such AIs would inevitably be imperfect, just as we are imperfect, but better than the alternative.

Reading the Algorithm

The cultural imaginary of AI plays an important role in helping us contend with the real-world consequences of autonomous, intelligent systems. Classic narratives like HAL and Asimov's Three Laws continue to colour our interpretations of AI, providing a vocabulary and influencing our emotional responses to thinking machines. However, mythic stories about AI rarely capture the complexities or real limitations of the machines we actually use every day. True algorithmic literacy requires a different kind of toolkit that combines cultural knowledge and helpful metaphors with a basic understanding of both narrative and the technical abilities a new system might present. This is critical for journalism where future stories will heavily depend on an informed reporter who can convey this information to the general public. For the journalist wishing to concretize their analysis of AI algorithms in modern society, common points or features for characterizing an algorithm are presented below and summarized in Table 1.

User Experience and Privacy. First, it is important to understand the relationship between a user and a system. This forms the basis of computer science research into human-computer interaction and will be increasingly important as AI tools collaborate with human users for complex tasks. As dystopian narratives like *Tron* have taught us, poorly designed AI can dehumanize and disempower the people it is meant to serve. As our myths warn us, appearances can sometimes be deceiving, particularly when a relationship is presented as "get something for nothing." Common examples of prototypical user-AI experiences that are worth studying in detail include Amazon's Alexa for smart homes, the Netflix recommendation system for movies and TV shows, and targeted advertisements on a platform like Facebook. Understanding every interaction with an AI system as a kind of trade or two-way conversation can be a helpful starting place: most systems, especially those that offer free services, are also collecting personal data or observing our actions in other ways that may not be immediately obvious or explicit. Developing some understanding of the real terms of our relationships with intelligent systems is an important step in building transparent and mutually beneficial collaborations.

Infrastructure and System Design. AI is forged through technical feats of engineering: data, algorithm design, and hardware and software are welded together for the final system. A critical look at these parts will yield valuable information for the observant eye. For instance, AI algorithms depending on machine learning utilize large datasets with implicit bias or discrimination encoded into their data samples and the statistical correlations that are drawn from them. The choice of classical versus deep learning algorithms for an AI can have downstream effects on the performance, explainability and interpretability of the system. As technology advances, both newer sensors and software will replace their respective counterparts in legacy systems, enabling unseen capabilities to sense and interact with the outside world.

Culture Machines and Machine Ethics. A clear understanding of the terms of a human-AI relationship also offers some insight into "what an algorithm wants"-how the individual user fits into a broader system of people, processes, and organizations that may all be interacting with a computational system according to different cultural frames or narratives. A child's relationship with an Amazon Alexa device may be very different from an adult's relationship, which again might be radically different from how an Amazon engineer or a third-party data broker perceives the same system. Further, the ethical dilemmas that accompany the building/design and deployment of AI systems require a holistic, systems viewpoint rather than the traditional normative or prescriptive-based ethical tradition. AI systems now contribute large amounts of carbon emissions in their training of neural networks, and the ethical footprints of these systems can exacerbate existing systemic

Cultural Myths and Narratives about Artificial Intelligence

inequalities and marginalization of minorities. Any insights should inform a grounded understanding of an AI system as both a technical and a cultural entity.

Speculative Futures. Science fiction teaches us that AI systems can “be” and “do” many things, just like people are infinitely diverse in their interests and capacities. Having a richly informed frame of reference for AI gives us the tools to understand intelligent systems as complex objects that we can interpret and interact with. “Reading” AI in this way is a skill we have all begun to develop, practicing with science fiction and then deploying our skills in the real world. Understanding that AI systems in reality are telling stories and performing roles about what they are and what they can do is a way to bridge the space between imagined and real AI systems, and in fact to recognize that our shared imaginary of AI plays a significant role in determining what these systems really can do.

Table 1: Reading the Algorithm Toolkit

Point of analysis	Characteristics	Examples/Evidence
User Experience	<ul style="list-style-type: none"> • User interface design (e.g. text-based, visual, VR, haptic) • How does the user interact with the algorithm • Data collection procedures 	<p>Amazon’s Alexa, Apple’s Siri, and other digital assistants</p> <p>Netflix recommendations</p>
Infrastructure and System Design	<ul style="list-style-type: none"> • Sources of data for machine learning/training • Basic signal/information flow for the algorithm • Learning-based or classical AI algorithm design • Hardware and software requirements (e.g. uses sensors, online data/resources, power needs) 	<p>Block diagram/flow chart for algorithms</p> <p>Data sheets and specifications for the system</p> <p>Algorithm taxonomy and classification</p>
Cultural Machines and Machine Ethics	<ul style="list-style-type: none"> • Stories/narratives of the algorithm (linked to cultures and regional hegemonies) • Stakeholders/companies that leverage the algorithm • User demographics 	<p>News articles</p> <p>Social media</p> <p>Companies involved in the development of AI and its dissemination</p>
Speculative Futures	<ul style="list-style-type: none"> • Imaginative futures for AI technology • Literary and poetic visions for AI 	<p>Science fiction books and literature</p>

Art for Art's Sake: The Uses of AI in Arts



Bernhard J. Dotzler, Professor, University of Regensburg, Germany, the co-editor of the book *Götzendämmerung – Kunst und Künstliche Intelligenz* (transcript, 2021, together with Berkan Karpac)

Computers have been used to produce art – poetry, music, drawings, paintings – from the very beginning. Exhibitions like *Cybernetic Serendipity: The Computer and the Arts* (London 1968), *Tendencije 4: Computers and Visual Research* (Zagreb 1968/69), and *Software – Information Technology: Its New Meaning for Art* (New York 1970) give proof to that. As early as 1960, Denis Gabor worried whether the machine would ‘cut out the creative artist’. He went on to answer his question himself: ‘My answer is that I sincerely hope that machines will never replace the creative artist, but in good conscience I cannot say that they never could.’

The computer, so far, has been used as a tool for making art. The same goes for AI. Up to now, so-called AI-generated art exists only by way of ‘collaboration between human and machines,’ as the art collective Obvious puts it. Obvious succeeded in selling the painting entitled *Edmond de Belamy* at Christie’s for \$432,500: ‘the first portrait generated by an algorithm to come up for auction’ (Christie’s) in 2018. One also might think of Sougwen Chung’s painterly performances with robots: *Drawing Operations* (2018), *Artefacts* (2019), or *F.R.A.N.* (2020), as well as of Patrick Tresset’s *Human Studies* installations (since 2011), implementing drawing robots as alter egos of the artist.

Strictly speaking, however, we can say that there is AI-generated art only if – and when – its creation is not solely based on a collaboration with AI, but when the replacement of the (human) artist has happened completely. In this respect, such art teaches the same lesson about AI as any other field of application (e.g., the driverless car): AI can make artist to redundant.

But the other way round, is there any lesson to be learned? A lesson that AI teaches about art? In order to present itself as art, every work of art has to have at least a touch of *l’art pour l’art* (‘art for art’s sake’). Once truly AI-generated art will have become state-of-the-art it will be ‘art’ (meaning technology by its etymology: the Latin *ars* as the translation of the Greek *techne*) that produces ‘art’ (meaning works of art). Thus AI-generated art will be *l’art pour l’art* in the strictest sense, verifying that both, the artist (human or artificial) as well as the work of art, are only, to quote Martin Heidegger, ‘on account of a third thing, which is prior to both; on account, that is, of that from which both artist and artwork take their names, on account of art.’

Using AI to imagine art about cats playing piano in the style of well-known artists.



Source: <https://ev.medium.com/cat-playing-piano-in-the-style-of-b5f1191dd1c2>



Suggested Assignments



READ AND REFLECT: Choose one of the films mentioned above in pre-assignments (question 2, preparatory tasks) and discuss how it reflects the archetypal narratives mentioned in the outline.



DISCUSS: Discuss what do the following concepts of science fiction mean, how they have they been manifest in the public imaginaries and thus affected our understandings of AI.



IDEATE: Envision the development of an AI-infused object. Create a timeline indicating the milestones of technological, economic and social development of the object. Suggest some potential directions for future as well. Choose one of the following objects or come up with an alternative of your own: a) telephone; b) car; c) child's toy d) eyeglasses; e) restaurant; f) voting machine; g) language translation device.



SEARCH AND SOURCE: What is (the state of current) "computer art" or "digital art"? Try to find recent examples and/or interview a computer/digital artist to write a feature article of the aesthetic phenomenon. What does computer- and algorithm-based art teach us?



PRODUCE: Write a column with a critical approach to presentations of AI and human-machine interaction based on a specific type of cultural product: box office films, comics, or a specific genre (science fiction, cyberpunk, manga, anime). Choose one cultural product or an entire genre. You can also make use of the films and film genres indicated in pre-assignments (question 2, preparatory tasks). Reflect upon the ways how this specific cultural product or genre has affected the ways in which we think of technologies and human-computer interaction. What kind of biases or blind spots are there, for example, in terms of diversity of perspectives or gender?



Readings

Barrat, J. (2015). *Our final invention: Artificial intelligence and the end of the human era*. St. Martin's Publishing.

Bloomfield, B. P. (2003). Narrating the future of intelligent machines: The role of science fiction in technological anticipation. In: Czarniawska, B., & Gagliardi, P. (eds.) *Narratives we organize by*. John Benjamins Publishing Company, 174–212. <https://doi.org/10.1075/aios.11.16blo>

Gibson, R. (2020). *Desire in the age of robots and AI: An investigation in science fiction and fact*. Springer.

Nilsson, N. J. (2010). *The quest for artificial intelligence: A history of ideas and achievements*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511819346>

Pilkington, A. G. (2017). *Science fiction and futurism: Their terms and ideas*. McFarland & Company.

Thompson, S. J. (Ed.) (2018). *Androids, cyborgs, and robots in contemporary culture and society*. IGI Global.

Yodkowsky, E. (2015). *Rationality: From AI to zombies*. Machine Intelligence Research Institute.



Module 3:

Policy Frameworks and Recommendations for Artificial Intelligence



Summary

This module focuses on international policy frameworks supporting ethically sustainable use of the development of AI. The module traces the most common characteristics identified in major policy frameworks and outlines a set of criteria for fair, trustworthy and responsible uses of AI in democratic societies. It discusses how AI policies intend to work towards equality and inclusion, harm prevention, and crisis response. Furthermore, it discusses journalism and journalists' relationship to policy work and how to cover it by finding appropriate source practices. The module examines how to concretize abstract ideas, approach expert sources related to various aspects of AI, and how to report.

Key concepts: *policy, ethics, strategy, regulation, responsible AI*



Module Aims

The aims of this module include the following:

- ▶ Understanding the role of policymaking for future AI development.
- ▶ Identifying existing policies and policy areas related to AI development and identify central principles included in the current policies.
- ▶ Increasing the learner's knowledge about weak and strong signs of future technological development.



Learning Outcomes

After the completed module, the student is expected to have learned the following:

- ▶ To recognize the policymaking infrastructures and landscape in the learner's own geographical area and beyond.
- ▶ To identify the core ethical principles that are of public concern, constituting sustainable and ethical AI.
- ▶ To be able to follow ongoing ethical discussions of public relevance in the AI area.

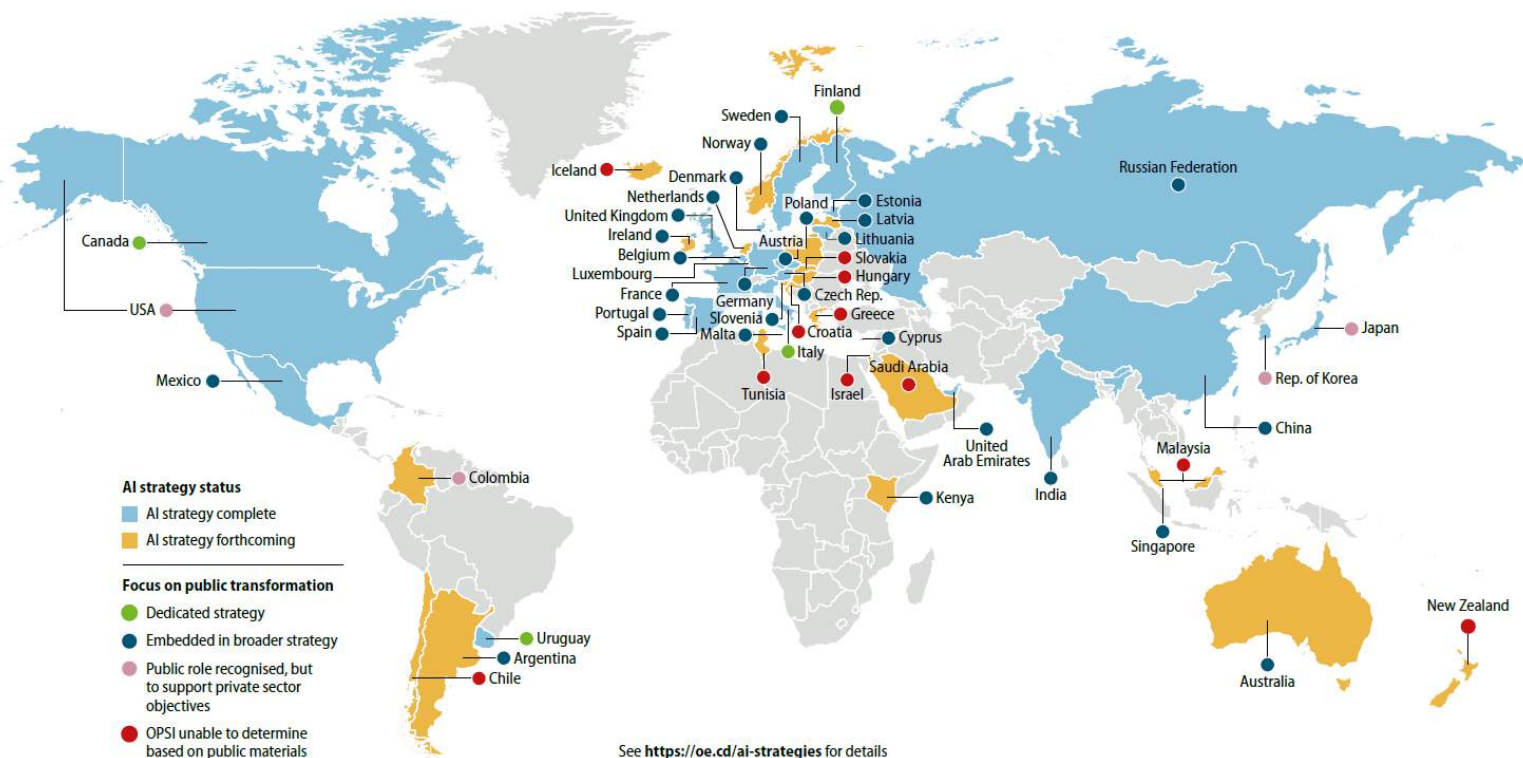


Questions for Reflection

Here are some questions to reflect upon in advance:

1. With regard to the technological device used, how has your environment changed during the last decade, two decades, three decades? For example, what kind of a line of development can be seen in using the telephone or the computer? How has it changed the social interaction in different contexts?
2. How has AI ethics been discussed in your local news media during the past months? What kind of technology- or AI-related public debates have there been recently in your country in the fields of a) schooling and education (e.g. privacy of school platforms); b) economy, banking and financing (e.g. bitcoin); c) privacy of internet users (e.g. face recognition issues)?
3. What are the main regulatory bodies in your country that have the power of creating AI-related policies?

Figure 6: AI Global Landscape



Source: Original map from *Hello, World: Artificial Intelligence and its Use in the Public Sector*, by OECD.

Here are some preparatory tasks to do prior to entering the topic:

1. Take a look at the map AI Global Landscape (OECD) below, describing the status of national AI strategies in different countries across the world in 2020. Select two countries that are expected to be quite different from each other and search for their most recent AI strategy documents in the OECD AI Policy Observatory’s database of national AI policies: <https://oecd.ai/en/dashboards>. Summarize the main points of both strategies and compare them to each other. In which points do the strategic approaches differ from each other, and which features do they share? Where may the differences derive from?
2. What are the leading countries of AI ethics frameworks development and how can you motivate their leading position in terms of numbers, such as the amount of investments and the number of scientific papers? Which areas are less developed when it comes to developing strategies for AI, and why? What could be done to advance the development in less developed areas? Search for country positionings in different rankings and reports and find out on which base indexes are developed.



Outline

By Fredrik Heintz

The world is scrambling to get a grip on AI. Tech companies are exploiting this technology in often impressive, sometimes scary applications. Research institutes are pushing the science and developing AI even further. Policymakers aim to control the development and use of this rapidly developing technology. The public is torn between using the latest applications and confronting the potential risks, often implicitly paying with their personal data. This chapter presents an overview of the major policies and frameworks used by governments, businesses and societies to govern the development and use of AI.

The overall ambition of these guidelines and policy frameworks is often to maximize the opportunities while minimizing the risks and to ensure that the benefits reach as many people as possible. AI technologies bring significant improvements to many areas of society, such as automatic translation between languages, improved cancer detection in radiology and (sooner or later) self-driving cars. As the technology becomes increasingly powerful and pervasive, and as our understanding of the consequences of the technology improves, the risks and potential negative consequences become clearer. Amplified by the speed and scale of automated decision-making, even small issues can have significant impacts.

An important consideration is the data-driven nature of many modern AI techniques, meaning that they find patterns and make recommendations based on collected, often historic data. This leads to risks, such as cementing historical injustices and making biased decisions based on non-representative data. A well-known example is the COMPAS system in the US, which is used for estimating the risk of criminals committing further crimes if released on parole. Another questionable example is an Argentinian system designed to predict which teenage girls are most likely to have an adolescent pregnancy (<https://www.wired.com/story/argentina-algorithms-pregnancy-prediction/>).

A second consideration is the fact that most AI methods optimize some objective function. They are designed to achieve the highest possible score by changing their behaviour. Since typical objective functions could involve, for example, maximizing ad revenue or time spent in an app, they could lead to both undesired and unethical behaviour. One type of unwanted behaviour is when a system finds a way to significantly increase its score while violating implicit constraints that are usually taken for granted, such as personalized marketing or dynamic pricing to make people buy more things, increasing our consumption beyond what is necessary and good for us. The reason for this is that there are important constraints that are never made explicit in the simplified objective functions used by AI systems. This leads to the so-called *value alignment problem*. How do we ensure that the values an AI system is trying to achieve are aligned with human values?

A third consideration is how to maintain meaningful human control over a system. AI systems are often much faster than humans, processing vastly larger volumes of data, and a single system can be used to determine thousands or millions of cases. Taken together, these systems work at a completely different scale than us and often use different methods than we do. This leads to the challenging question of how to maintain control of such a system. There have already been incidents, such as algorithmic trading systems driving the market out of control, requiring significant manual roll-back to correct.

Another interesting question is how to evaluate how well an AI system really works. One can use Google DeepMind's AlphaGo as an example. Go is an ancient two player game in which the players place stones on a grid to control as much territory as possible. This is a game where the best players often refer to intuition and beauty when they describe how they play. In the famous Go games between AlphaGo and Ke Jie, there was a move where the human expert commentators thought the computer had made a mistake, but it turned out

to be a novel and winning move. In this case, it was objectively clear that the move made by the computer was good (since it won the game), while at the same time it was perceived as a bad move based on the experience of the human experts. Now, imagine that instead of Go the computer is helping a physician diagnose a patient. In this case, how should the human doctor act if the computer suggests a diagnosis or treatment that is completely contrary to his or her experience and knowledge? This is a non-trivial question.

These are just some of the complex and intriguing issues that policy frameworks such as the UNESCO Recommendations on the ethics of AI, the OECD AI principles and the European Ethics Guidelines for Trustworthy AI attempt to address. The OECD AI Policy Observatory has collected more than 700 AI policy initiatives from more than 60 countries. According to this study, the US and the UK are the countries with the most AI policy initiatives.

Human-centred and Trustworthy

Policies and policy frameworks refer to general sets of principles to guide the course of action of AI development, agreed on by high-level authorities. Many of the existing policy frameworks take the form of recommendations or guidelines, which means that they are not strict laws but rather voluntary. This does not mean that there are no laws regulating AI – on the contrary. All existing legislation also applies to AI. This includes regulation related to discrimination and privacy. There are also several ongoing initiatives, most notably the EU AI Act, targeting dedicated AI regulation. However, there are also those who argue against hard regulation of AI since it is hard to define exactly what AI is. It is a rapidly evolving field, and regulation normally comes after best practices have been established to promote good practices while prohibiting the bad ones.

The first and most detailed guideline was the European Ethics Guidelines for Trustworthy AI, which is currently being transformed into law. Subsequently, the OECD AI Principles and the UNESCO recommendations on the ethics of AI were released, which are much more global in scope.

The European approach is based on human-centred and trustworthy AI. This means that AI is not a goal in itself but rather a way to improve circumstances for humans. The idea is that we want AI, but we do not want just any AI. The overall objective is to maximize the opportunities and minimize the risks. This is well aligned with other global AI policy frameworks.

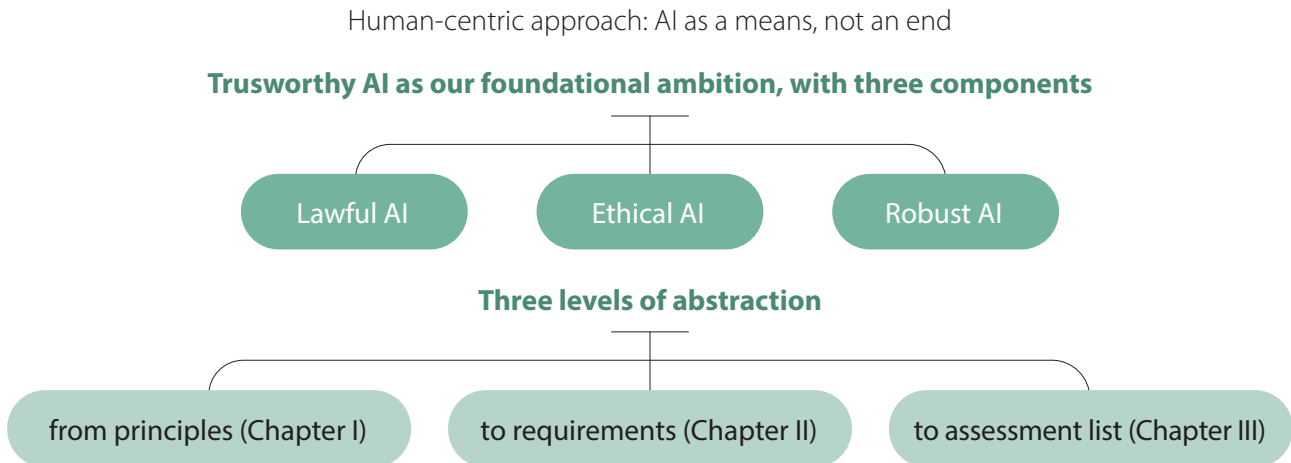
Trust in the development, deployment and use of AI systems concerns not only the technology's inherent properties but also the qualities of the socio-technical systems involving AI applications. It is not simply components of the AI system but the system in its overall context that may or may not be trustworthy. Thus, striving towards trustworthy AI entails more than the trustworthiness of the AI system itself; it requires a holistic and systemic approach, encompassing the trustworthiness of all actors and processes that are part of the system's socio-technical context throughout its entire life cycle (High-Level Expert Group on AI 2019).

According to the High-Level Expert Group on AI, trustworthy AI has three main aspects, which should be met throughout the system's entire life cycle:

1. It should be *lawful*, ensuring respect for all applicable laws and regulations;
2. It should be *ethical*, ensuring adherence to ethical principles and values; and
3. It should be *robust*, ensuring that the implementation actually lives up to expectations both from a technical and social perspective since, even with good intentions, AI systems can cause unintentional harm.

An overview of trustworthy AI in the EU's vision is shown in Figure 7. The ethical framework put forward by the high-level expert group on AI, there are three levels of abstraction, from high-level principles to requirements to ensuring that a system satisfies the requirements. The operationalization of high-level principles is challenging, so these steps are useful for moving us toward realizing the vision.

Figure 7: Ethics guidelines for trustworthy AI - an overview



Source: Composed by authors, based on *Ethics Guidelines for Trustworthy AI*, by Independent High-Level Expert Group on Artificial Intelligence, set up by the European Commission.

Each of these three components is necessary but not sufficient on its own to achieve trustworthy AI. Ideally, all three reinforce each other. However, in practice there may be tensions between these elements, for example, when breaking the law might be necessary to save lives or when the scope and content of the existing law might be misaligned with ethical norms. According to ethical guidelines, it is our individual and collective responsibility as a society to work to ensure that all three components help to achieve trustworthy AI.

The four ethical principles, shown in Figure 8, are as follows:

1. *Respect for human autonomy*. This means that AI systems should augment, complement and empower humans. It also means that people should be allowed to make mistakes and even bad decisions. An interesting case is nudging. If you install an app that nudges you to eat healthier or to exercise more, it is probably ok since you are in control. However, if your insurance company asks you to install the same app to get a discount on your insurance, it is starting to become questionable. This can be taken even further, for example, by a country that requires all citizens to use such an app to improve the health in the country, making it even more questionable.
2. *Prevention of harm*. This is probably the most straight forward principle, as it states that AI systems should prevent physical as well as mental harm to humans. It should be safe to use it.
3. *Fairness*. Again, it is relatively obvious that we want AI systems to be fair, in the sense that the benefits and the costs should be equally divided. At the same time, there are many ways to define fairness. For example, do we want the outcome of a system to be equally shared, or is it the resources that should be equally shared? If we want an equal outcome, it will likely be necessary to allocate resources unevenly to compensate for individual differences.

4. *Explicability*. The fourth principle is the most challenging to define, as it relates to the ability to understand what the system does and on whose behalf. Important concepts include transparency, explainability and traceability.

Figure 8: Four ethical principles based on fundamental rights



Source: Composed by authors, based on Ethics Guidelines for Trustworthy AI, by Independent High-Level Expert Group on Artificial Intelligence, set up by the European Commission.

The real challenge, of course, is how to operationalize these principles. To take a step in this direction, the high-level group defined seven key requirements for trustworthy AI based on these principles. An overview of these requirements is presented in Figure 9. To assist organizations using AI systems and developers to build trustworthy AI systems, the group also developed an assessment list for trustworthy AI.

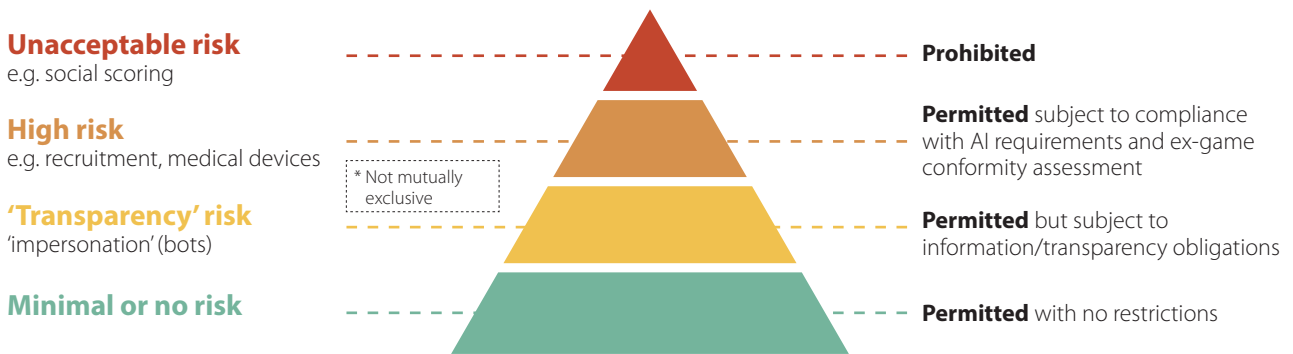
Figure 9: The assessment list for trustworthy AI (ALTAI)

- ✓ **Requirement #1 Human Agency and Oversight**
(Human Agency and Autonomy; Human Oversight)
- ✓ **Requirement #2 Technical Robustness and Safety**
(Resilience to Attack and Security; General Safety; Accuracy Reliability; Fall-back plans and Reproducibility)
- ✓ **Requirement #3 Privacy and Data Governance**
(Privacy; Data Governance)
- ✓ **Requirement #4 Transparency**
(Traceability; Explainability; Communication)
- ✓ **Requirement #5 Diversity, Non-discrimination and Fairness**
(Avoidance of Unfair Bias; Accessibility and Universal Design; Stakeholder Participation)
- ✓ **Requirement #6 Societal and Environmental Well-being**
(Environmental Well-being; Impact on Work and Skills; Impact on Society at large or Democracy)
- ✓ **Requirement #7 Accountability**
(Auditability; Risk Management).

Source: Composed by authors, based on The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment, by Independent High-Level Expert Group on Artificial Intelligence, set up by the European Commission.

Currently, the European Commission is working on the AI Act, which is expected to be the first AI regulation in the world (see Figure 10 for an overview). The regulation takes a risk-based approach, meaning that the amount of regulation is dependent on the level of risk associated with the application. Applications that are considered high risk are those that involve people, especially when the well-being of individuals is involved. It has been argued that some applications have unacceptable risks, such as subliminal manipulation, social credit scoring and large-scale remote biometric identification (e.g. face recognition in public spaces). The intention of the commission is that most applications should be low risk. However, some experts believe the effect will most likely be the opposite, that is, most applications will be considered high risk because either the developer or the user of the system will prefer to stay on the safe side and assume that it is a high-risk application. High-risk applications are heavily regulated, while low-risk applications have virtually no additional regulation (beyond existing regulation). The demands on high-risk applications are largely based on the requirements for trustworthy AI.

Figure 10: A risk-based approach



Source: European Commission.

The questions of monitoring and enforcement are still largely unsolved. According to the proposed regulation, each country should have its own competent body that can both provide advice and monitor the enforcement. Two major challenges are the shortage of highly skilled people needed for these positions and the potential conflict when both giving advice and monitoring the regulation.

The OECD AI principles and recommendations (Figure 11) promote the use of AI that is innovative and trustworthy and that respects human rights and democratic values. They focus on how governments and other actors can shape a human-centric approach to trustworthy AI. As an OECD legal instrument, the principles represent a common aspiration for its member countries.

The OECD uses the following definition of AI:

An AI system is a machine-based system that is capable of influencing the environment by producing an output (predictions, recommendations or decisions) for a given set of objectives. It uses machine and/or human-based data and inputs to 1) perceive real and/or virtual environments; 2) abstract these perceptions into models through analysis in an automated manner (e.g., with machine learning), or manually; and 3) use model inference to formulate options for outcomes. AI systems are designed to operate with varying levels of autonomy.

This definition can be seen as an elaborated version of the one used by the European High-Level Expert Group.

Figure 11: The OECD AI Principles and Recommendations**Values-based principles**

	Inclusive growth, sustainable development and well-being
	Human-centred values and fairness
	Transparency and explainability
	Robustness, security and safety
	Accountability

Recommendations for policy makers

	Investing in AI R&D
	Fostering a digital ecosystem for AI
	Providing an enabling policy environment for AI
	Building human capacity and preparing for labour market transition
	International co-operation for trustworthy AI

Source: Original imagery from OECD AI Principles overview, based on OECD's Recommendation of the Council on Artificial Intelligence.

The UNESCO Recommendation on the Ethics of AI (Figure 12) is another interesting example of soft-law regulation. It is the first globally accepted set of guidelines that takes a broader perspective on AI, including the need for cultural diversity and appropriate education. Like the EU guidelines, it highlights 11 important areas where policy development is needed, including ethical impact assessment, ethical governance and stewardship and development and international cooperation. It also emphasises the need for monitoring and evaluation, including the development of a UNESCO ethical impact assessment and readiness assessment methodology. Likely, these will be important tools for the practical realization of AI technology, as companies and organizations are trying to assess both their readiness and the ethical impact of their products and services.

AI is a truly global phenomenon. Therefore, it is important that global organizations like UNESCO are developing common frameworks that are gaining widespread adoption and acceptance. Only by working together will we be able to harness the potential of AI.

Figure 12: The UNESCO Recommendation on the Ethics of AI

Values	Principles	Policy Areas
<ul style="list-style-type: none"> • Respect, protection and promotion of human rights and fundamental freedoms, and human dignity • Environment and ecosystem flourishing • Ensuring diversity and inclusiveness • Living in peaceful, just and interconnected societies 	<ul style="list-style-type: none"> • Proportionality and Do No Harm • Safety and security • Fairness and non-discrimination • Sustainability • Right to privacy and data protection • Human oversight and determination • Transparency and explainability • Responsibility and accountability • Awareness and literacy • Multi-stakeholder and adaptive governance and collaboration 	<ul style="list-style-type: none"> • Ethical impact assessment • Ethical governance and stewardship • Data policy • Development and international cooperation • Environment and ecosystems • Gender • Culture • Education and research • Communication and information • Economy and labour • Health and social well-being

Source: Composed by authors, based on the UNESCO recommendation on the Ethics of Artificial Intelligence.

This chapter has provided a brief and introductory account of some of the challenges that different forms of AI regulation are trying to address along with two complementary approaches. This is a very active area, and most countries and global organizations are working on different aspects of these issues. We should expect to see a great deal of activity and development in this area in the coming years.

Raising General Awareness of AI :Creating an Open Online Course



Teemu Roos, professor, initiator of the public MOOC
“The Elements of AI” by the University of Helsinki, Finland

“In 2020, we launched a MOOC (Massive Open Online Course), a free online course, called “The Elements of AI”, which was by the end of 2021 attended by 750,000 people. The purpose of the course, designed by the University of Helsinki and the ICT company Reaktor, was to demystify AI. We wanted to encourage as broad a group of people as possible to learn what AI is, what can (and can’t) be done with AI, and how to start creating AI methods.

One of the initial sparks was the National AI Strategy (“AI Era”). In 2017, the Finnish government started the implementation of the strategy and one part of it is to raise awareness among the general public. The University of Helsinki, as well as other higher ed institutions, was asked if we have any generally accessible educational resources. We didn’t but we had done some MOOCs on coding and cybersecurity, so I thought why not do one on AI as well.

As a happy coincidence, we also met people from the company Reaktor at that time, and mentioned about this plan. They wanted to help us because they felt that there’s need for more AI awareness. We had lots of meetings and workshops to come up with a great concept and as a part of it, we chose 1 per cent of the Finnish population as the goal. The point is to have a concrete, ambitious goal, which helped us communicate to others that we are serious about what we’re about to do, and to encourage others to join the initiative.

We wanted empower more people to get into tech in some ways. Some may want to start learning skills by which they can start solving problems by AI in their work. But more than that, we hope that people will be able to form their opinion on what kind of tech should be developed, and how it should be regulated.

The accelerating adoption of AI poses serious challenges to privacy, equality, and democracy. Social AI-powered media algorithms are a part of complex socio-technical phenomena that may lead to polarization; platform business models may lead to extreme income inequalities among the workforce (so called “clickworkers”); centralization of computational and financial resources may even challenge existing political systems; ...

These types of challenges cannot be solved by technology. However, understanding the underlying technology is a prerequisite to being able to address them. Everyone should have access to the basics, so that solving the above challenges is not left only to technologists.

We need to support general awareness: to make sure that everyone has access to the basic knowledge that underlies societally impactful applications. In addition, we need reskilling and upskilling initiatives: people who want to contribute to building technological solutions should have access to the required technical skills. This means that we should offer education ranging from the basics to advanced skills.

It’s important to note that the Elements of AI project is not only an online course but a wider initiative involving an extensive collaboration network with multiple partners in dozens of countries. This comes with significant coordination overhead and requires a fair amount of fundraising. The project is also quite unique in mix of educational, scientific, industrial, public policy, communication, components, which makes

it hard to place it in any existing categories of projects. By the tailored funding structure and collaboration model we have been able to maximize the accessibility of the course (no end-user fees, support in dozens of languages, etc.). Hopefully this success will encourage governments and organizations to support other similar projects around the world."

Link to the MOOC "The Elements of AI": <https://www.elementsofai.com>.

Another examples of MOOCs "AI and the Rule of Law", an introductory course engaging judicial operators in a global and discussion around AI's application and impact on the rule of law, available in seven languages (https://www.judges.org/ai_and_law/english/), and "Defending Human Rights in the Age of Artificial Intelligence", addressing specifically the youth (<https://www.edapp.com/course/defending-human-rights-in-the-age-of-artificial-intelligence-2>).

You can find more MOOCs on course platforms such as the journalism course site by the Knight Center for Journalism in the Americas (<https://journalismcourses.org>), edX (<https://mooc.org>) and Coursera (<https://coursera.com>).



Suggested Assignments



READ AND REFLECT: Read the paper below in which the authors are discussing the potential impacts of AI technologies on society and group the potential accidents in machine learning systems into five areas of risk. Try to come up with concrete examples of your own concerning these risk areas.

Amodei, D., Olah, C., Steinhardt, J., Christiano, J.S. & Mané, D. (2016). Concrete problems in AI safety. *arXiv*, 25 July 2016.



DISCUSS: How can AI promote the global sustainability development goals (SDGs) set up by the United Nations (UN) in 2015? The 17 SDGs include: (1) No Poverty, (2) Zero Hunger, (3) Good Health and Well-being, (4) Quality Education, (5) Gender Equality, (6) Clean Water and Sanitation, (7) Affordable and Clean Energy, (8) Decent Work and Economic Growth, (9) Industry, Innovation and Infrastructure, (10) Reducing Inequality, (11) Sustainable Cities and Communities, (12) Responsible Consumption and Production, (13) Climate Action, (14) Life Below Water, (15) Life On Land, (16) Peace, Justice, and Strong Institutions, (17) Partnerships for the Goals.



IDEATE: Produce a story idea that deals with the European Union's guidelines for trustworthy AI for a) a national daily newspaper; b) a girl's magazine; c) a podcast for senior citizens; d) a reality show for a commercial television channel. You can also make use of information available on your own country at the European Commission's knowledge service AI Watch: https://knowledge4policy.ec.europa.eu/ai-watch_en.

Madiaga, T. (2019). *EU guidelines on ethics in artificial intelligence: Context and implementation*. Brussels: European Parliamentary Research Service (EPRS). Briefing available at [https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRS_BRI\(2019\)640163_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRS_BRI(2019)640163_EN.pdf)



SEARCH AND SOURCE: Choose one platform or online service – for example, Facebook or Spotify – and explore the rules: how does it function for a user? What kind of explainability can be observed, and what kind of shortcomings are there in terms of the system's explainability? How could journalists trace explainability in computer systems?



PRODUCE: Conduct a telephone interview with a system developer – a developer of an app or a platform or a digital service, preferably a small local startup – and try to get him or her to explain how the system works, asking critical questions. Write a short story explaining the system works. What kind of challenges did you face during the interview, and in reporting on it?



Readings

- Amershi, S., Weld, D., Vorvoreanu, M., Fournay, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P.N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). Guidelines for human-AI interaction. *CHI19: Proceedings in the CHI Coinference on Human Factors in Computing Systems*, 1–13. Glasgow. <https://doi.org/10.1145/3290605.3300233>
- Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines*, 30(1), 99–120. <https://doi.org/10.1007/s11023-020-09517-8>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Liao, S. M. (Ed.) (2020). *Ethics of artificial intelligence*. Oxford University Press. <https://doi.org/10.1093/oso/9780190905033.001.0001>
- Loi, M. (2021). *Automated decision-making systems in the public sector: An impact assessment tool for public authorities*. AlgorithmWatch. <https://algorithmwatch.ch/en/adm-publicsector-recommendation/>
- Shane, J. (2021). *You look like a thing and I love you: How artificial intelligence works and why it's making the world a weirder place*. Little, Brown and Company.
- Taddeo, M., & Floridi, L. (2018). How AI can be a force for good. *Science*, 361(6404), 751–752. <https://doi.org/10.1126/science.aat5991>



Module 4:

Reporting in Algorithmic Cultures



Summary

This module helps journalism students understand and investigate AI- and data-driven environments in a critical way. The module discusses the essence and importance of algorithms in contemporary society and in the current market economy, addressing potential issues and calling for ways of exploring them in journalistic coverage. It further discusses the difficulties journalists face in a society increasingly reliant on algorithms and explores ways how journalists can handle, expose, explore and counteract the manipulation of the gatekeeping process by algorithms.

Key concepts: *algorithmic culture, algorithmic bias, accountability, data literacy*



Module Aims

The aims of this module include the following:

- ▶ Getting familiar with the concept of algorithm and algorithmic culture and relating it to discourses on AI.
- ▶ Identifying algorithmic biases in different environments.
- ▶ Increasing the learner's knowledge about equal, fair and responsible AI in terms of the concept of implicit or unwanted bias.



Learning Outcomes

After the completed module, the student is expected to have learned the following:

- ▶ To understand how algorithmic cultures work and are related to AI.
- ▶ To be able to explore and respond to the obstacles that journalists encounter when writing about topics in the algorithmic society.
- ▶ To have ideas on what journalists can do to report comprehensively on an algorithmic society and how can journalists be supported in their reporting.



Questions for Reflection

Here are some questions to reflect upon in advance:

1. In your opinion, is an algorithm a neutral technology? Or are only people and their data biased? Can the algorithm be subjective? How, and why (not)?
2. Is it problematic that people in the Global South are dependent on algorithms that are designed (and therefore based) on the values that prevail in the Western world, or the other way round? How do values guide the production of AI applications?

Here are some preparatory tasks to do prior to entering the topic:

1. Watch a video addressing filter bubbles, for example, Eli Pariser's TED talk (2011): https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles. Reflect upon your own encounters with "filter bubbles" and whether the situations could have been avoided. In more general, what solutions are there to avoid people becoming isolated from each other, and being exposed to some kind of information while not seeing alternative aspects?
2. Search for statistics showing how people in different genders are engaged with technologies in terms of education, occupations, and technological applications – as developers and producers of technological applications and content, and as consumers and audiences. Can you argue for a gender bias, and if you can, why, and where does it derive from?



Outline

By Frederic Heymans

Gatekeeping, the process by which information is filtered through dissemination, has been considered an exclusively human experience for a long time. This practice has changed over the years. First, the rise of mass media began curating news. After this came the development of digital media, which centred around the World Wide Web. More recently, we have observed the emergence of the use of algorithms and their manipulation of the gatekeeping process.

In essence, an algorithm is a set of instructions that, when executed correctly, can solve a problem or complete a task. Algorithms are increasingly used for a variety of purposes: advertising, romantic matchmaking, recommender systems, credit and insurance scoring, evaluation and assessment in education, job application procedures, policing, welfare management, and more. Vastly increased computing power, the presence of large volumes of data and the use of artificial intelligence (AI) and its applications, such as machine learning and computer vision, have led to new momentum in the socio-economic environment.

Algorithmic culture is a society where human thought, conduct, organisation and expression are enfolded into the logic of big data and algorithms. Striphas (2015) was the first to apply the term to the current social context. Striphas sees culture as the sorting, classifying and hierarchising of people, places, objects and ideas, as well as the habits of thought, behaviour and expression related to these processes. According to Striphas, culture has increasingly been outsourced to algorithmic processes in recent years.

The emergence of algorithmic culture was coupled with a community process. People have gradually come to regard algorithms and their functions as normal and to accept them. In many circumstances, people have even been found to prefer the advice of an algorithm over that of a human (Logg et al., 2019). This algorithmic acceptance is based on the fact that algorithms often make life easier. Algorithms have gradually become more accurate and have become a trusted information instrument that ranks, classifies, associates and filters information.

A next step in the evolution of the online public sphere might be a shift from the current digital environment to an immersive space for interaction that is also powered by algorithms. Here, we distinguish between virtual reality (a technology for creating a 3D environment that can be seen, felt and heard), augmented reality (a virtual layer placed over the real world) and the metaverse. The metaverse generally refers to a virtual world in which people can connect with each other online. It is a network of virtual 3D spaces in which users can perform different activities via an avatar.

Several small metaverses have already been set up, but the ultimate idea of a general virtual universe where all these applications are connected does not yet exist. This may also take some time, as several technological issues remain unsolved. Whether these new forms of interaction will be normalized as quickly as algorithms remains to be seen. In addition to the technical hurdles, questions can be raised about the ecological footprint, and people's acceptance of an immersive environment is also uncertain. Recent studies indicate for example that working in virtual reality comes with a lot of discomfort (Kalamkar, et al., 2022). Are people eager to perform daily tasks in an immersive way, or do they just want these tasks done quickly and efficiently? The verdict on this is still out there.

When Algorithmic Culture Fails

Despite the acceptance of algorithms, various issues exist in algorithmic culture. Problems that may be caused by an algorithm include incorrect predictions due to poor programming, the reinforcement of social inequalities and discrimination because of biased data, a lack of diversity and fair representation of the information distributed to the public, and censorship. In addition, the people who design and program algorithms and their individual agendas and biases can also cause issues.

One should never assume that an algorithm is a neutral gatekeeper. It is important to consider the political economy of algorithms (e.g. the ownership structure; O'Neill, 2016), which often directs the results an algorithm will produce. An algorithm can also be used to manipulate human behaviour. This manipulation can take various forms, such as personalized strategies for the consumption of goods, actions aimed at an individual's emotional state and other ways that help increase a company's profits. These are often accompanied by clever designs and marketing strategies. Problems with algorithmic culture are further explored in discussion of two specific social domains: government culture and corporate culture.

In many countries, public administrations have been using algorithms to work more efficiently. Simple tasks are outsourced, and sensitive assignments, such as fraud detection, the distribution of social security and the grading of students in the education system, are now also being entrusted to algorithms. The use of algorithms in a government context may be promising, but there are numerous examples of how it can go wrong, such as in the Netherlands (Geiger, 2018), the United Kingdom (Elbanna & Engesmo, 2020), Brazil (Laranjeira de Pereira & Guimarães Moraes, 2022), Argentina (Jemio, Hagerty & Aranda, 2022) and South Africa (Merten, 2022). Algorithms can be biased, designed badly or implemented with too little ethical sensibility, which can lead to major errors with serious consequences for thousands of people.

Algorithms are increasingly being used to achieve efficiency gains, among other things, in the corporate sector as well. However, the use of algorithms can lead to lower-quality jobs for workers. For example, algorithms are increasingly being used to monitor employees and check whether they are sufficiently productive. This practice flourished during the COVID-19 pandemic and the associated teleworking phenomenon (Cater & Heikkilä, 2021). In another example, a recruitment algorithm that was supposed to screen job applications not only maintained but reinforced an existing bias, thus discriminating against women (Dastin, 2018). In addition, software designed to create employees' schedules caused the staff to be under severe time pressure (Kantor, 2014).

We must not overestimate the impact of algorithms, and we must not allow ourselves to be lured into technological determinism. For example, a major concern in society today is increasing polarization and disinformation. The algorithmic operations of search engines, social media and other digital platforms that work with personalized content are suspected of creating filter bubbles. A filter bubble creates a situation in which an internet user only sees information that reinforces and conforms to their own opinions. These bubbles are mainly produced by ranking algorithms that engage in passive personalisation without any active choice. However, several empirical studies based on survey data and tracking data have indicated that the use of gatekeepers, such as search engines and social media, is, in most cases, associated with more diverse news consumption, the opposite of the filter bubble hypothesis (Ross Arguedas et al., 2022).

Acerbi (2020) nuanced the importance of algorithms in society, indicating that the introduction of digital technologies does not bring about such large and negative changes in human behaviour as digital doomsayers like to claim. For example, digital media and algorithms allow every internet user to reach far more people than traditional means of communication. This has not led to unnatural social interactions, and personal contacts continue to dominate social relationships (Acerbi, 2020).

Algorithmic Hurdles for Journalism

The rise of algorithmic culture has certainly not made the work of journalists and media professionals any easier. First, the introduction of new technologies has had a major impact on the journalistic profession. Today's journalists are expected to have expertise in digital platforms, multimedia tools, the internet and social networks, and they have to be able to react quickly to new developments.

Furthermore, journalists now have to work in a world where decision-making and choices are increasingly influenced by algorithms. Therefore, it is important that journalists understand how algorithms work and what impact they have on both individuals and society. Journalism has the important duty of spreading relevant information, which is not a simple task; it requires a decent understanding of rapidly advancing technology, and an algorithm needs to be studied over a certain period of time. Moreover, organisations – not only platforms but also government bodies – that own algorithms rarely provide insight into how they work. In part, this is because of trade secrets, but it is increasingly because the systems are not designed to explain their decisions (also known as (black box algorithms)). Algorithms are often opaque, and technical complexity is a barrier to reporting. Moreover, the data used to train the model must also be understood. Now that algorithms are driven by complex AI systems that mimic the behaviour of human agents, the task becomes even more difficult. Finally, it is challenging to explain these matters to the general public in a way that is simple and accessible.

In addition to reporting, algorithms challenge the relevance of journalistic work. One good example is cultural journalism (Hu, 2018). The status of cultural reviews is threatened by platforms and their algorithms. For example, who is going to read a music review when Spotify recommends new music on a daily basis via different playlists? Who will consult film reviews when Netflix or the Internet Movie Database (IMDb) personally guides you through the offerings?

With the advent of digital news accompanied by algorithmic technologies that scrutinise web traffic and activity, newsrooms have more insights into what kinds of content their audiences prefer. These metrics have a direct impact on news companies' strategies. Indeed, traditional newsroom positions are coming under pressure as audience metrics and the attention paid to them on the news floor influence journalistic intuition (Ferrer-Conill & Tandoc, 2018; Welbers et al., 2016).

Moreover, news reporting is increasingly influenced by the role played by social media, such as Twitter (McGregor, 2019). The opinions expressed there are often mistakenly interpreted as public opinion. In political reporting as well (Broersma & Graham, 2015), the use of social media has led to an increased interest in subjects such as the private lives of politicians, leading to a "depoliticization of politics". On the other hand, journalists themselves feel pressure to market themselves on platforms such as Twitter or Instagram in an attempt to remain relevant.

Journalistic Resilience in an Algorithmic Age

The increasing presence of algorithms in society requires journalists – just like a large percentage of the working population – to sharpen their skills. A foundation of both data and algorithmic literacy and elements of the computational science of reporting must be important parts of a journalist's skillset, while the core of the journalistic profession must also be practised. This leads to the question of how journalists can retrain themselves.

First, the composition of a structured toolbox for reporting on algorithms can help. Reporters can compile this themselves or use existing toolkits: The Knowledge Centre Data and Society in Flanders and Journalism.ai in the United Kingdom have developed accessible toolboxes to familiarize journalists with AI. A toolbox may consist of an extensive AI glossary and a repository of articles on AI reporting or scientific publications with which a journalist can become acquainted with the terminology; a contact list of AI experts and stakeholders for reflecting on the subject of a report (differentiating between technical, social/ethical and legal experts); a list of AI tools with which

one can experiment (e.g. speech-to-text applications or tools that can be used to unmask fake news) and become accustomed to the technology; and links to professional communities that share new trends and insights.

A second suggestion is to cultivate a critical attitude towards data, which several studies have shown is often lacking among journalists (Simons et al., 2017; Stalph, 2017; Kouts-Klemm, 2019). Accuracy of data tends to be insufficiently questioned; this is important not only for interpreting press releases, research data or company results, but also for scrutinizing algorithms in times of big data. A good level of data awareness and data literacy is therefore an important prerequisite for being able to fulfil the mission of journalism. Improving data literacy does not require training as a data journalist. There are a number of organisations that offer limited, free English-language data literacy courses (see e.g. the University of Sheffield, 2022).

That said, journalistic opportunities do increase when one can evaluate algorithms through the lens of a data journalist. Courses for data journalists vary in content but often include components such as an elementary understanding of arithmetic and probability, advanced statistical techniques and data analysis, allowing journalists to dig deeper into the risks of algorithms and better assess what can go wrong in an investigation. A good starting point is the *Data Journalism Handbook* (Bounegru & Gray, 2021). Also inspiring is scientific research on how data journalism is practised in newsrooms around the world, such as in Norway (Karlsen & Stavelin, 2014), China (Zhang & Feng, 2019), Pakistan (Jamil, 2019) and the Middle East (Bebawi, 2019). A number of initiatives promote data journalism in the Global South, such as Code for Africa (Code for Africa, n.d.) and CFI, a subsidiary of the France Médias Monde group promoting the development of the media in Africa, the Arab world and South-East Asia (CFI, n.d.).

Another possibility is to learn algorithmic accountability reporting, a new branch of journalism that investigates the social power exercised by algorithms (Diakopoulos, 2014). Algorithmic accountability (Diakopoulos, 2021) reporting, an umbrella term for various research methods, is a mechanism for investigating, clarifying and expressing the power structures, prejudices and influences that algorithms can exert in/on society. Effective algorithmic accountability reporting requires all the traditional skills of journalism in reporting, interviewing and domain knowledge and the ability to translate the results clearly and concisely. It also requires new skills, such as (at least) minimal knowledge of data journalism (e.g. scraping and cleaning data) and the use of advanced statistical techniques.

Guidelines for Journalists Covering Algorithms

- Familiarize yourself with the terminology: read research papers and journalistic articles on AI to get acquainted with the correct terms; however, translate technological and scientific terms into everyday language and expressions whenever possible.
- Complicate the narrative: the “true” narrative may not be that simple; problematize oversimplified narratives on AI and triangulate with several experts and stakeholders, i.e. check with a number of experts how they reflect upon the issue.
- Beat the hype: examine technologies with a critical lens, avoid buzzwords and ask yourself: what does this mean in practice, or in the life of your audiences?
- Assess research results: when reporting on research on AI technologies, ask how the research was conducted, what data was used and how large the dataset was, what initial assumptions did the researchers have when approaching the study, and what consequences might come out of the technology.
- Test: if possible, try to acquire personal practical knowledge about the applications you are reporting about by testing new algorithms and systems by yourself.
- Follow professional online communities: national and international journalistic communities, both institutional projects and peer groups on social networking sites, provide journalists with up-to-date tools for thinking and action when reporting technologies and related issues.

Source: Sivasdas & Argoub, 2021

Excluded from Data, Eroded from History?



Jenny Bergenmar is Professor of Comparative Literature at the University of Gothenburg, Sweden, and affiliated researcher at the Center for Digital Humanities, University of Gothenburg

“As the amount of digitized materials increases, it begs the question: what knowledge can one gain through a computer-assisted analysis of these materials? There are many possibilities, but also risks putting your faith in the analysis of large textual corpora with statistic and probabilistic methods.

One problem is the unequal compilation of texts in large digital corpora. There is often a lack of traces in archives left by marginalized groups, and they therefore risk being excluded from historiography once more. On a global scale, there is also an acute unevenness in what languages are represented in digitized corpora, with English as a leading academic language having a clear advantage over smaller languages. Furthermore, not all parts of the world can afford to digitize their national historical and literary sources, which also creates unequal conditions if too much focus is put on data-intensive methods in research funding.

Although algorithmic analysis can be used to identify certain aspects of texts on a large scale, however, there are other aspects that require human reading. One example is the representation of LGBTQI motifs and characters in literature. To identify these, you need to be sensible to queer metaphors and to what is left unsaid in a text, since there are no consistent features on the surface of the text that capture LGBTQI identities and actions.

Finally, a problem is that technology itself becomes the carrier of certain positive values. Algorithmic analysis has in some cases meant the return of formalism and positivism to the humanities, because there is an idea that a computer can analyze more objectively than a human can. The digital, large-scale methods have brought with them a re-evaluation of evidence-based research, sometimes at the expense of critical perspectives, such as feminist, queer, and postcolonial perspectives.”



Suggested Assignments



READ AND REFLECT: Read Nicholas Diakopoulos’ article below, which introduces “the notion of algorithmic accountability reporting as a mechanism for elucidating and articulating the power structures, biases, and influences that computational artifacts exercise in society”, and discuss the methods journalists can use to detect bias in algorithms.

Diakopoulos, N. (2015). Algorithmic accountability: Journalistic investigation of computational power structures. *Digital Journalism*, 7(8), 1–23.



DISCUSS: How much power does an individual user of an app have? How can the algorithmic structures be contested and even manipulated? Take an app with a personalized feed and a recommender system to your example (e.g. Instagram, YouTube), exploring the suggestions that the algorithm presents to an individual user. What do you think the recommendations are based on, and how can the user distract or change the algorithm?



IDEATE: If a central platform would be closed for a) a day; b) for a longer period of time, what would be the consequences for different organizations, communities and individuals? Choose one platform and imagine the possible consequences.



SEARCH AND SOURCE: Conduct online searches to find a current example of biases in AI systems, specifically a bias in a) gender (i.e. how female or transgender people are discriminated by an algorithm), b) ethnicity or race (i.e. how people are discriminated based on their skin colour), c) language (i.e. how algorithms cannot overcome linguistic differences). For each case, discover possible solutions for eliminating or minimizing the bias.



PRODUCE: Select a topic related to algorithms based on a scientific article in one of the special issues of the *International Journal of Interactive Multimedia and Artificial Intelligence* (IJIMAI).

-> Artificial intelligence, digital marketing and neuroscience, 5(6)

-> Soft computing, 6(1)

-> Artificial intelligence and blockchain, 6(3)

-> Artificial intelligence, spirituality and analogue thinking, 7(1)



Readings

van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199970773.001.0001>

van Dijck, J, Poell, T., & de Waal, M. (2013). *The platform society: Public values in a connective world*. Oxford University Press.

Fuchs, C. (2021). *Social media: A critical introduction*. 3rd ed. Sage. <https://dx.doi.org/10.4135/9781446270066>

Hristova, S., Slack, J. D., & Hong, S. (Eds.) (2020). *Algorithmic culture: How big data and artificial intelligence are transforming everyday life*. The Rowman & Littlefield Publishing Group.

Lindgren, S. (2017). *Digital media and society*. Sage.

Seyfert, R., & Roberge, J. (Eds.) (2016). *Algorithmic cultures: Essays on meaning, performance and new technologies*. Routledge. <https://doi.org/10.4324/9781315658698>



© Rodnae Productions Pexels.com

Module 5:

Reporting on Malicious Uses of AI Technologies



Summary

This module discusses issues related to the potentially harmful effects created by the use of artificial intelligence (AI). Technology's position in the public sphere, the inequality of information dissemination and raising larger questions about the appropriate use of technology have been the broad talking points. These concerns call for critical AI literacy among journalists, so they can cover issues with uncertain ramifications. How should the field of journalism be alert to and monitor social developments in terms of AI? This module identifies and accounts for ethical challenges by covering the aspects of excessive use of AI and big data, online propaganda, deepfakes, bad bots, the dark web and lapses in cybersecurity.

Key concepts: *ethics, bad bots, cybersecurity, dark web, deepfakes, online propaganda*



Module Aims

The aims of this module include the following:

- ▶ Identifying the ethical boundaries for the use of AI-powered technologies.
- ▶ Identifying and understanding the different phenomena of malware and misuses of AI in the online public sphere.
- ▶ Learning about emerging verification and mitigation methods to minimize the risks and harms of inappropriate AI use.



Learning Outcomes

After the completed module, the student is expected to have learned the following:

- ▶ To be able to identify the structures of AI-generated online propaganda.
- ▶ To know how to detect bot behaviour and verify AI-generated content.
- ▶ To know how to be engaged in discussions on the uses of AI with a malicious intent.



Questions for Reflection

Here are some questions to reflect upon in advance:

1. Can you recall some recent examples of fraudulent uses of data in your country or region? What happened and how was it covered in journalism?
2. What kind of unpreferred effects of technologies can you identify in your or others' daily life and how do we overcome these ill-effects?
3. How do platform providers try to avoid harms and misuses? Take a look at one specific app (e.g. Instagram) and list the measures that the platform provider has set to avoid cyberbullying, nudism, paedophilia, and so on.
4. Would it be better for journalists to focus on the positive effects of AI only and diminish the coverage on "bad AI"? Are there some mis- and maluses of AI that journalists should not report on in your opinion? Why (not)?

Here are some preparatory tasks to do prior to entering the topic:

1. Study the infographic produced by the non-profit organisation EAVI (the European Association for Viewers Interests) and figure out how the ten types of misleading news could be created through AI. Can they also be detected by AI tools or technologies?

Figure 13: Beyond 'Fake News'



N.B. The impact and motivation assignments are not definitive and should just be used as a guide for discussion

Source: EAVI (the European Association for Viewers Interests).

- Search for online discussions related to AI and algorithms and try to find out what people are afraid of when it comes to their use of applications and platforms. Identify at least 2–3 concerns and try to figure out whether the concerns are realistic or not, and how journalists could mitigate the concerns.
- Find out what kind of automatized tools there are available for detecting fake and fraudulent content on platforms you and/or your friends use.



Outline

By Santosh Kumar Biswal

Despite the fact that the past decades have seen many policies developed, recommendations and practices implemented to mitigate the non-preferred uses and effects of AI and efforts to make AI work for the social good, AI applications may still do harm and cause unwanted effects. Under the command or will of human beings, since AI consists of machine-driven actions that are dependent on humans' intentions, it can also be consciously misused in many ways. Briefly put, AI can be used for the good, but also for the bad.

When covering malicious uses of AI, journalists are faced with the question of how to report on issues related to intentionally harmful contents and processes in a proper way. Like criminal actions, injustices and crises need to be communicated to wide audiences, but the sense of alarm about the effects of AI may also make people feel worried and frightened about an "automated enemy". When addressing the malicious uses of AI, journalists have a responsibility to contextualize these functions in a way that does not make the malicious actions appear too alarming. Journalists need to see the relevance, volume and expected consequences in realistic proportions. Moreover, the general journalistic ethics rules that are applied to crime and justice reporting are valid even in the coverage of AI; individuals should be protected, no one should be publicly suspected or prosecuted by journalists and facts should be validated, to avoid spreading rumours. What makes the situation complex is that disinformation may be generated by AI. At this point, a consciousness of AI's workings is necessary for journalists. Being knowledgeable about the dark web – also known as the deep, invisible or hidden web – has become an important element of AI literacy, both for journalists and citizens.

AI as a Dual-Use Area of Technology

In terms of AI-related threats, experts distinguish between digital security, physical security and political security (Brundage et al., 2018):

- Digital security: automation of social engineering attacks and system vulnerability discovery, more sophisticated automation of hacking and criminal cyber-offence, human-like denials-of-service
- Physical security: commercial systems are used for terrorist aims, the volume of attacks and related damage can be increased and removed in time and space
- Political security: state-initiated use of automated surveillance resulting in the suppression of debate; fake news and fabricated but realistic video and audio; automated, hyper-personalized disinformation and influence campaigns; bot-driven large-scale information-generation attacks; manipulation of information availability and user behaviour

Besides these security issues, the question is often about how to balance the harmful effects of AI against its beneficial effects. A group of experts gathered at a workshop in Oxford in 2018 (Brundage, 2018) suggested an alternative narrative for the common tropes of a "robot apocalypse" and "automation boon": the "dual-use narrative". They defined AI as a "dual-use area of technology", meaning that AI technologies can be used to both beneficial and harmful ends, and the celebratory aspects of AI should always be balanced against misuse-related considerations. AI is a double-edged sword.

Next, we will focus on two AI-technology aspects that are of relevance for journalists when reporting on AI and working in AI-driven environments: malware and bots, and the most recent advancement of everyday applications of AI, deepfakes.

Malware, Bad Bots and Information Disorders

Malware (a portmanteau of the term “malicious software”) can leak private information, gain unauthorized access to information systems, deprive users of access to information and cause harm to individuals and organisations. As computer users, we are already familiar with malware, such as computer viruses, programs that replicate themselves by modifying other computer programs and inserting its own code. There are also other types of malware:

- *Spyware* gathers information about a person or an organisation and sends it to another computer in a way that is harmful for the user.
- *Adware* supports advertising by trying to generate income for advertisers, making the user see advertisements or click on them, while often simultaneously collecting and reporting data about the user to be sold further.
- *Scareware* attempts to cause shock and anxiety via psychological manipulation to make users take actions that may be against their interests, for example:
 - *Ransomware* prevents users from accessing their systems or intimidates them with threats to publish the victim’s personal data unless the user pays a ransom.
 - *Rogue software* misleads users into believing that there is a virus on their computer and persuades them to pay for a fake removal tool.
- *Wipers* erase the data on the hard drive of the user’s computer.

The designed processes that try to trick users into doing something that is harmful to them is called *social engineering*. With the uses of AI in different systems and environments rising, automated processes such as these are becoming intensified. According to the Internet Security Report, malware variants are on the rise. AI is becoming increasingly accessible for actors causing harm, and adversaries can adopt neural networks to harness and reduce the cost of generating novel and highly variable malware. For example, in 2019, the California-based research lab OpenAI, started by Elon Musk, announced the launch of GPT-2, a text-generating program that was able to write a convincing essay on a topic that the researchers disagreed with. Because of the robot’s outstanding performance, the lab decided to withhold the full version of the program, out of fear that it would be used to spread fake news, spam and disinformation in too effective a way.

A specific example of malware is that of malicious bots, or “bad bots”. Bots are software applications that are developed to run automated assignments. They are typically used for web crawling, which refers to bots – also called “spiders” – systematically browsing the internet and, with the help of an automated script, fetching, analysing and indexing information from web servers. The internet is crowded with such crawlers, which are constantly at work. Benign bots aggregate content and deliver automated information, such as weather reports and traffic schedules, sport results or niche news, or they assist in editing, such as on Wikipedia or warn for earthquakes. They may also respond automatically to citizens’ requests and perform other useful services or respond to audiences’ questions during events, such as in the case of so-called “conversational AI”. Such benign bots can, for example, contact and motivate journalists to mention women who are still missing an article on Wikipedia (Flores-Saviaga et al., 2016).

While some bots are useful in their nature, there are also bots that can, among other things, arrange automated cyberattacks or steal content in the fashion of malware discussed above. Malicious bots can steal valuable content from other websites. Bots can be used to distort and mislead marketing analytics, influencing a website’s search engine rankings, spamming community forums with undesirable messages and influencing political elections. Bad bots may work for various criminal activities, including drug menace, child pornography

and credit card fraud, or as a surveillance tool of authoritarian governments for various oppressive activities. Bots that are in interaction with users, including journalists who are doing their daily work, are often called *social bots* because they are designed to act in ways similar to how a real person would act in the social sphere (Stieglitz et al., 2017). Social bots are difficult to detect because of their human-like behaviour, compared to, for example, *spam bots*, which are rather easy to distinguish because they publish a large number of almost identical messages within a short period of time. With AI, social bots can learn to impersonate human users and become increasingly difficult to distinguish from them.

Many social bots are *influence bots* trying to exert influence on human users' views and behaviours. For example, *astroturfing bots* try to create the impression for a user that the vast majority is in favour of a certain position, even if it is powered, for example, by a political campaign organized by one person or a limited group of people. Astroturfing thus refers to the intention to influence opinion in a political debate. *Smokescreening bots* use a certain hashtag in social media discussions to distract or misdirect the reader from the main point of the debate. *Sybils* are bots operating in social media with a fake identity, appearing as *sockpuppets*, accounts with a false identity. These can be, for example, a duplication of a celebrity's account. Bots can also be operated in networks, or "social botnets", where the workings of automated social actors (bots) are synchronized.

Social bots can lead to manipulating others' perceptions and provoking users into displaying strong emotional responses. Bots may like or post large amounts of content in an automated fashion, or they can generate text in response to others' posts and comments. Their effects become more influential the more their automatically generated texts look and feel similar to those generated by humans. As automated fake users, they can influence conversations for commercial or political reasons on a large scale. They can, for example, push selected content to the top of people's news feeds, search results and shopping carts, or promote certain kinds of information and set agendas that benefit specific groups. With this in mind, bots need to be examined as a democratic concern.

Deepfakes – Artificial Misinformation

The development of deep learning has brought about a new and rapidly evolving concern that can be best seen in artificial videos, voice recordings and data that are called *deepfakes*. Deepfakes – a portmanteau of "deep learning" and "fake" – refer to the use of deep learning to produce fake content. They are synthetic media generated with AI-driven processes by distorting visuals and speech that tend to arouse strong emotions. In their most common form, they are videos where a person's face has been replaced by a computer-generated one. In addition to face-swapping that uses facial data for humorous purposes, deepfakes can be encountered in forensic analysis and pornographic material.

In other words, today's AI technologies can be used to fabricate texts and images that look increasingly convincing, and we can expect that their use will continue to be refined and become ever-more sophisticated. AI can, with the help of so-called generative adversarial networks (GANs), create fabricated images that look real because the technology has made great strides within the past few years, as shown in a tweet by Ian Goodfellow, the inventor of GANs, in Figure 14.

Figure 14: Ian Goodfellow's tweet (14 January 2019)

Source: Original image by Ian Goodfellow - from his own Twitter post.

Deepfakes, as falsified content, counteract the historical conviction of images and visual materials as records of reality. Deepfakes easily go viral and, when depicting news events, as in the case of men on a motorbike kidnapping a child in India in 2018, may lead to mass panic among the audience, with fatal consequences that lead to deaths of several people (BBC, 2018). The production of deepfakes also seems to be intensified during elections, where presidential candidates' comments and appearances are a popular topic for deepfake videos and audio. The producers of the documentary *Roadrunner*, directed by Morgan Neville, aroused vivid ethical discussions about AI in 2021 when they used AI-generated audio to re-produce the voice of the late Anthony Bourdain, who died by suicide in 2018, to make him speak posthumously. The synthetic voiceover was used without disclosing it to the viewers, and even if AI was only used to narrate text that Bourdain himself had written, it aroused many troubled reactions on the potential for AI use. What if more dead people were brought back to life? Is that really ethical?

In 2021, a team of computer scientists at a start-up based at Rutgers University, in the United States, trained AI to mimic Ludwig van Beethoven's musical style to complete a symphony based on his sketches. Beethoven, who died in 1827, never finished his 10th symphony. By using the sketches he left behind, mainly written notes, and the idea of variations on a theme, a typical structure in classical symphonies, the researchers tasked AI with learning from the musical data and allowed it to evolve a whole new movement of the symphony. Professor Ahmed Elgammal described the AI learning process as follows:

We first trained the AI to generate the composition as two lines of music, not as a full symphony, which is a typical way of a composer works – by just composing first and then orchestrating. So then, we had another AI that would take that composition and learn how to orchestrate it. I believe this is very similar to the way humans

learn – you cannot really master fourth-level college without going through the first and second and third levels first. It's always incremental (Goodyer, 2021).

These examples show that AI can create non-existent, but imaginable, new worlds that are both immersive and powerful in their effect – while this composition exercise was designed for positive purposes, there is also potential to use it to bring about bad ends. Because of their increased reality effect, or intrusion into the observed reality, deepfakes pose a potential threat to political communication, in particular, and even democracy in general. More particularly, deepfakes have increased the potential to strengthen mis- and disinformation that is both realistic and efficacious. Deepfakes thus need active debunking measures from journalists.

Mitigating Threats to Democracy

The negative potential of AI can be especially harmful if it is harnessed in transitioning democracies, where the governance is still weak and regulatory frameworks are lacking. If the potential of AI is harnessed by challengers to the state and democracy to attack nascent democratic institutions or processes, the consequences may be voluminous. For example, it is imaginable that a supporter of a terrorist group could fabricate a provocative video or audio recording attributed to a governmental authority in an effort to add fuel to political or religious polarisation, and pertinent hate speech can be amplified on frequently used social media platforms. AI-powered drones or automated small arms could all serve as vehicles of violent actions at a relatively low cost.

Alternatively, AI may be a threat to democracy and human rights in authoritarian states seeking to suppress political adversaries or otherwise non-preferred, marginalized or vulnerable groups. For example, an authoritarian government could use AI for increased surveillance by collecting data and individuals' facial imagery to be used by existing monitoring applications. Automated monitoring and surveillance may be a good thing if someone is suspected of planning a school shooting or running a child pornography ring, but it can equally be used to monitor people not suspected of anything. State-led cyber-intelligence could use it to give domestic actors advantages over their foreign counterparts, or AI could be leveraged to influence opinions in favour of one's own interests.

In addition to the above, the harmful effects of mass-scale automatized information production and transmission may not arise only because of false or fraudulent content. Venturini (2019) suggests the term *junk news* to refer to viral news content that saturates public discussion and leaves little space for alternative debates, arousing strong but polarising emotional reactions. Through appealing emotional engagement and clickbait, large audiences can be mobilized via large-scale automated productions involving propagandistic, ideologically extreme or conspiracy-oriented (pseudo-)news and information, overshadowing some democratically more relevant concerns.

To curtail the potential for negative AI, a common framework for the effective governance of internet-enabled technologies has become essential. In 2018, the software corporation Microsoft took the initiative to make a case for a Digital Geneva Convention to advance industry and civil society efforts to protect civilians from state-organized cyber-attacks. In the same year, the Paris Call for Trust and Security in Cyberspace identified a number of target areas for governance, including protecting the integrity of the internet, preventing interference in electoral processes and clamping down on offensive online actions by non-state actors. The Internet Governance Forum, a multistakeholder governance forum with the aim of bringing together governments, the private sector and civil society for policy dialogues, discussed issues related to digital AI technologies on a regular basis. Inter-governmental organisations can help government officials in AI to partner with entrepreneurs and corporate actors to improve their skills, practices and routines in digital forensics. Improved digital forensics may, in turn, help government institutions to identify more quickly fabricated media content and help them remove inflammatory content to de-escalate violence. Multistakeholder initiatives based on specific needs can

also help by establishing toolkits that can support balanced democratic discourse and action. AI tools can be used to detect, decelerate and remove, for example, extreme speech and other harmful behaviours online, such as along the lines of AI4Dignity, a project funded by the European Research Council to tackle online extreme speech with technology. This has been put forward jointly by AI developers, fact checkers, anthropologists and policy-makers.

Given the many potential risks and harmful effects of the malicious uses of AI, preparing for “a good AI society” is becoming an oft-discussed topic. With this in mind, the term friendly AI (or “FAI”, coined by Yudkowsky, 2001) is used to refer to the goodwill uses of artificial intelligence. A central idea behind it is that AI can have a benign effect on humanity, or at least align itself with human interests or contribute to the fostering of improvements in societies. While AI ethics, in general, as discussed in *Module 3: Policy Frameworks and Recommendations for Artificial Intelligence* in this book, are concerned with how an AI agent should behave so as not to produce harmful effects, FAI directs our attention even more towards ensuring that such effects are adequately constrained by asking how to practically bring about positive kinds of behaviour.

The accumulation of power in the hands of a few – the power of money, those with political and religious power or the power over infrastructures for democracy and discourse – must be seen together to inform the debate about ethics and laws for AI. Indeed, ethical concerns and efforts to track the shortcomings of AI should be deliberated as part of the public discourse. AI should be used to support people’s formation and exchange of political beliefs, rather than for persuasion based on the vested interests of some entity’s political will. Factors such as human rights, surveillance and freedom of speech are the larger discussion points in this context. For that to be possible, in a broader perspective, the development of AI digital literacy is required, and journalists are a part of that mission.

Guidelines towards More Benign AI:

- Ensuring citizens’ digital, physical and political security
- Fostering collaborations among policy-makers, AI experts and researchers to check the wrong intentions of AI tools
- Formulating best practices that could be wholesome in nature in terms of the judicious use of AI tools
- Engaging and deliberating among each and every stakeholder of society on the applications of AI tools
- Ensuring the quality of information and transparency in the course of using AI
- Eliminating the digital divide and assuring the space of equality
- Ensuring the freedoms of expression, assembly and association, and the right to work
- Promoting AI literacy globally

How could a journalist contribute to these global goals?

Why Analytic Scepticism is Important for Reporting AI



Steve Woolgar, Professor, Linköping University, Sweden

New technologies often quickly lose their relevance. While newsworthy and often surrounded by much hype when introduced, once even slightly old are of much less interest. An important consequence of this is that general lessons about the emergence and reception of technology can easily get lost. Not many are now particularly interested in the social dynamics of cybercafes, yet the analysis of the hype generated around their potential impact illuminates much about technology in general.

Between 1997–2002, I directed a national interdisciplinary programme in the UK researching the social impacts of (the then) new electronic technologies: 22 projects based in British universities and 3 abroad. The central assumption, and hence the urgency of the research, was that the many new electronic technologies were about to radically change society, culture, organisations, personal identity, and so on. The funding body, the Economic and Social Research Council (ESRC), designated this initiative the “Virtual Society programme”.

As the incoming programme director, I queried the initial assumption of a wholesale movement to a virtual society. Should we not maintain some caution in adopting this premise at the outset? However, the wording of the title – having been approved by various committees – was unchangeable. How then to avoid embracing technological determinism but at the same time address anticipations of widespread impact? After some discussion, I managed to persuade ESRC to retain the title but add a question mark: “Virtual Society? programme”.

The research coming out of “Virtual Society?” included some surprisingly counter-intuitive findings. For example, that the emergence of e-mail networks encouraged more travel, not less (so much for the “death of distance”); the transition to online encouraged greater (not less) physical visits to art galleries and museums; and the adoption of Internet technologies was far from ubiquitous among certain populations of potential users (Woolgar, 2002). In 2000, this last finding generated a media storm. Our research programme office was overwhelmed by journalists asking if this marked the end of computers!

Notwithstanding the iconic distinction between the working practices of academics and journalists – “the worst thing you can say to an academic (journalist) is that they write like a journalist (academic)” – the experience of “Virtual Society?” demonstrates the clear need on both sides for analytic scepticism: the stance of distanced reflection in the face of beguiling claims about novelty and impact. We need that question mark. This is especially important when it comes to artificial intelligence (AI), since its supposed effects are not just societal impacts – they also question our deepest assumptions about what it is to be human.

It is important to recall that what we consider AI today is one in a series of different waves of AI (Woolgar, 1985) and that AI is just the latest in an historical sequence of phenomena construed as a challenge to the essence of intelligence, or humanness, and to the supposed distinctiveness of humans vis-à-vis animals and machines. We can think of seventeenth century mechanical puppets, clever Hans the horse, talking apes, and so on.

How then to speak of the future prospects for AI? I recall a conference in the late 1990s, largely comprising techno-enthusiasts from IT companies. When invited to speculate about future technology, many participants offered visions of dramatic societal changes, both positive and negative. Yet a far more compelling vision was offered by Sir Richard Sykes. In 20 years’ time, he said, we will all be much more relaxed about new technology and much less easily spooked by cyberbolic* speculations of utopian and dystopian technological consequences. And what is the reason for this relaxed and calm state of affairs,

for the cool with which future generations can be expected to respond to exaggerated claims about new technology? The education and influence of the principles of analytic scepticism.

If the recent BBC Reith lectures on AI are anything to go by, Sykes' prediction is still some way off. In reporting the latest technological phenomena, we must try harder not to forget the question mark in "Artificial Intelligence?"

* Cyberbole: a neologism denoting exaggeration (hyperbole) of the nature and consequences of cyber technologies



Suggested Assignments



READ AND REFLECT: Read the article on Wikipedia's bots and their behaviour indicated below. Can you detect bot behaviour on Wikipedia? Try to identify users that are likely to be bots and trace their edits. How have they contributed to the articles in this online encyclopaedia, and how does their contribution possibly differ from the humans?

Tsvetkova, M., García-Gavilanes, Floridi, L. & Yasseri, T. (2017). Even good bots fight: The case of Wikipedia. *PLoS one* 12(2).



DISCUSS: Media and information literacy (MIL) has often been introduced to tackle challenges in people dealing with malicious online content. Could media literacy, or its subarea "AI literacy", help eliminate fraudulent uses? In which ways, and what kind of challenges are there? What kind of skills would "AI literacy" consist of?



IDEATE: Identify a context or situation where deepfakes of a certain kind could be used for a) benign intentions, b) malicious intentions.



SEARCH AND SOURCE: Search for verification tools online and use the handbooks for journalists indicated below to discover existing software for the detection of a) deepfakes, b) fake accounts (sockpuppets) on popular social networking sites, c) junk news.

Silverman, C. (ed.) (2014). *Verification handbook: A definitive guide to verifying content for emergency coverage*. European Journalism Centre (EJC). Available at <http://verificationhandbook.com>

Silverman, C. (ed.) (2015). *Verification handbook for investigative reporting: A guide to online search and research techniques for using UGC and open source information in investigation*. European Journalism Centre (EJC). Available at <https://datajournalism.com/read/handbook/verification-2>



PRODUCE: Write either a mini essay (opinionated text) or a mini feature article (including an interview) to address what we know on state-led surveillance, focusing on China, Russia, or the U.S. Discuss, perhaps with a recent case that has been covered in public, what kind of a role AI plays in it, how it affects individual citizens, and what are the future prospects.



Readings

- Brundage, M. et al. (2018). *The malicious use of artificial intelligence: Forecasting, prevention, and mitigation*. Future of Humanity Institute. <https://arxiv.org/pdf/1802.07228.pdf>
- Diakopoulos, N. (2021). The algorithms beat: Angles and methods for investigation. In: Gray, J., & Bounegru, L. (eds.) *Data handbook 2: Towards a critical data practice*. <https://datajournalism.com/read/handbook/two/>
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- Graves, L. (2018). *Understanding the promise and limits of automatic fact-checking*. Technical report. Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-02/graves_factsheet_180226%20FINAL.pdf
- Horák, A., Baisa, V., & Herman, O. (2021). Technological approaches to detecting online disinformation and manipulation. In: Gregor, M., & Mlejnková, P. (Eds.) *Challenging online propaganda and disinformation in the 21st century*. Palgrave Macmillan, 139–166. <https://doi.org/10.1007/978-3-030-58624-9>
- Reuter, C. (Ed.) (2019). *Information technology for peace and security: IT applications and infrastructures in conflicts, crises, war, and peace*. Springer Vieweg. <https://doi.org/10.1007/978-3-658-25652-4>
- Yampolskiy, R. V. (Ed.) (2019). *Artificial intelligence safety and security*. CRC Press. <https://doi.org/10.1201/9781351251389>



© Rodnae Productions, Pexels.com

Module 6:

Patterns of Storytelling about Artificial Intelligence



Summary

This module discusses AI as an object of news reporting and a topic of public coverage, and the related aesthetics and practices of presentation. As journalism plays a crucial role in mediating ideas related to new technologies to the public, it is important for journalists to be able to select societally relevant topics and frame them in a way that is not overstating or sensational. This chapter provides insight into how AI has been covered in professional media so far and discusses the consequences and implications of different journalistic choices. It also addresses some central shortcomings in the infrastructures of production and reporting of AI and encourages future journalists to reflect upon their role as potential changemakers and audience educators.

Key concepts: *aesthetics, journalistic storytelling, visualization, journalistic roles, specialization*



Module Aims

The aims of this module include the following:

- ▶ Examining how AI is typically presented in journalism and examine the strengths and weaknesses of existing AI news coverage.
- ▶ Understanding how recurring patterns in journalistic storytelling affect the public presentation of AI.
- ▶ Encouraging students to develop reporting on AI.



Learning Outcomes

After the completed module, the student is expected to have learned the following:

- ▶ To be able to reflect upon the practices of storytelling and journalistic performance related to AI.
- ▶ To be able to understand the constraints and possibilities of journalistic AI coverage, as well as assess the quality of such coverage.
- ▶ To be able to create ideas on balanced, sound and versatile AI reporting.



Questions for Reflection

Here are some questions to reflect upon in advance:

1. What professional and personal challenges do you feel when approaching new technologies and applications as a potential topic of coverage?
2. Try to recall recent articles on AI that you have read. What patterns have you noticed? Try to reconstruct prototypical articles on AI in the following genres: a) short piece of news, b) television news report, c) person portrait, d) column, e) tweet by a journalist.
3. Use a search engine to discover images of AI (e.g., Google's image search or a stock-photo library). How would you describe the typical images displaying AI?
4. If you would like to sell a story related to AI to a newsroom, which desks would you contact first?

Here are some preparatory tasks to do prior to entering the topic:

1. Find an article discussing AI in a newspaper or magazine. In which ways is AI presented as an object of inquiry? Analyze it in terms of its discursive strategies:
 - a) What is the main message that the article intends to construct about AI?
 - b) What terms and expressions are used about AI? How do they characterize AI?
 - c) Who are interviewed and what are the main sources? Who else could have been interviewed and are there any biases in sourcing?
2. Find 4–5 articles on AI, either in one news outlet or from different outlets. Pay attention to how they are illustrated. What kind of patterns do you notice? What kind of alternative ways of illustration had there been, and why do you think they were not chosen?

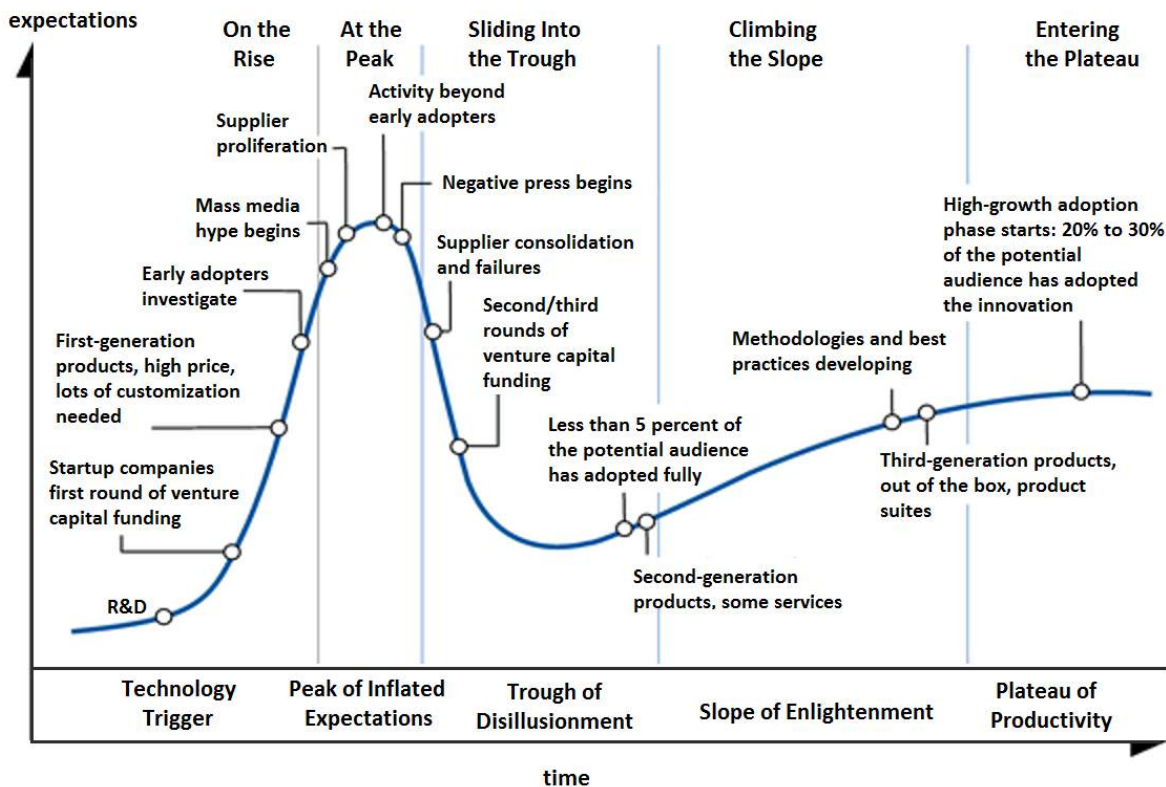


Outline

By Maarit Jaakkola

Journalism plays a pivotal role in mediating central ideas and conceptions on nascent technologies such as AI to the public. Especially in the context of emerging technologies, there are often high insecurities and even disinformation about innovations. Because ordinary people typically lack a direct and explicit experience with AI, even if they may be in touch with it, the general public is largely dependent on journalists' reporting about AI. The public sense-making and imaginaries contribute to the public understanding of the role of technology in different sectors of society and the ramifications of such emerging phenomena as big data and deep learning. It is thus important to observe the dynamics, patterns, and possible biases in reporting that result from daily journalistic practices. How do media typically cover AI?

New technologies are not necessarily easy to cover, and they are typically projected with hopes and fears – particularly true for AI. While juggling pessimistic and optimistic perspectives, there is always a high extent of uncertainty concerning what the new technologies actually are, how they work, and what consequences they may have for citizens and society in the long run. It may be difficult to assess what is actually important and societally or democratically relevant and what are mere buzzwords. New inventions are observed to follow dynamics where expectations initially inflate and spur wide attention, which over time first relapses into disillusionment and criticism, and thereafter gradually begins to develop into enlightenment and understanding of the emerging technology. The dynamics displayed in Figure 15, called Gartner's hype cycle, according to the American company that detected it (Fenn, 2010), applies even to the AI winters discussed in Module 1: Defining Artificial Intelligence, that can be understood as valleys of disillusionment with regard to AI technologies. The S-curve model of attention may not be a precise description of the attention received by any innovation, but it may vary from case to case, which has also been demonstrated in empirical studies of media attention dedicated to emerging technologies (Shi & Herniman, 2023).

Figure 15: Media attention dedicated to emerging technologies (Gartner's hype cycle)

Source: Composed by Olga Tarkovskiy (CC-BY-SA 3.0), based on Gartner.com original piece.

Textual Representation

Typical topics in which AI is regularly followed include healthcare, science, economy, employment and job market, business and technology, education, as well as media and cultural industries (Zhai et al., 2020; Fast & Horvitz, 2017; Chuan et al., 2019). More specifically, topics that have been found to be of special interest in the empirical analyses of general journalistic coverage of journalism include, above all, AI in autonomous vehicles, chatbots and other applications, computer games, mobile devices, warfare, development of computing, and philosophy of consciousness (Vergeem, 2020).

Several studies have discovered that in daily newspaper coverage, industry sources dominate and the discussion about AI is typically conducted by male stakeholders who can have financial gains from AI (Brantner & Saurwein, 2021; Brennen et al., 2018; Ouchchy et al., 2020). In the study of AI coverage in the UK, the Reuters Institute for the Study of Journalism report (Brennen et al., 2018) found that almost 60% of news articles across different outlets covered new industry products or initiatives. Products with AI that had caught journalists' attention ranged from smartphones to running shoes, and from sex robots to brain preservation, the latter being more sensational than the mundane devices. A third of unique sources across the articles were affiliated with industry, almost twice as many as those from academia, and six times as many as those from government. Nearly 12% of all articles included a reference to Elon Musk. The dominance of insider and expert views over non-expert views indicates that major global industries are more agile and capable of making new innovations visible and available for newsrooms as potential topics of coverage, while nation-states are less prone to arrange promotional events that would provide journalists with intriguing topics to discuss. The discussion should

obviously be more diverse and multivoiced in terms of sourcing, and journalists should consciously give space to actors other than major industry sources. Activists, non-governmental organizations (NGOs), and other civil-society representatives, as well as customers, should be included in reporting, as they are experiencing the consequences of AI in their everyday lives.

In addition to coverage with a focus on industry and company representatives, AI technologies tend to be personified or anthropomorphized in news texts. Bunz and Braghieri (2019) examined the coverage on AI systems in healthcare in three English language newspapers from the UK and the US over four decades (1980–2019) and found that the trend of addressing AI systems as persons evolved gradually. In the first two decades, AI systems were still discussed more as a brain than a person, but during recent years, the person trope has become the dominant frame. This may be because AI systems – even when they are weak AI systems – are increasingly configured in an active role of having agency and an individual name. Researchers suggest that coverage focusing on the development of a technology that is still speculative shows a stronger tendency towards anthropomorphising compared with coverage that reports a concrete technology that is already being tested or used. This suggestion yields a more far-reaching question: Does AI become less anthropomorphized and otherized the more integrated it becomes in our everyday lives?

Studies show that AI has also been approached as a value-laden area and seen through politicized lenses. According to the Reuters Institute report (Brennen et al., 2018), right-leaning outlets tend to highlight issues of economics and geopolitics, including automation, national security, and investment, whereas left-leaning outlets underscore issues of AI ethics, including discrimination, algorithmic bias, and privacy. Studies seem to suggest that AI coverage does not have an overtone of alarmist, pessimistic, or dystopian attitudes underlining the risks and threats, but rather tends to lean towards optimism (Brantner & Saurwein, 2021; Garvey & Maskal, 2020; Vergeer, 2020; Fast & Horvitz, 2017). However, coverage may have unintentional consequences: The pursued neutrality, person-centredness, or industry perspective may lead to a disclosure of power structures.

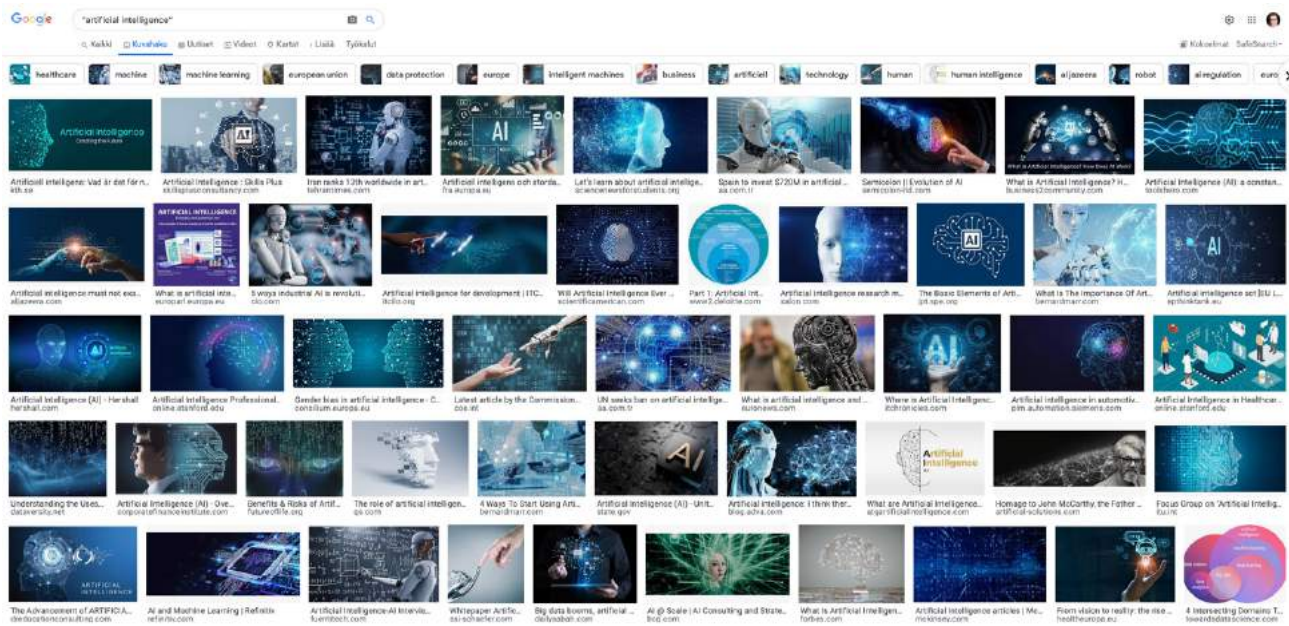
The world's monitoring eye does not sweep the globe equally, but a geographical bias in AI reporting can be found. The countries that are the most frequently monitored in international outlets are the US and China (the “power players” of AI) as well as the UK, Canada, South Korea, Japan, and Germany. These countries also rank high on the list of the Global AI Index, which covers 62 countries in the world and assesses their AI implementation, innovation, and investment (see <https://www.tortoisemedia.com/intelligence/global-ai/>). While some Western and Asian countries strike out, wide regions are never discussed in terms of AI technologies.

Visual Representation

When it comes to illustrating AI in visual news discourse, AI, like many other technological phenomena – such as “big data” or “cloud computing” – provide some challenges for journalists because of their abstract quality. AI is difficult to represent visually. The visuality related to technology has often been described as monotonous or even relatively neutral (Pentzold et al., 2019). The UK-based initiative *Better Images of AI* – launched by the research and development section of the BBC, non-profit organization We are AI, and The Leverhulme Centre for the Future of Intelligence – wrote on their homepage that the visual pattern of AI images is very standardized (as illustrated in Figure 16):

Type “AI images” into your search engine and you will notice a pattern. ... The result is striking, and it's the same on photo libraries and content platforms. In fact, the lack of variety, and the inaccuracy is almost inescapable. The predominance of sci-fi inspired and anthropomorphised images, and the lack of readily accessible alternative images or ideas, make it hard to communicate accurately about AI.

Figure 16: A screenshot of a Google search for images of AI (February 2022)



Source: Google search screenshot done by authors.

The initiative argues that the monotony of the photos matters because they undermine the wider public comprehension of AI technologies, leaving people in the dark regarding the important changes that impact their lives. According to them, stock images of science communication, as in the case of AI images, add to the public mistrust of AI. Romele (2022) terms the visual stream of AI representations as “aesthetic”, communicating in a repeatable way disengagement, inaccessibility, and “black-boxness”. Neutrality in visuals runs the risk of reinforcing a belief in the possibilities of objective quantification of technologies or other myths of dataism and singularity (discussed in more detail in Module 2: Cultural Myths and Narratives about Artificial Intelligence). Societal discourses on ethical problems, such as surveillance and privacy issues, may be difficult to integrate into illustrations.

To tackle the visual challenges, the *Better Images of AI* initiative launched a free stock-image library to provide and support the production of alternative images (<https://betterimagesofai.org/images>). The initiative-takers argued that images should portray technology more realistically and show people behind this technology, and they formulated the following principles for what better images of AI should be:

- Represent a wider range of humans and human cultures than ‘caucasian businessperson’
- Represent the human, social and environmental impacts of AI systems
- Reflect the realistically messy, complex, repetitive and statistical nature of AI systems
- Accurately reflect the capabilities of the technology; it is generally applied to specific tasks, it is not of human-level intelligence and does not have emotions
- Show realistic applications of AI now, not in some unspecified science-fiction future.
- Don't show physical robotic hardware where there is none
- avoid monolithic or unknowable representations of AI systems
- Don't show electronic representations of human brains
- Constitute a wider variety of ways to depict different types, uses, sentiments, and implications of AI

Source: Better Images of AI (<https://betterimagesofai.org/about>)

Examining the visual rhetoric applied to big data in American newspapers, Pentzold and colleagues (2019) found that the concept of data was often illustrated with visualized numerical information in infographics depicting numbers, artistic collages displaying statistics, apps depicted on screens of mobile devices, or through logos, such as IT company logos. Illustrations were also realized by showing the material dimension of computing, such as gadgets like computers, smartphones, cameras, drones, or smart gear, or hardware necessary for processing and storing data, for example, microchips, server farms, or IT company premises. Furthermore, the industry was personified through prominent individuals, who according to Pentzold and colleagues' findings, showed the already male-dominated industry almost purely as a "man's world", resembling findings from science journalism more generally (see, e.g., Kim et al., 2016).

Professional Role Differentiation

Journalistic organizations and professions are traditionally differentiated according to topics, such as the traditional news beats of politics, economics, sports, and culture. This means that there are news desks and professional subgroups of journalists with a specific professional identity specialized in these areas. When it comes to AI, at the current moment there are no permanent structures for covering it, and it varies from newsroom to newsroom how the reporting is organized. The global modern-day challenges ranging from climate change to pandemics have increased journalists' need for specialization, even if these science-related topics can also be covered by generalists. Even for generalists, the global challenges – of which AI constitutes one emerging field attached to science and technology as expert fields – pose challenges to understanding and synthesizing complex areas.

The group of journalists covering AI as a journalistic topic is no homogenous group. Unlike environmental journalism, there is not yet a distinct tradition of AI journalism. Instead, journalism following AI as a field must draw on existing journalistic traditions in each country, of which science journalism, technology journalism, economic and finance journalism, consumer journalism, as well as solutions or constructive journalism may come closest to the substance of AI. In addition to these cross-sectoral approaches, AI is expected to be covered as part of news and local journalism, often carried out by generalists, as well as in international affairs and even cultural journalism, carried out by specialists.

AI coverage can, in other words, include the following types:

- Coverage by generalists: AI is covered as part of news journalism in general. Stories are written by journalists with specialized knowledge, often following communicating organizations and spin sources.
- Coverage by specialists: AI is covered as part of traditional news beats or trans-sectorally, without having a permanent place in the organizational structures of a newsroom, similar to environmental and consumer journalism.
- Coverage by AI-specialized journalists: AI is covered as an area of specialization and monitoring by a journalist specialized in this particular field. Specialization often forms a niche for a science- and technology-oriented journalist who may optimally work as a freelancer, contributing to several journalistic outlets and thus being able to fully, or to a high extent, focus on specifically AI-related topics.

Journalism genres are likely to embrace AI differently, pursuing their agendas in various ways. It is also possible, although not yet richly studied, for news outlets with different profiles – ranging from political to religious positionings and from quality to popular press – to evaluate the topic in differing ways, with regard to the observation that in different studies, AI is encountered with various value propositions. The central ideas of computing related to AI, such as singularity and the creation of intelligence as a "God Machine", may contradict the Christian worldview, as seen in discussions on medical vaccinations and nanotechnology (Vergeer, 2020). Popular press may endorse the sensationality of ideas of computers "taking over" more eagerly than quality dailies and public broadcasting.

Economic and finance journalism endorses the aspects of economic growth, profit, and employment in terms of AI, while arts and cultural journalism may put forward philosophical debates about mind and consciousness and human–machine interaction. Last but not least, regional orientation of news outlets makes newsrooms prioritize the topic differently. While leading national outlets in general dedicate more attention to global currents than regional and local outlets, which do not even have specialized staff to monitor technologies, AI may appear as somewhat less important, or even a foreign topic, to local newspapers and broadcasters. This does not, of course, need to be so; elucidating complex questions may be an equally meaningful task for local journalists. However, delving deep into issues may require time and education that in newsrooms with scarcer resources and less role differentiation may not always be realistic.

Furthermore, it is noteworthy to pay attention to the gendered patterns of technologies and the related structures in reporting in the context of AI. Over the past years, there has been a call for increasing the share of female professionals in the science, technology, engineering, and mathematics (STEM) fields, as women have been observed to be underrepresented in decision-making and technical positions in these industries. The related industries and educational institutions have recognized the gender differences and created educational interventions to encourage more female candidates. In 2019, 25% of American teen girls said that they would pursue a career in the STEM fields, while the corresponding share among teen boys was 57% (Women's Media Center, 2019), and according to the Adeva IT network, in the computing job market, the share of female employees is only 25%, and only 6% of users with profiles on the software sharing platform GitHub are female (<http://adevait.com>). Some debaters have brought the poor representation of females in the STEM fields partly back to media cultures, suggesting that there has traditionally been a lack of female STEM characters in film, television, and online entertainment, and therefore, to influence females' career choices, highlighting female personalities in technological contexts would be important. In 2019, the Women's Media Center found that less than a third (27%) of STEM characters in Western entertainment media were female. The archetypal characters representing science and technology tend to be male physical scientists, engineers, and computer professionals, even if media industries have consciously advanced the representation of females. On the other hand, in many cases, AI has been represented in the form of female characters – gynoids or fembots – ranging from films such as Fritz Lang's *Metropolis* (1927) to Alex Garland's *Ex Machina* (2015), underlining the divine and erotic aspects of such technologies.

In journalistic reporting, the share of female journalists in general, and female journalists in tech reporting specifically, has been a highly debated and regularly monitored issue. Around 35% of American tech news was covered by female writers in 2014 (Women's Media Center, 2014). Besides the fact that top editorial positions have traditionally been male dominated (Andi et al., 2020), even horizontal segregation has been reported, where there are gendered divisions of tasks associated with reporting in terms of topics and the so-called soft and hard news (see, e.g., North, 2016). As AI may, as a broad and heterogeneous topic of coverage, foster a diversity of approaches that may not necessarily align with previous binaries of technology and science, future studies will show to which extent the practices of reporting on AI are gendered.

In 2019, over 71% of STEM characters in Western entertainment media products were white, while Black people comprised about 17%, Asians 6%, Latinx 4%, and those of Middle Eastern descent 2% of STEM characters (Women's Media Center, 2019). AI has also predominantly been portrayed as white in terms of race and ethnicity. Cave and Dihal (2020) suggest that the dominance of the whiteness may partly reflect the predominantly white environments from which the technologies and artefacts arise. However, they also argue that the AI technologies' attributes of being intelligent, professional, and powerful also reinforce the white racial frame that tends to ascribe these attributes to white people. According to Cave and Dihal, AI racialized as white could exacerbate bias by erasing people of colour from the imaginary. Biases may also occur when scientific scholars from different parts of the world study public representations of AI and their datasets are typically retrieved from Anglo-American

journalism, thus reinforcing the Western public sphere as the space for universal representation of AI technologies. It must be observed how AI is being discussed at the local and regional levels, where more situated, innate, and self-referential discourses and representations drawing on local conditions can emerge.

Educating the Audience

The role perceptions of science journalists indicate that making complex information accessible to the public is a shared mission (Brüggemann, 2017; Viswanath et al., 2008). Science and technology journalists aim at building bridges between scientists and experts making discoveries and citizens who want to learn about these new discoveries and their value. To position AI as a fully public problem requires a diversity of voices to balance the perspectives. Brennen and colleagues (2018) suggest that AI should be presented neither through overstatements and sensationalizing content nor as a solution to every problem. Rather, they suggest that journalists negotiate more between different aspects to find a balance:

While media should explore in detail the promise and pitfalls of AI, they would be well served to treat it less as a world-shaking revolution and more as a set of technologies in the process of being designed, as set of choices in the process of being made, and a set of problems in the process of being collectively solved (Brennen et al., p. 10).

The diversification of sources and aspects addressed calls for more collaboration across news desks.

It can thus be said that approaching AI in a balanced, fact-based, and nuanced way requires new literacies from both journalists and audiences. As for many specialized topics of coverage, such as in science, journalists do not need to be experts in the AI field, but they need to be familiar with the topic and act as translators, mediating ideas for a highly specialized technological or scientific area into the everyday lives of the audiences. At the same time, audiences increasingly need skills to understand and cope with media and digital technologies in their personal lives, which has been referred to as *media and information literacies* or *digital literacies* or, more specifically, *technology*, *algorithmic*, or *platform literacies*. To assist individuals in their relationship with new technologies, journalists in a way become the educators of their audiences. Journalists can pinpoint what is new in the products and what is, after all, not; how technologies and applications actually work and what everybody should now about their workings; and they can cast light on the background and implications of the technologies. In this respect, AI dovetails with the ambitions of journalistic approaches that have been recently emerged and developed. These include not only the traditional areas of science and consumer reporting but, above all, *solution journalism* or *constructive journalism*, which seek to focus on responses to problems instead of highlighting shortcomings, risks, and threats. One central question for future development is how AI coverage could be made more participatory, dialogical, reciprocal, and engaging.

AI is currently an area dominated by spot news events, but ideally it should be covered in more in-depth approaches following trends and long-term processes that affect society. If monitored in terms of individual events, which are typically product releases, investment decisions, or initiative launches by major industry organizations, reporting may relapse into following major players' agendas. As influential public actors, journalists can counteract AI-related representation of gender and race, as well as other biased structures, and contribute to increasing diversity in different fields. In this respect, covering AI requires from journalists a proactive meaning-making approach where phenomena are identified and contextualized, rather than reproducing and reacting to news occurrences. Questions to think of as a journalist when covering AI can be the following:

- What kind of AI is the topic about? As AI is a very general term, broad in scope, it often needs to be specified. What is so special about this particular technology? What does it do, and possibly change?
- What is known about how the AI in question actually works? Is it clear to the audience how an app looks or how it is used in practice? Can the workings be illustrated in word (fact file) or images (illustration of the process) somehow?

Patterns of Storytelling about Artificial Intelligence

- Who are the stakeholders of the AI in question, and what are the financial interests in the background of this particular application? Who owns the product, and who has financed it?
- How does the AI affect the lives of different groups of people? How do these people, as citizens and customers, perceive these changes?
- What consequences does the AI have for the target groups of your media?
- What kind of image does your reporting deliver of the phenomenon?

Reporting on Algorithms: The Math is the Least Important Thing



Nicolas Kayser-Bril, reporter at AlgorithmWatch, Germany

“Journalists, as well as politicians and many others, tend to see automated systems as too complex to understand. They are often blinded by the techno-solutionist discourse on Artificial Intelligence and do not dare to start investigating. But the math behind these systems is the least important thing. More important are their effects on society, and whether the systems work at all.

AlgorithmWatch is a non-profit based in Berlin that sheds light on automated decision-making processes that are relevant to society. We do research as well as campaigning and advocacy. We work with journalists across Europe and help them lift the veil on automated systems in use by public authorities, for example in welfare services, schools or the police, and by private companies, such as credit scoring or content management by social media platforms.

We are aware of the fact that many systems are very complex and reading through the code or the technical details without a background in math or statistics is very hard. But the technical intricacies are relevant to statisticians, not to society. Among the very few metrics that are relevant to society, there are the false-positive and false-negative rates. These can be understood in minutes by watching a tutorial. And if a journalist is still feeling coy about these terms, it might reassure them that researchers (principally from Gerd Gigerenzer’s lab) have shown that many people who think of themselves as experts do not understand these concepts very well.

There’s a mistaken belief – even among journalists – that highly technical systems can only be understood by highly technical people. On the contrary, the code and the algorithms themselves are only a very small part of these systems. More important is who built them, to what end, and what effects they have on society. Journalists are in a much better position to answer these questions.

I once saw a journalist complain on Twitter that her application at the welfare management agency had been rebuked one day after she sent in a whole batch of documents. To me, this was a red flag that an automated system, and not a human, had processed her file. I asked her if she would investigate, which she did, and this led to many more stories on the topic, published at AlgorithmWatch and beyond.

Another time, an Instagram content creator told us she felt she did not reach her followers when she posted pictures where she was fully clothed. She had to be in a bathing suit just to be shown to her followers, she said. We wanted to check if that was true. From a statistical perspective, attempting to understand how Instagram’s newsfeed algorithm worked was a fantastic challenge. That’s why we worked with professional statisticians. The rest was traditional journalism: talking to affected people and to the men who designed the system. When we joined both ends of the investigation, we had a powerful and insightful story.”

AlgorithmWatch’s homepage: <https://algorithmwatch.org/en/>




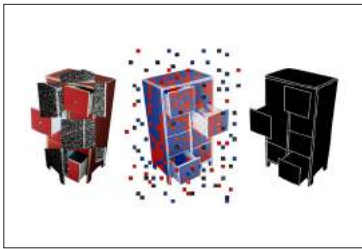
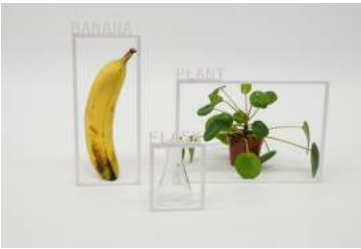
Suggested Assignments



READ AND REFLECT: Professional magazines and community sites regularly publish articles providing journalists with advice on how to report on AI in an appropriate way. Search for such online texts, often written by journalists to journalists, and collect the pieces of advice together, possibly grouping them thematically. Based on the pieces of advice that you have found, create a communication charter for journalists – a practical one-pager that outlines the key principles of covering AI in journalism. Present your list of advice to your peers, discussing the challenges and comparing the differences, if possible.

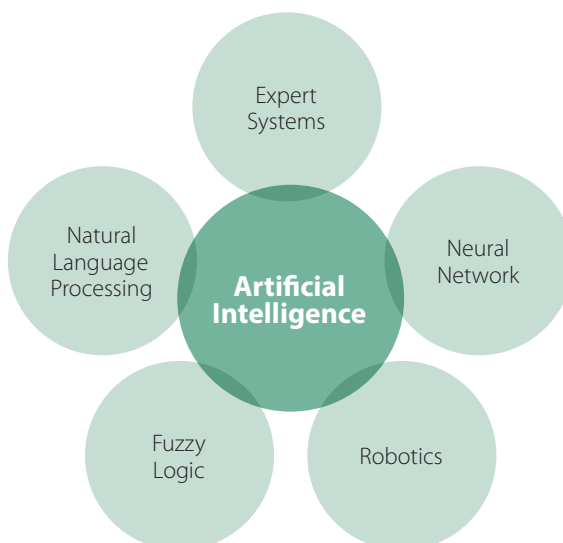


DISCUSS: Take a look at the images below, illustrating AI. Try to verbalize what you see in the picture. Thereafter, go to the website and get a description to compare with your own interpretation. Did you catch the idea from the image, and if you didn't, why was that? How would you illustrate the idea in another way? Which search terms could be used, instead of "artificial intelligence", to retrieve appropriate images?

Quantified Human	Classification Cupboard	Banana / Plant / Flask
		
<p>Alan Warburton, © BBC, Better Images of AI (CC-BY 4.0) https://betterimagesofai.org/images?idImage=0</p>	<p>Anton Grabolle, Better Images of AI (CC-BY 4.0) https://betterimagesofai.org/images?idImage=5</p>	<p>Max Gruber, Better Images of AI (CC-BY 4.0) https://betterimagesofai.org/images?idImage=15</p>



IDEATE: To reach out to academics working in the AI field, you need to find experts in different academic fields. Try to find scientific experts for a hypothetical interview in the fields shown below at universities near you. Find one expert in each area, identify what kind of societal relevance their research might have, and formulate 1–2 interview questions to ask them. Discuss with a peer what challenges you experienced.



Original material by Authors



SEARCH AND SOURCE: Choose one of the following, depending on whether you want to focus on text (a–b) or image (b):

- Pay attention to headlines in newspaper articles on AI. You can search for 5–6 examples in an online database in your country or region or a news outlet of your choice, and/or take a look at the examples from major news outlets below. List the 20 most recent headlines. What news values typically apply to AI as a journalistic topic? Are you content with these news values or should some other values be highlighted instead? Which, why, and how?
- Explore a selected social media platform (e.g., Twitter) by searching with a chosen hashtag (e.g., #ai, #artificialintelligence, #deeplearning, #machinelearning) and find out how different user groups tend to talk about AI. Take 100 recent posts using the hashtag and group them according to what they represent (industry companies, media/journalism, NGOs, ordinary people, etc.). Group the posts also according to whether they represent a positive, negative, or neutral attitude towards AI (if that can be decided). What differences do you recognize between different producers of information? Can you identify some other patterns within the groups?
- Examine how AI is illustrated in news articles today. Search for news articles on AI in a news outlet that allows articles to be categorized under this topic (e.g., *The Guardian's* topic tag “artificial intelligence (AI)”), or, alternatively, search for stock photos with the search term “artificial intelligence”. Pay attention to the images retrieved. How has AI been typically illustrated and what kind of associations do the pictures deliver? Do the images represent *Better Images of AI's* ideas or differ from them? Can you figure out alternative ways of illustrating to avoid potential biases, such as the gender bias, or mythologization of technology?



PRODUCE: Develop a story synopsis of one of the following topics by filling in the table below: a) female AI advocates; b) female technology journalists; c) people of colour and AI; d) challenges of AI policy implementation; e) misuses of a particular type of AI; f) inequalities in the uses of AI; g) a question related to AI that personally occupies you. For the synopsis, conduct background research on scientific journal articles and journalistic coverage on the topic.

Topic: formulate your topic.

Medium: choose a medium.

Pursued length: assess the length of the story in its final form.

Area	Clarification	Your answer
Main argument	What is the main message you want to convey? Formulate a statement.	
Genre	Which genre do you apply?	
Intended audience	Describe the expected receiver of the story.	
Intended effect	What is your story expected to change in the audiences' attitudes and in the world?	
Positive arguments	What are the main arguments for your statement?	
Negative arguments	What are the main arguments against your statement?	
Sources	Who will you interview? Try to discover a real person.	
Case	Which AI application(s) or initiative(s), etc., do(es) your story deal with?	

Area	Clarification	Your answer
Title	Propose a title that includes your message and arouses interest in the topic.	
Lead	Write 2–3 sentences	
Motivation	Why does this story matter (in terms of AI)?	
Self-reflexivity	What do you feel as challenging when reporting on your topic? Why?	



Readings

- Angler, M. W. (2017). *Science journalism: An introduction*. Routledge. <https://doi.org/10.4324/9781315671338>
- Bains, S. (2019). *Explaining the future: How to research, analyze, and report on emerging technologies*. Oxford University Press. <https://doi.org/10.1093/oso/9780198822820.001.0001>
- Cave, S., & Dihal, K. (2020). The whiteness of AI. *Philosophy and Technology*, 33, 685–703. <https://doi.org/10.1007/s13347-020-00415-6>
- De Vuyst, S. (2020). *Hacking gender and technology in journalism*. Routledge. <https://doi.org/10.4324/9780429262029>
- Johnson, D. G., & Verdicchio, M. (2017). Reframing AI discourse. *Minds & Machines*, 27, 575–590. <https://doi.org/10.1007/s11023-017-9417-6>
- Natale, S., & Ballatore, A. (2017). Imagining the thinking machine: Technological myths and the rise of artificial intelligence. *Convergence*, 16(1), 3–18. <https://doi.org/10.1177/1354856517715164>
- Sudmann, A. (Ed.). (2019). *The democratization of artificial intelligence: Net politics in the era of learning algorithms*. Transcript Verlag.
- Sun, S., Zhai, Y., Shen, B., & Chen, Y. (2020). Newspaper coverage of artificial intelligence: A perspective of emerging technologies. *Telematics and Informatics*, 53, 101433. <https://doi.org/10.1016/j.tele.2020.101433>
- Verdegem, P. (Ed.). (2021). *AI for everyone? Critical perspectives*. University of Westminster Press. <https://doi.org/10.16997/book55>



© George Morina, Pexels.com

Module 7:

Artificial Intelligence in Journalism and Journalistic Practice



Summary

This module briefly introduces the use of AI in journalism and journalistic practices, for example, in machine-written articles and robot journalism, and calls for critical self-coverage and reflection on ethics, as well as best practices in editorial decisions in news gathering, production, and distribution. It also encourages educators and teachers to make innovative uses of AI in journalism to enhance transparency, creativity, and problem solving.

Key concepts: *AI in journalism, algorithmic journalism, automated journalism, computational journalism, robot journalism*



Module Aims

The aims of this module include

- ▶ Understanding the ways AI is being used and can be used in future for journalism.
- ▶ Understanding the potentials and risks that AI entails for journalism and journalistic practice.
- ▶ Understanding the human value of creating journalism.



Learning Outcomes

After the completed module, the student is expected to have learned the following:

- ▶ To identify AI-powered technologies that are the most relevant for journalistic practice and understand how they work.
- ▶ To know to which journalistic genres and tasks AI can the most appropriately be harnessed.
- ▶ To know how the journalistic community deals with AI technologies in terms of professional and ethical values as well as be aware of some contemporary best practices.



Questions for Reflection

Here are some questions to reflect upon in advance:

1. In which ways are computers assisting journalists in newsrooms nowadays, according to your own experience?
2. What kind of AI-related terms of “journalisms” can you identify? Please make a list of such technology-related terms referring to journalism aided by computer and technologies by using the database from the project X Journalism (<https://xjournalism.org>) by the Leibniz Institute for Media Research and Hans-Bredow-Institut in Hamburg. What do the scholars mean by the term “X Journalism”?

Here are some preparatory tasks to do prior to entering the topic:

1. Talk to a journalist and ask them about their experiences of using AI in journalistic practice.
2. Take a look at the following robot-written news texts and discuss whether a human would have written some parts in a different way:
 - ReporterMate. (2019, January 31). Political donations plunge to \$16.7m – down from average \$25m a year. *The Guardian*. <https://www.theguardian.com/australia-news/2019/feb/01/political-donations-plunge-to-167m-down-from-average-25m-a-year>
 - GPT-3. (2020, September 8). A robot wrote this entire article. Are you scared yet, human? *The Guardian*. <https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3>



Outline

By Jenny Wiik

In journalism, there is a lot of buzz about the possibilities of new technology, and not least AI and automation. Many see this as a lifeline at a time when the industry is weak, while others see it as a threat that mechanizes a creative profession and pushes commercialization further, perhaps even causing redundancies. Not surprisingly, there are some justifications for both views: The new technology brings opportunities to journalism but also challenges. One thing is certain: AI will imbue and change journalism at its core – in every part of the production chain. The data-driven development of journalism, increasingly sustained by AI technology, means breaking and renegotiating established business models, professional expertise, and future aims.

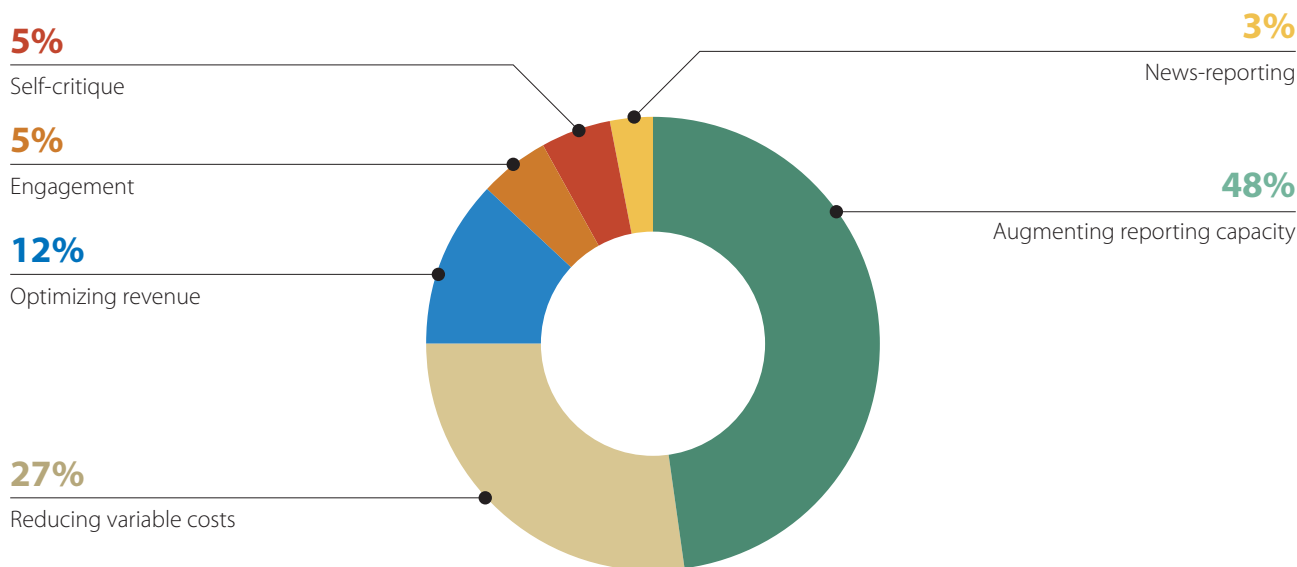
This process may therefore be regarded as a “creative destruction” of journalism. The “gale of creative destruction” describes, according to the economist Schumpeter (1994, pp. 82–83), the “process of industrial mutation that incessantly revolutionizes the economic structure from within, incessantly destroying the old one, incessantly creating a new one”. Such processes are in many respects painful, but – according to this perspective – also necessary, unavoidable, and revitalizing. Is the automation of news production part of a “creative destruction”, a process of transformation where the human–machine hybridity becomes the heart of the profession? And if so, what does it mean for journalism, as a profession and as a democratic institution?

This chapter introduces the main advantages of AI to journalism, as well as the most important challenges. It contextualizes the technological development with a professional perspective but also discusses what this means to media houses and leaders.

How Can AI be Used in Journalism?

AI is a collective name for a variety of data-driven algorithms, robots, and systems, with the common denominator that they can replace human labour. Some algorithms are self-learning and can improve their precision as the data increases, but in many cases, it is about static codes performing the same thing every time. This means that robots can conduct much of the mundane work of journalists, potentially freeing human resources to do more qualitative tasks. In his book *Sådan forandrer automatiseringen medierna* [How automation changes the media], Danish journalist Andreas Marckmann Andreassen (2020) writes that it is now possible to identify the use of AI in all blocks of the journalistic production chain. However, some parts of the process have developed further than others. Most common automation practices include data collection and research, as well as auto-generated news items.

A global survey of 130 AI projects, conducted by the Knights Foundation in 2021, showed that the primary purpose of implementing AI in journalism is to augment reporting capacity, and secondly to reduce costs.

Figure 17: Primary Purpose of AI in Journalism (2012–2021)

Source: Composed by authors, based on *The Presence and the Potential of AI in Journalism*, Knight Foundation 2021.

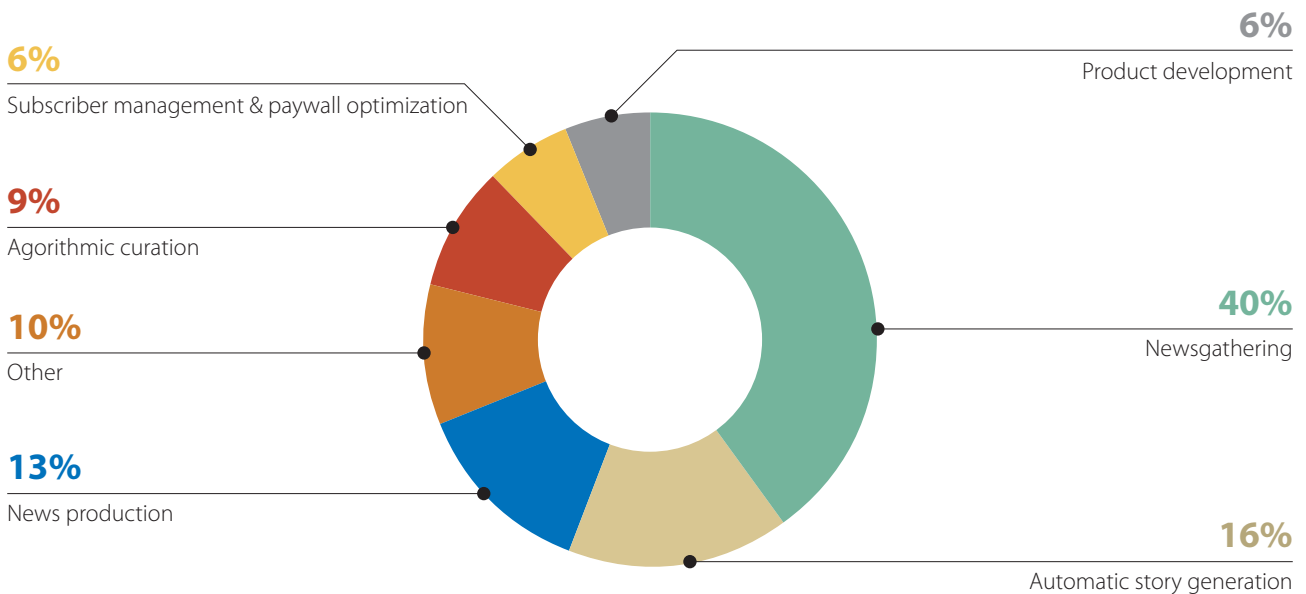
One example of how AI can contribute to information search and research is Reuters' algorithm News Tracer, which automatically searches for breaking news on social media. By detecting clusters of similar information and reviewing the accounts that published the information, the robot can notify a journalist about a news item the moment it occurs (Xiaomo et al., 2017).

Another example is BBC's Juicer (<https://bbcnewslabs.co.uk/projects/juicer/>), which functions like a news aggregation 'pipeline'. It ingests news articles and extracts the best from them – just like a juicer does with fruit. BBC Juicer pipeline is watching RSS feeds of news outlets. When a new article is published on one of these RSS feeds, BBC Juicer scrapes the news article, both raw text and metadata (e.g., date, time, title, news source). In the next step, BBC Juicer identifies and tags concepts mentioned in the article text, making them searchable and therefore useful for trend analysis. By streamlining media workflows like that, AI enables journalists to focus on what they do best.

AI is currently being implemented in most European and North American newsrooms, but other regions are also part of this development. Munoriyarwa, Chiumbu, and Motsaathebe (2021) investigated the practice of AI in South African newsrooms, finding that although some media houses are technologically advanced, the adaptation of AI is slow overall. The optimistic attitudes to AI in Western newsrooms are not echoed by South African journalists, whose scepticism is driven by "fear of job losses, ethical issues around AI, its efficacy in the democratic process and the costs of adopting AI for newsrooms in Africa" (Munoriyarwa et al., 2021). They point out that in developing countries, the meaning of new technology may be different than in the Global North, such that automation of an already unstable profession may increase the risk of losing autonomy and control. On the other hand, in countries with weak democracy and state-controlled news, algorithms may be regarded as means to avoid bias and false news. An experimental study of Korean news consumers showed that they credited robot-generated news items with higher accountability than those produced by human journalists – a reaction most likely caused by the public's negative attitude toward journalists' credibility and a craving for new ICT products and services (Jung et al., 2017).

Looking more closely at the story pipeline, AI is predominantly used as a tool to gather news, and generate stories automatically. This is often done in the form of short data-driven notices within real-estate sales, sports, or economy.

Figure 18: Where in the story pipeline is AI commonly implemented? (2012–2021)



Source: Composed by authors, based on *The Presence and the Potential of AI in Journalism*, Knight Foundation 2021.

One example of this is the Swedish company United Robots, which delivers text-generating solutions to several news firms. They use AI and Natural Language Generation (NLG) to automatically produce publishable news texts from large datasets, including sports, property sales, traffic news, weather, the stock market, and local business registrations. Their robots analyse big datasets that are open and available to everyone and identify unusual events or patterns. The information can then be used to build extended articles or as alerts for the newsroom to follow up on (United Robots, 2021). Many publishers have also built gender analyses and other values into these publishing tools, allowing automated monitoring of their own news output.

AI is also fostering a new approach to content management more generally. By applying semi-automated publishing tools, where AI is used to combine modules of content in a multi-faceted way, news is being reshaped into a more personalized and changeable form. This means that authoring is no longer based on a “story” but instead around “nested blocks” that allow better connections across stories, making it easier to reassemble content in potentially limitless ways (Leitner, 2018).

Another growing area for machine learning applications is the use of machine learning for churn prevention. By automatically collecting and analysing data about user behaviour, market analysts may work directly towards the reasons for why users are likely to churn, that is, stop using an app, which typically occurs when users simply lapse in use or uninstall the app from their device.

In sum, AI has several potential positive applications in journalism:

- better personalized distribution of content
- more efficient, automated production of content

- dynamic pricing, both for ads and subscriptions
- ability to find more stories in data, and more data in stories
- better automated transcriptions
- manageable content moderation
- recognition of fake news and deep fakes
- new tools for debunking
- enhanced image and video search
- deeper sentiment analysis on user-generated content (UGC)

For the news business, which has been struggling with financial difficulties and depleted business models for years, the possibilities of doing more with less is exciting. Especially in local markets with shrinking newsrooms, AI opens opportunities for improved digital coverage of local politics and economy.

The Organizational and Market Context

The development and implementation of AI in journalism is part of a broader effort to rebuild business models and innovate in newsrooms. Many media houses have for a long time struggled with declining advertising revenues, mostly because of the exploding competition for user attention in online environments. During the Covid-19 pandemic, the gradual decline accelerated and dropped drastically. For American news media, as an example, the advertising revenues went down by around 30% between 2019–2020 (Adgate, 2021). Simultaneously, online subscription revenues went up as the health crisis launched a massive need for accountable information. Rapidly, market and wider societal tendencies have moved the news industry further into the “gale of creative destruction”, with new technology playing a key role.

Considering these changes as “creative destruction” emphasizes newsroom innovation not as a strategy or an end goal, but as a process – as a “series of dynamics, mechanisms, means, and changes that lead to a particular outcome” (Siles & Boczkowski, 2012, p. 1386). The cross-sectional multi-platform news production model which has emerged during the past decades got an extra spark due to the changing economic conditions described above, but there has also been an increasing interest and investment in newsroom innovation. Based on an extensive literature review, Paulussen (2016) underlines the complex interplay between structural and individual, and human and non-human, factors and actors staging the processes of innovation in journalism. The impact of AI on journalism thus calls for a holistic perspective of the organization and production of news.

Collaboration is key

Collaboration is an important factor for successful automation of any business, and journalism is no exception. Collaboration involves sharing resources and utilizing them where they are needed and leads to new ways of combining skills, abilities, and material assets. Novel contact surfaces lead to a better flow of information, new insights into problems, and the opportunity to learn from the mistakes of others (Jia & Xia, 2008). To meet the growing complexity of the outside world and the market, media companies are increasingly trying to organize themselves based on principles of cooperation. Internal boundaries that previously separated departments, such as marketing and editorial, from each other are now eroding at a fast pace, as companies move to project-based and interdisciplinary working methods. Although these approaches are relatively new to journalism, they have been applied in other knowledge-intensive sectors for decades. Research from, for instance, healthcare has shown that strategies for achieving comprehensive goals require both cooperation and meaningful commitment (Winfield et al., 2017). When it comes to product and business development, the media industry

has a tradition of monitoring its close competitors, but rarely extends the view to other industries. An American media manager says:

I think journalism should look at ALL other industries. In fact, I worry that journalism as an industry looks too often for guides within journalism itself. What is the New York Times doing? What is the Washington Post doing? We need to look at other industries and see how their innovations, stumbles, and concerns could apply to us. (Beckett, 2019, p. 86)

“Strangers” in the newsroom

Innovation and new technology thus require new skills, and a set of new players has now entered the newsrooms. The question of how the new actors and their expertise are being integrated into established news processes is relevant from a journalistic perspective (e.g., Eldridge, 2018; Holton & Belair-Gagnon, 2018). Research shows that outsiders who enter the journalistic field are sometimes treated as “unwelcome strangers” (Holton & Belair-Gagnon, 2018, p. 71). However, Eldridge (2018) could see a spectrum of reactions among journalists, from resisting to embracing change. He believes that the key to progress lies in the normalization process of journalists adopting new technology and incorporating it into their norms and routines. To achieve this normalization, cooperation is required within the editorial staff and between the editorial staff and external parties. Successful collaborations have, for instance, involved designers and journalists to strengthen the journalistic narrative (Doherty, 2016), or technicians and journalists as in the open-source network Hacks / Hackers (Lewis & Usher, 2014). The latter was founded in 2009 with the aim of rethinking news and information. It is based on an equal model of collaboration, where journalists work with an open attitude together with both technicians and audiences, resulting in a fruitful hybrid logic where different skills and interests are taken advantage of (Lewis, 2012).

Journalistic renewal often comes from the side and seeks its way into the centre as “newcomers” bring added value to news production in the form of innovative practices and approaches. Holton and Belair-Gagnon (2018, p. 72) state that “strangers” in the journalistic field (e.g., web developers and programmers) are “importing qualities to it that do not originally stem from the journalistic profession [and have] helped to introduce new ways of identifying what news is, how to deliver it more effectively, and how to better engage with news audiences”. A study of these new actors’ own attitudes shows that they have a good understanding of journalism as a field and its conditions, but that they also carry an identity based on a “Silicon Valley ethos”, including 1) focus on the user experience, 2) a desire to explore new applications for data, and 3) a drive to develop easy-to-consume products (Wu et al., 2019). Although the clear market logic opposes the professional jurisdiction of journalism, there is also common ground, especially in the striving to meet user demands and to make journalism more relevant.

What Are the Challenges?

A growing body of research addresses the automation of journalism, showing that machines are indeed transforming the news process in several ways. Based on an extensive literature review of this research, Thurman, Dörr, and Kunert (2017) suggest several aspects where the influence of algorithms and increasing use of digital data present ethical challenges. These include how journalists acquire, validate, and use digital data in news production (Bradshaw, 2014) and the risk of bias in the algorithms that power automation (Carlson, 2015; Gillespie, 2014; Thurman et al., 2016). The transparency of code and data is of course an issue, both from a professional and a democratic perspective (Diakopoulos, 2015; Diakopoulos & Koliska, 2017). It is also a question of machine autonomy: whether algorithms can “reason” and draw appropriate conclusions in all contexts (Young & Hermida, 2015).

From a professional perspective, Carlson (2015) points to some areas of potential concern: the future of journalistic labour, the rigid conformity of news compositional forms, and the normative foundation of journalistic authority. He concludes that although the growth of automated journalism expands the amount of available news and frees journalists to pursue fewer mechanical stories, it conversely involves negative predictions such as increased layoffs, polarizing personalization, and the commoditization of news writing.

Asking journalists about their perceptions of working with automation, Thurman, Dörr and Kunert (2017) identify several limitations:

1. Its reliance on single, isolated data streams.
2. The one-dimensional nature of the quantitative data feeds it relies on.
3. The difficulties of interrogating that data.
4. The lack of human angles in the texts generated.
5. The requirement to template stories – predicting “top lines” – in advance.
6. The difficulty of working creatively with the data in the templating process.

(Thurman et al., 2017, p. 1254)

Ethical and professional challenges aside, there are also organizational challenges of developing and implementing AI. Lack of will, level of knowledge, or resources are the most common, but also organizational structures, difficulties in recruiting the right skills, and lack of strategies are factors that come into play (Beckett, 2019). AI is not only about technology but also about the organization’s ability to take advantage of technology. Although there are variations, the AI maturity of news organizations is often low. There is a lack of strategic thinking and active leadership in this area, often due to low knowledge and awareness of new technologies. New innovations are many times initiated by tech personnel in IT departments, who may find it difficult to truly immerse them into organizational activities. New technology may be met with scepticism, as both professionals and organizations strive for status quo, and it can be difficult to find resources and the right skills. Therefore, the issue of AI is not just an issue for IT departments, but it very much concerns the management and editors as well.

How, then, should media companies act to succeed better with their AI strategies? Beckett (2019) gathers a few pieces of advice from his study of media innovators and editors, which could be used as point of departure for increased self-reflection and transparency for reporting when AI is involved:

- Create an understanding of what type of AI you need, what problem it should solve, and how it relates to other editorial strategies and policies.
- Which parts of the organization are affected and need to collaborate on this?
- Identify the obstacles that may arise, such as resources, culture, and management, and address these in a systematic way.

Beckett (2019) concludes by pointing out the importance of collaboration, clear roles, and credible ways of measuring what one wants to achieve.

AI Transforms Journalism

These days, journalists not only have the company of users, bloggers, citizen journalists and the like, they also share workplaces with programmers, user experience designers, and algorithms. This has led to a reassessment of journalistic skills and a renegotiation of professional identity (van Dalen, 2012), especially as audiences cannot always tell the difference between robot-generated material and texts written by humans (Clerwall, 2014). What does this mean for the professional identity of journalists?

While journalists have historically sought to resist attempts to standardize or rationalize their operations, these acts have become increasingly unavailing. Analytics dashboards constantly monitoring their performance and ranking them against their peers have emerged as a ubiquitous part of news work and as a managerial tool (Petre, 2018). Apart from the challenging logics of quantification, the new technological support systems may also convey changing professional standards of journalism, for instance, by algorithmization of news valuation. A study of the suppliers of web analytics shows that, although they actively seek to understand the fundamentals of journalism, they tend to foster profit-oriented norms and values in newsrooms by introducing disruptive, connective, and routinized technology (Holton & Belair-Gagnon, 2018). Carlson (2018) conceptualizes this tendency as “measurable journalism” and argues that the desire of industrialized news media to monitor its audience will have irreversible effects on journalistic judgement and autonomy.

The journalistic gut feeling – or a nose for news – has for a long time been the very foundation of the journalistic profession but is increasingly perceived as obsolete by news editors and publishers (Milosavljević & Vobič, 2019). Traditionally, news valuation of a social event is built on one or more criteria. These can be, for example, that it is about power elite, conflict, sensation, celebrity, or is affecting a large part of the population (Harcup & O’Neill, 2017). Such assessments make the core of journalistic intuition, that is, what is deemed important and interesting to cover. Quite contradictory to the professional imagination, there are many successful attempts to translate this intuition into algorithms. One thing that robots can easily learn is to detect deviation from a norm or expectation: a feature which has been used, for instance, to generate news about earthquakes over 3.0 on the Richter scale (LeCompte, 2015) or to scan the Swedish criminal record for deviations in comparison with historical data and then notify a journalist (Magnusson et al., 2016). But algorithms can also be used to work more closely with news valuation standards and praxis within an organization. One example is the “public service algorithm” News Values developed for Swedish Radio (SR). The purpose of the algorithm is to personalize the news feed based on news value ratings made by the journalists, both to create a more relevant news feed but also to streamline the organizational news valuation standards. In the development process, employees from all over the company were engaged in identifying and formulating the core values which they believed characterize strong public service publications. Olle Zachrisson, digital strategist at SR, says: “The system of news values stimulates a discussion about how we do our journalism. It inspires positive change” (Zachrisson, 2021).

The professional logic of journalism has always been negotiated towards organizational goals, technical innovations, and more recently, an increasingly strong management culture (Waldenström et al., 2019). Historically, technological innovations have been integrated with the professionalisation of journalism and supported the consensus between journalists, audiences, media owners, and other stakeholders that granted journalism its professional legitimacy (Nerone, 2013, p. 452). However, innovation work in the media industry has so far been relatively slow, and the applications of automation and AI introduced are seen as subordinate to the human actors (Beckett, 2019). In interviews with innovation strategists, editors, and software developers for news production, similar perspectives emerge, namely that journalists are still central to the business. Most AI that is currently applied in newsrooms is not intelligent enough to be able to replace more skilled work, but even if it was, there is a reluctance to abandon human control of the process. Milosavljević and Vobič (2019) show in a study of editorial managers that those would rather emphasize the liberation of human labour

than an actual replacement of it. The authors interpret this as a strategy for balancing professional values and organizational goals against each other.

Creative Destruction of Journalism?

AI offers a range of opportunities for journalism, and newsrooms are increasingly adopting different solutions. While research has shown the human–machine interface to be surrounded by tensions, for instance, between competing competences and logics, the general tendency is now to move into more automation throughout the production chain. New technical solutions extend the reach, frees journalists to do more qualitative tasks, and improves the curation of news to users. The data driven development is already changing the fundamentals of news routines, in the sense that quantification and measurability have become predominant logics.

Does this mean the “creative destruction” of journalism? Considering the transformative impact of data and AI, the answer would be yes. Technological development is reframing established business models, incorporated routines, and professional values and ideals. This window of change calls for a recapitulation and revaluation – not only of data and technology, but also of the human capital, the journalists. Relocation of organizational resources raises opportunities for higher work being done by humans. AI provides new tools for journalists to elevate their investigations and produce unique, original content. It may also be applied as a tool for monitoring content to ensure that it meets editorial quality standards. But to use AI for augmenting journalistic methods, reporters must learn more about the possibilities. Up to this point, responsibility for newsroom innovation is often located with IT departments (Westlund et al., 2021). To ensure a professionally and financially sustainable development of journalism, editorial staff and management must get involved too. The main reason is that “the core elements of journalism – story telling in different formats, critical thinking, verification, human values, ethics, autonomy and integrity – become even more important at a time when “smart” machines are entering every part of society and professional life” (Lindén, 2020). The professional identity of journalists is likely to change as the fundamental conditions change, but the development should not be seen as a threat. Rather, it indicates an upvaluation of the tasks, traits, and functions that are *not* possible to automate. It is around these phenomena that journalism will refine, in close collaboration with algorithms and robots.

However, the analysis of AI implementation in journalism must be set into the cultural context. While the development of AI technology is first and foremost driven by Western countries, the professional, democratic, and organizational conditions of journalism in developing countries are often completely different. In such a profound transformation that is being brought about by AI, those aspects must be carefully considered. As some journalistic communities struggle to establish professional standards and autonomy, new technological solutions must be built to support those aims.

As technology advances, so too must journalism education. With an increasing number of journalistic tasks being automated, it is essential that students are taught how to use these new tools to remain competitive in the industry. Automation is already having a significant impact on the business of journalism. For example, many news organizations are now using bots to write simple stories, such as sports scores or weather reports. While this frees up reporters to focus on more complex stories, it also means that there is less demand for entry-level journalists.

As a result, journalism schools need to adapt their curricula to prepare students for a future in which automation is increasingly prevalent. This could involve teaching students how to use bots and other automated tools, as well as how to identify stories that are best suited for automation. In addition, schools should place a greater emphasis on teaching the soft skills that are needed to succeed in an automated news environment, such as creativity and adaptability. By adapting their programmes to the increasing automation of the business, journalism schools can ensure that their students are prepared for the future of journalism.

AI in African Newsrooms: Promises and Pitfalls



Ammina Kotari, Professor in Journalism and Director of Harrington School of Communication & Media, University of Rhode Island, U.S.

The discourse on AI technologies is very much centered on the Global North, and this applies to research and its impact. Based on my research, very few newsrooms in African countries have integrated AI technologies in their work. In a recent article, “Artificial Intelligence and Journalism: An Agenda for Journalism Research in Africa”, I and my colleague Sally Ann Cruikshank (see Kothari & Cruikshank, 2022) discussed the status of African newsrooms when it comes to using AI. Through systematic searches in databases, we wanted to discover cases from Africa and increase the knowledge about the African initiatives, as the innovation of AI in newsrooms is led by Western countries and China, and, consequently, the discussions on AI in journalism have predominantly focused on the Western perspective.

Some notable examples of AI use on the African continent include Morocco, where the project “Fake news à l’épreuve des faits” worked with investigative journalists to train them in using AI to assist in fact-checking and data journalism. In 2020, the South African news organization News24 announced it would use AI to moderate comments on its stories (News 24, 2020). Bloomberg South Africa has incorporated AI tools into their news practices, and the South African TV station MSNBC also uses AI tools to manage their newsgathering equipment, including cameras.

Some promising AI work is currently being done in some African countries, including two IBM Research labs – one in Nairobi, Kenya, and another in Johannesburg, South Africa – along with a Google AI lab in Accra, Ghana. A more recent one in Tanzania, the AI learning lab and incubator #AIForGood, focuses on the social good, advancing open-data solutions to strengthen the national AI ecosystem. There are also other machine learning initiatives across the continent.

A lot of research has focused on the positive impact of AI on various industries. Still, there are some serious policy implications for countries with limited press freedom and even authoritarian governments. AI has been touted as a means for improving economic growth, social progress and governance across the continent, but the potential of AI may also be harnessed by states that seek to suppress or monitor political opposition or marginalized groups.

The AI and journalism discourse in the Global North gives the impression that AI integration would be the norm in many newsrooms. Still, when you look more closely, you realize that the integration is primarily done at national or big-market news organizations. Many small or mid-size news organizations in the Global North face similar resource challenges that hinder their ability to integrate AI tools in their newsrooms. There is a bias on the positive implications of using AI in journalism. Still, more research and conversations are needed about using AI tools for newsgathering, especially when it comes to in-built biases of privileging viral stories and prominent sources in digital spaces. As we know, advances in AI technology have also increased the creation of deep fakes, making it harder for news readers to differentiate between factual information and disinformation.

To reduce the technological divide between the Global South and Global North, some tech companies and scholars are working on transferring the technology or training journalists to use tools developed in the West. But this approach does not really solve the inequity issues. As the biases built into the AI tools developed in the Global North will be amplified in the Global South, especially without local expertise to mitigate and customize the applications. The solution would be to train Africans to develop their own tools that can be scalable and reflect professional and cultural norms in their countries.

In sum, while AI integration in various sectors is promising, there are some specific challenges to integrating AI tools in newsrooms in the Global South. One is the lack of training for journalists and resources to hire

programmers who can help build AI tools in the newsroom. Second, integrating AI tools requires financial resources to develop and maintain technology infrastructure to support the effective use of automated tools; many newsrooms operate with slim budgets. Third, journalists' skepticism about AI tools being able to make journalism better, and finally, concerns about human jobs being replaced by automated journalism, will be a challenge for integrating AI tools in newsrooms.



Suggested Assignments



READ AND REFLECT: Read one of the below articles addressing the possibilities and limitations of AI in a specific genre of journalism, and search for recent examples of uses in your own geographical area. Choose between a) investigative reporting, b) sports journalism, c) science journalism, d) immersive journalism, or e) visual journalism.

Stray, J. (2019). Making artificial intelligence work for investigative journalism. *Digital Journalism*, 7(8), 1076–1097. <https://doi.org/10.1080/21670811.2019.1630289>

Galily, Y. (2018). Artificial intelligence and sports journalism: Is it a sweeping change? *Technology in Society*, 54(1), 47–51. <https://doi.org/10.1016/j.techsoc.2018.03.001>

Tatalovic, M. (2018). AI writing bots are about to revolutionize science journalism: We must shape how this is done. *Journal of Science Communication*, 17(1), E. <https://doi.org/10.22323/2.17010501>

Sánchez Laws, A. L., & Utne, T. (2019). Ethics guidelines for immersive journalism. *Frontiers in Robotics and AI*, 6, Article 28. <https://doi.org/10.3389/frobt.2019.00028>

Gynnild, A. (2014). The robot eye witness: Extending visual journalism through drone surveillance. *Digital Journalism*, 2(3), 334–343. <https://doi.org/10.1080/21670811.2014.883184>



DISCUSS: Go through the national ethical code for journalists and discuss how the use of a chosen AI technology or methodology in journalism may alter the ethics. Alternatively, you may analyse the Society of Professional Journalists' (SPJ) Code of Ethics, available in different languages (<https://www.spj.org/ethicscode.asp>).



IDEATE: Identify working strategies for the following procedures in terms of where to find open data, which software to use, and what to do with the software. In addition, please reflect what kind of competence requirements the methodology creates for you and how you could pursue them: a) finding the most frequent words, phrases, and topics in governmental documents (e.g., textual analysis, topic modelling); b) visualizing statistics on a map (e.g., data visualization); c) automatic collection of news topics from different sites (e.g., aggregation).



SEARCH AND SOURCE: How could interactive robots – related to the so-called conversational AI – be of help to journalists in their work engaging with audiences? Read the papers below on conversational AI and journalism and discuss how the ideas of dialogical or reciprocal journalism can be fulfilled but possibly undermined by robot-aided journalism.

Gao, J., Galley, M., & Li, L. (2018, June 27). Neural approaches to conversational AI. *Proceedings for SIGIR18: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, pp. 1371–1374. <https://doi.org/10.1145/3209978.3210183>

Shin, D. (2021, March 12). The perception of humanness in conversational journalism: An algorithmic information-processing perspective. *New Media & Society*. Advance online publication. <https://doi.org/10.1177/1461444821993801>

Veglis, A., & Maniou, T. A. (2019). Chatbots on the rise: A new narrative in journalism. *Studies in Media Communication*, 7(1), 1–6. <https://doi.org/10.11114/smc.v7i1.3986>

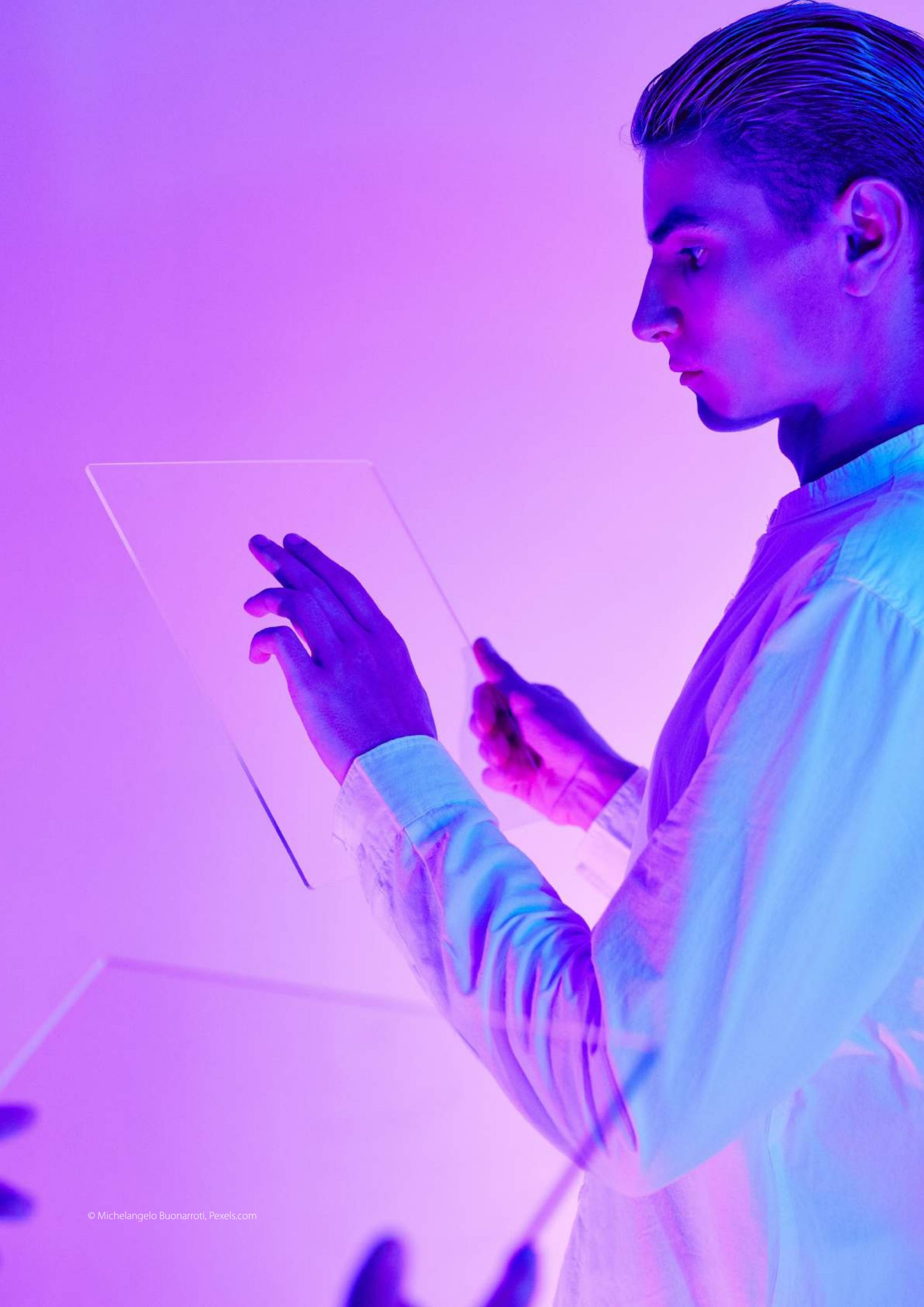


PRODUCE: Make a Gallup interview round with ordinary people, asking 5–6 randomly or purposely selected individuals how much they would trust a robot creating a news text, and what kind of ideas they have on AI-generated journalism. What kind of misunderstandings of journalism and AI possibly occur? Based on the interviews, how should AI journalism be clarified for general audiences? What makes people trust journalism?



Readings

- Beckett, C. (2019). *New powers, new responsibilities: A global survey of journalism and artificial intelligence*. Report from POLIS journalism and Society, think tank at the London School of Economics (LSE). <https://www.lse.ac.uk/media-and-communications/polis/JournalismAI/The-report>
- Bradshaw, P. (2014). Data journalism. In: Zion, K., & Craig, D. (Eds.) *Ethics for digital journalists: Emerging best practices*. Routledge, 202–219. <https://doi.org/10.4324/9780203702567>
- Broussard, M. (2015). Artificial intelligence for investigative reporting. *Digital Journalism*, 3(6), 814–831.
- Broussard, M., Diakopoulos, N., Guzman, A. L., Abebe, R., Dupagne, M., & Chuan, C.-H. (2019). Artificial intelligence and journalism. *Journalism & Mass Communication Quarterly*, 96(3), 673–95. <https://doi.org/10.1177/1077699019859901>
- Carlson, M. (2015). The robotic reporter. *Digital Journalism*, 3(3), 416–431.
- Carlson, M. (2017). Automated judgment? Algorithmic judgment, news knowledge, and journalistic professionalism. *New Media & Society*, 8(4). <https://doi.org/10.1177/1461444817706684>
- Diakopoulos, N. (2019). *Automating the news: How algorithms are rewriting the media*. Harvard University Press.
- Gillespie, T. (2014). The relevance of algorithms. In: Gillespie, T., Boczkowski, P., & Foot, K. (Eds.) *Media technologies: Essays on Communication, materiality, and society*. MIT Press, 167–194. <https://doi.org/10.7551/mitpress/9780262525374.001.0001>
- Lewis, S. C., Guzman, A. L., & Schmidt, T. R. (2019). Automation, journalism, and human–machine communication: Rethinking roles and relationships of humans and machines in news. *Digital Journalism*, 7(4), 409–427. <https://doi.org/10.1080/21670811.2019.1577147>
- Marconi, F. (2020). *Newsmakers: Artificial intelligence and the future of journalism*. Columbia University Press.
- Thurman, N., Dörr, K., & Kunert, J. (2017). When reporters get hands-on with robo-writing. *Digital Journalism*, 5(10), 1240–1259.
- Wu, S., Tandoc, E. C. Jr., & Salmon, C. T. (2019). Journalism reconfigured. *Journalism Studies*, 20(10), 1440–1457. <https://doi.org/10.1080/1461670X.2018.1521299>



A checklist of eighteen pitfalls in AI journalism

Sayash Kapoor, Arvind Narayanan

September 30, 2022. Read the blog post which introduces this checklist: <https://aisnakeoil.substack.com/p/eighteen-pitfalls-to-beware-of-in>

⇒ Flawed human-AI comparison

What? A false comparison between AI tools and humans that implies AI tools and humans are similar in how they learn and perform.

Why is this an issue? Rather than describing AI as a broad set of tools, such comparisons anthropomorphize AI tools and imply that AI tools have the potential to act as agents in the real world.

Pitfall 1. Attributing agency to AI

Describing AI systems as taking actions independent of human supervision or implying that they may soon be able to do so.

“Artificial intelligence is starting to take over repetitive tasks in classrooms, like grading”

– The Machine Are Learning, and So Are the Students, The New York Times

Pitfall 2. Suggestive imagery

Images of humanoid robots are often used to illustrate articles about AI, even if the contents of the article have nothing to do with robots. This gives readers a false impression that AI tools are embodied, even when it is just software that learns patterns from data.

– How elite investors use artificial intelligence and machine learning to gain an edge, CNN

▶ This article has nothing to do with robots and is about AI tools to find patterns in financial data.

Pitfall 3. Comparison with human intelligence

In some cases, articles on AI imply that AI algorithms learn in the same way as humans do. For example, comparisons of deep learning algorithms with the way the human brain functions are common. Such comparisons can lend credence to claims that AI is “sentient”, as Dr. Timnit Gebru and Dr. Margaret Mitchell note in their recent op-ed.

“[The study] focused on an AI technique called deep learning, which employs algorithms, big data, and computing power to emulate human intelligence.”

– AI may be as effective as medical specialists at diagnosing disease, CNN

▶ “Emulating human intelligence” is not an accurate description of what deep learning does. It gives readers a false intuition that deep learning algorithms compare with human intelligence.

Pitfall 4. Comparison with human skills

Similarly, articles often compare how well AI tools perform with human skills on a given task. This falsely implies that AI tools and humans compete on an equal footing—hiding the fact that AI tools only work in a narrow range of settings.

“A new scientific review has concluded that artificial intelligence (AI) may be able to diagnose disease as successfully as human healthcare professionals”

– AI may be as effective as medical specialists at diagnosing disease, CNN

- ▶ This sentence hides the fact that AI tools only perform a narrow slice of the variety of steps that comprise a diagnosis.

⇒ **Hyperbolic, incorrect, or non-falsifiable claims about AI**

What? Claims about AI tools that are speculative, sensational, or incorrect can spread hype about AI.

Why is this an issue? Such claims give a false sense of progress in AI and make it difficult to identify where true advances are being made.

Pitfall 5. Hyperbole

Describing AI systems as revolutionary or groundbreaking without concrete evidence of their performance gives a false impression of how useful they will be in a given setting. This issue is amplified when the AI tool is deployed in a setting where they are known to have past failures—we should be skeptical about the effectiveness of AI tools in these settings.

“For years, people have tried to re-engineer learning with artificial intelligence, but it was not until the machine-learning revolution of the past seven years that real progress has been made.”
(emphasis ours)

– The Machine Are Learning, and So Are the Students, The New York Times

- ▶ This statement has no evidence in the article to back it up. In fact, EdTech has proven to be notoriously failure-prone in the last decade.

Pitfall 6. Uncritical comparison with historical transformations

Comparing AI tools with major historical transformations like the invention of electricity or the industrial revolution is a great marketing tactic. However, when news articles adopt these terms, they can convey a false sense of potential and progress—especially when these claims are not backed by real-world evidence.

“In Altman’s view, the unfolding AI revolution may well be more consequential for humanity than the preceding agricultural, industrial and computer revolutions combined.”

– Is AI finally closing in on human intelligence?, Financial Times

- ▶ The article uncritically quotes Sam Altman, co-founder of OpenAI, in comparing AI with historical transformations.

Pitfall 7. Unjustified claims about future progress

Claims about how future developments in AI tools will affect an industry, for instance, by implying that AI tools will inevitably be useful in the industry. When these claims are made without evidence, they are mere speculation on the part of the article, and like before, can give a false impression about these developments.

“Chatbots, for example, can be clumsy and frustrating today, but they will eventually become truly conversational, learning our habits and personalities and even develop personalities of their own.”

– A.I. Here, There, Everywhere, The New York Times

Pitfall 8. False claims about progress

In some cases, articles can include false claims about what an AI tool can do.

“The system’s auto-grader teaches itself how to score.”

– The Machine Are Learning, and So Are the Students, The New York Times

- ▶ The article falsely claims that the tools can learn how to grade homework on their own, when in fact the tool merely scores student answers against the correct answers fed into the system.

Pitfall 9. Incorrect claims about what a study reports

News articles often cite academic studies to substantiate their claims. Unfortunately, there is sometimes a gap between the claims made based on an academic study and what the study reports.

“Studies show that [ML] systems can raise student performance well beyond the level of conventional classes and even beyond the level achieved by students who receive instruction from human tutors.”

– The Machine Are Learning, and So Are the Students, The New York Times

- ▶ The study cited here does not refer to machine learning even once.

Pitfall 10. Deep-sounding terms for banal actions

As Prof. Emily Bender discusses in her work on dissecting AI hype, using phrases like “the elemental act of next-word prediction” or “the magic of AI” implies that an AI tool is doing something remarkable in the course of its operation. It hides an understanding of how mundane the tasks are, and that AI tools are functioning exactly as expected.

“I ask the gods of artificial intelligence to turn on the light”

– A.I. Here, There, Everywhere, The New York Times

⇒ Uncritically platforming those with self-interest

What? News articles often use PR statements and quotes from company spokespeople to substantiate their claims without providing adequate context or balance in their news stories.

Why is this an issue? Emphasizing the opinions of self-interested parties without providing alternative viewpoints can give an over-optimistic sense of progress.

Pitfall 11. Treating company spokespeople and researchers as neutral parties

When an article only or primarily has quotes from company spokespeople or researchers who built an AI tool, it is likely to be over-optimistic about the potential benefits of the tool.

AI tested as university exams undergo digital shift

– Financial Times

- ▶ Almost the entire article is written from the perspective of the company selling AI tools. As a result, the article reads more like a PR piece and less like a news story.

Pitfall 12. Repeating or re-using PR terms and statements

News articles often re-use terms from companies' PR statements instead of describing how an AI tool works. This can lead to misleading wording that misrepresents the actual capabilities of a tool.

"She uses the platform Bakpax that can read students' handwriting and auto-grade schoolwork"

– *The Machine Are Learning, and So Are the Students*, The New York Times

- ▶ The article repeatedly reuses PR terms such as "read students' handwriting" and "auto-grade" homework. Though Bakpax has since shut down, we found these PR terms on the company's archived website.

⇒ Limitations not addressed

What? The potential benefits of an AI tool are emphasized, but the potential limitations are not addressed or emphasized.

Why is this an issue? A one-sided analysis of AI tools can hide the potential limitations of these tools.

Pitfall 13. No discussion of potential limitations

Limitations such as inadequate validation, bias, and potential for dual-use plague most AI tools. When these limitations are not discussed, readers can get a skewed view of the risks associated with AI tools.

The Machine Are Learning, and So Are the Students

– The New York Times

- ▶ There is no discussion of potential limitations of the use of AI in EdTech. The article mentions in passing that there could be privacy concerns, but quickly provides a quote from the developer of the tool to assuage these concerns.

Pitfall 14. Limitations de-emphasized

Even if an article discusses these limitations and quotes experts who can explain them, limitations are often downplayed in the structure of the article, for instance by positioning them at the end of the article or giving them limited space.

"The big caveat is, in my opinion, that the story is not 'AI may be as good as health professionals', but that 'the general standard of evaluating performance of AI is shoddy'..."

– *AI may be as effective as medical specialists at diagnosing disease*, CNN

- ▶ In this story about a research study, despite an expert taking pains to point out what the real message of the study is, the article buries this information at the far end.

Pitfall 15. Limitations addressed in a "skeptics" framing

Limitations of AI tools can be caveated in the framing of the article by positioning experts who explain these limitations as skeptics who don't see the true potential of AI.

"Some skeptics argue that the software is capable only of blind mimicry ..."

– *A.I. Is Mastering Language. Should We Trust What It Says?*, The New York Times Magazine

- ▶ Instead of engaging with the substantive points of Prof. Emily Bender and others, this NYT Mag piece circumscribed their viewpoints to a skeptic's framing. Prof. Bender discusses issues with this article in much more detail in her response to this article.

Pitfall 16. Downplaying human labor

When discussing AI tools, articles often foreground the role of technical advances and downplay all the human labor that is necessary to build the system or keep it running. The book *Ghost Work* by Dr. Mary L. Gray and Dr. Siddharth Suri reveals how important this invisible labor is. Downplaying human labor misleads readers into thinking that AI tools work autonomously, instead of clarifying that they require significant overhead in terms of human labor, as Prof. Sarah T. Roberts discusses.

“A plethora of online courses and tutorials also have freed teachers from lecturing”

– *The Machine Are Learning, and So Are the Students*, *The New York Times*

- ▶ This phrase does not acknowledge the labor that goes into recording these lectures, maintaining online resources, and shifts the attention away from the human labor that goes into maintaining this system.

Pitfall 17. Performance numbers reported without uncertainty estimation or caveats

There is seldom enough space in a news article to explain how performance numbers like accuracy are calculated for a given application or what they represent. Including numbers like “90% accuracy” in the body of the article without specifying the conditions under which these numbers are calculated can misinform readers about the efficacy of an AI tool, especially because AI tools are known to suffer performance degradations even under slight changes to the datasets they are evaluated on.

“A new computer algorithm can now forecast crime in a big city near you – apparently. The algorithm, which was formulated by social scientists at the University of Chicago and touts 90% accuracy...”

– *Algorithm Claims to Predict Crime in US Cities Before It Happens*, *Bloomberg*

- ▶ The article has no details that might help the reader to understand what an accuracy of 90% means.

Pitfall 18. The fallacy of inscrutability

Referring to AI tools as inscrutable black boxes is a category error. Instead of holding the developers of these tools accountable for their design choices, it shifts scrutiny to the technical aspects of the system. Journalists should hold developers accountable for the performance of AI tools rather than referring to these tools as black boxes and allowing developers to evade accountability.

Our Machines Now Have Knowledge We'll Never Understand

– WIRED

- ▶ The article's main point is that it is impossible to understand how models “reason”, simply because they have a large number of parameters or weights. It ignores an entire body of research on model interpretability and explainability. The author further argues that because we cannot understand the model's internal representations, there is no way to use them in a way that meets legal requirements for non-discrimination and explanation, such as in credit scoring. But those requirements are about the way that decision making algorithms interact with the world rather than their internals, and algorithms can always be understood at this level, as Dr. Kroll notes.



Afterword: Educating AI-savvy Journalists

Maarit Jaakkola

This handbook discussed artificial intelligence (AI), or how machines designed to be intelligent agents equipped with belief systems can be reported in journalism and addressed in journalism education to prepare future journalists. The overall objective is to provide journalism educators with basic knowledge about AI as a democratic concern so they can **teach journalism students to report on increasingly technology-saturated societies.**

The key lessons in this book, which many modules return to, deal with the dual character of AI technologies. AI applications have potential for both good and bad; thus, they can have either a positive or negative impact on society. However, at the end of the day, AI is always what humans make of it. At the same time, humans are the ones who decide how to employ the opportunities that AI technologies offer, although under conditions of differential power and control.

! We as humans have, however, to ensure – to use Microsoft’s terminology – the fairness, accountability, transparency and ethics (FATE) of AI use, which is a major societal mission.

Therefore, journalists are needed to address public apprehension and promote debate on this topic.

It is not the journalist’s task to make decisions about the design and use of AI, but **journalists play a pivotal role in mediating timely knowledge** to policymakers, developers and other practitioners, researchers and other producers of knowledge, educators and other mediators of knowledge, as well as the market and civil society actors. A prerequisite for this is that journalists must build an understanding of what AI is about. A basic understanding of AI and related debates requires people to feel apprehensive about some of the fundamental ideas related to this technology, such as the fact that AI agents are representations or models of the world and they are used for reasoning. Computers can perceive the world using sensors; they can learn from data. They can tackle the challenge of making AI agents interact with humans (Long & Magerko, 2020.) A critical understanding of AI requires tracing how system structures are based on decisions that veil economic, ethical or value-based political interests, which have consequences that maintain structures of power.

Technologies have been used in media and in society for a long time. Newsrooms can learn from the societal experiences. The contemporary introduction of AI into newsroom practices may not be considered revolutionary, as was the introduction of computers to newsrooms. The effects of the change are slower and have become gradually engraved in the existing structures. But, just as questions are posed as regards the wider society, we should also ask what does this mean for journalism education? What AI may demand is the partial *integration* of the topic into existing structures of education and journalistic routines, but some aspects of the existing

structures and values need to be *reinforced*. With the **proliferation of automated systems programmed to accomplish tasks and carry out functions without supervision, entirely new questions and practices are likely to emerge**. Instead of simply proclaiming the need for AI journalists and AI journalism programmes, journalism education can take a step back and reflect on how to **reconfigure educational thinking and practices**.

To conclude, I would like to summarise the conceptual, discursive, competence, public educational and didactic challenges that AI poses for journalism education. Ultimately, journalism educators need to ask:

- how to relate to what may be a post-humanistic shift where non-human actors are increasingly part of all interactions;
- how to be conscious of the discursive role of journalism in shaping future understandings of AI;
- how to encounter AI as a set of knowledge competences that need to be mastered;
- how to empower journalists contribute to the mission of educating citizens and consumers, which may be necessary to embrace the changes; and
- and how to advance their own activities by asking how AI could serve the production of learning experiences in newsrooms – as well as society at large.

⇒ Conceptual Challenge

In discourse on AI, the focus is often on how AI will change society, but it is also worth reflecting upon how AI may change humans. AI deployments increasingly affect our behaviour and influence our conception of reality.

While we can teach AI to work in a specific way, AI is also making us adjust our behaviours and further learn from these changes. AI-powered technologies are typically embodied, immersive and connected to our own capabilities of observation and decisions. The mechanism of selection, persuasion, networking and socialising affect the decision-making and behaviour of individuals.

Traditionally, primacy has been afforded to humans, but the proliferation of **intelligent agents** give rise to something that has been called **post- or transhumanism, post-dualism or a posthuman condition**. According to such philosophical approaches, human beings will no longer be considered to be the centre of everything. There is, therefore, an endeavour to recognize the value of the various non-human ways of acting in the world. The possible post-human condition encourages us to see society not just as an outcome of human behaviour, but as a result of the **interactions between humans** (and their power differentials) and **non-human or ahuman actors**. For journalists, the challenge is how to create storytelling without exoticizing or otherizing intelligent agents, or anthropomorphizing and over-humanizing material artefacts and abstract processes, without remaining too distant and clinical.

In journalism, the incorporation of non-human agents contributing to news work may pose an existential crisis. As has been discussed by many scholars, the basic understanding of journalism is that it is a human-made craft in which human contribution to society is crucial.

This shift generates the need for a new conceptual change, creating the need to sustain a new kind of relationship to the environment, where humans are seen more as collaborators at the same level with the non-living world. This requires more intellectual curiousness to discover how the non-human counterparts work

(albeit created by humans and particular configurations of power), but also an increased consciousness about the **ethical responsibility of humans and humankind**. Here, as the images of science fiction of machines and robots taking over the world is persistent, it is necessary to attempt to overcome the cultural myths and see technologies more realistically, positioning human, animal and machine intelligence more proportionately.

⇒ Discursive Challenge

Discourses around AI are characterized by a wide diversity of voices, mainly from the Global North, and a large breadth of topics that are of varying relevance to different regions of the world. AI technologies and their ramifications are discussed in a variety of scientific fields, including computer science, information science, sociology, anthropology, science and technology studies and political science, to name a few. The scope of scientific studies remains rather limited, drawing on the vocabulary, methodology and previous discussions of a specific disciplinary tradition, demanding pre-knowledge of concepts and contexts from the audience. Policy reports, whitepapers and strategy documents are produced by authorities, industry or civil-society initiatives and non-profit organizations with an often optimistic and hopeful perspective embracing innovation as a solution to problems. Industry perspective tends to highlight broad democratic issues and the common good, but tech giants are invested with their own interests. The everyday perspectives of users, as citizens and consumers, vary greatly.

Journalism is positioned in the middle of these discursive landscapes, negotiating with terminologies and traditions. Appropriately, **AI has been described as a “boundary object”**, a bridging concept used as an umbrella term in common discourse to bring together different actors in the public sphere to deliberate on a common issue in a language that may not be native to the fields they come from (Malinverni, 2015; Moran & Shaikh, 2022, Reddy et al., 2019). The differences between the discourses are not merely semantic; studies have shown that, for example, journalism and industry are deploying narratives and expressing concerns that can be contradictory (Moran & Shaikh, 2022). Part of the public understanding of AI is based on the public discourse on AI that is produced by journalism.

As media has the power to significantly frame the public debate and shape the discussion, journalism needs to ask key questions, such as how AI can work for society and how it should work, and how it is applied not only to the immediate sphere of journalism but to other sectors of society, in very different ways.

As discussed in this handbook, the public discourse in news media on AI has certain biases, such as the dominance of industry sources, major transnational organizations and their new products or initiatives, as well as insider and expert views. The **gender, racial and regional biases of AI systems**, such as when certain algorithms suppress voices with specific characteristics and backgrounds, may also easily be transferred into the journalistic discourse and coverage if journalists are not conscious of them, and the discrimination of individuals with a certain colour or gender may become normalized.

We are living in an era where new research is continually being published, new concepts are coined to label emerging and existing phenomena, and new methods are introduced to capture the multifaceted area. Journalism educators, journalism students and professionals need to follow news and research to be (re-) updated about the newest state of the art in machine learning, computer vision, speech recognition, natural language processing, expert systems and robotics, and their ramifications. Since it is likely that a journalist follows AI as a field in connection to another area, such as environmental sustainability, school education, industry,

disinformation and media literacy, or health care, curated services that collect information at the crossroads of different sectors are important.

⇒ Competence Challenge

Ultimately, journalism educators should teach students less about of technology and focus more on what we typically call soft skills. In Poynter’s industry report from almost a decade ago (Finberg & Klinger, 2014), the core competences that were foreseen as essential for future journalists were related to **personal characteristics** (curiosity, accuracy, ability to handle stress and deadlines), **personal competences** that can be gained (broad general knowledge, news judgment, awareness of current events, criticality), **abilities related to work and organisational behaviour** (social skills, team player skills), knowledge (of the media landscape, society, technologies, law and ethics, history) and **work practices and methods** (information search and acquisition, interview techniques, analysis of information and data, sourcing, techniques of presentation and storytelling, networking, audience development and engagement). These general skills apply to all aspects of journalism education, including the apprehension of AI as a disruptive technology framework.

The skills that could be added to this are the ability to embrace change and innovation with creativity and the capacity to assess them critically and with a historical perspective. Other essential skills include numeracy skills, such as the ability to understand data, statistics and scale, and the ability to translate ideas from one world to another. That mainly entails translating ideas from areas with highly specialized structures and terminology, such as technology (often times, technology in a specific subfield, such as natural language processing in the educational sector), to everyday contexts.

Both epistemologically and methodologically, journalists reporting on AI can draw on many established journalistic approaches. It is worth emphasizing that AI-themed questions can be incorporated into any journalism genre, ranging from news to cultural journalism, so it is not always journalism that needs to adjust and change. However, technologically-oriented approaches – whether referred to as computational journalism, data journalism, automated journalism, algorithmic journalism, robot journalism or augmented journalism – have deepened journalism’s relationship to data and computer systems. **Technology-anchored ways of reporting**, such as immersive journalism and mobile journalism, as well as the collaborative work of data journalists, journalist-programmers and hacks and hackers, have bearings on “AI journalism”. Approaches with a problem-solving and developing mindset, such as constructive journalism, solutions-based journalism, positive journalism, investigative journalism and fact-checking, have helped journalists envision alternative narratives to the science fiction fantasies, **analyze and validate information and “co-think” with audiences**. Furthermore, **participatory, reciprocal and dialogical approaches in journalism invite citizens and consumers to investigate the everyday dimensions of the especially narrow AI**.

⇒ Public Educational Challenge

AI also poses a challenge for society in general. The public understanding of AI-related issues is often limited. Currently, various efforts have been made across the globe to introduce AI into school curriculums and promote education for AI literacy in different countries, related to a country’s efforts to create national AI strategies. For example, in the UK, the Government’s Office for AI published an AI roadmap in 2021 with a vision “for

everyone to be able to live confidently with AI, and for those who go on to work with it and to build it do so with the very best foundation" (2021, p. 16). Tech giants and public organisations are introducing initiatives to raise citizens' awareness of AI. However, the perspective that remains in the background is the fact that, as public meaning-makers, journalists also educate their audiences. Thus, regarding the challenges AI imposes on societies, journalists should be viewed as educators in AI literacy (see Broussard et al., 2019; Deuze & Beckett, 2022; Jaakkola, 2022).

AI literacy refers to the competencies necessary in a future in which AI transforms the way we live, communicate and work with machines (Long & Magerko, 2020). An individual can be considered **AI-literate** when he or she is competent in using AI in an ethically responsible manner (Ng et al., 2021). The conditions for understanding AI systems or human-machine interaction look very different in different populations, as people may have a broad range of first-hand experiences with technology.

As always, there is no average reader or audience member. Individuals have varying connections to the digital environment. There are significant differences in people's skills in dominant languages, literacy, attitudes, data processing and numeracy and problem-solving skills – everything that can be subsumed under the concept of media or/and information literacy, or, perhaps in a more limited sense, data or AI literacy.

Accordingly, journalistic messages, educational interventions, and the journalistic interventions with an educational approach need to be designed differently for school children, university students, professional practitioners in different fields, senior citizens and minorities. Within these groups, the essence of AI is a common specific challenge for understanding the workings of AI. However, that concept is abstract, ubiquitous and often opaque to AI users. To increase understandings among people who have no first-hand experience in this area, it would be most effective to put individuals in an intelligence agent's shoes and ask them to make sense of its reasoning processes, its built-in artificial intelligence. Journalism can do this by diversifying the practices of presentation that do not only have to be based on the mediation of information. It can also engage people through immersion, gamification and involvement.

⇒ Didactic Challenge

Last, but not least, **journalism educators need to determine how best to harness AI for the sake of teaching and learning journalism.** The examples that have been seen, so far, ranging from simulations of human and human – computer encounters of immersive journalism to writing exercising tools, are only the beginning. AI applications can help train aspiring journalists to hone their skills in the areas of competence mentioned above. Creating **didactic metaverses** to simulate situations and worlds relevant to journalistic abilities may facilitate engaging and **embodied learning experiences.**

To produce new didactic solutions that would be feasible and scalable, that is, ready to be disseminated across institutions and countries, journalism educators need to partner with AI developers. Thus, learning resources are no longer produced as the solo work of an individual educator. The need to update individuals' competences is a collaborative effort that requires the participation of journalism educators, students and the journalism industry.

Glossary of Terms

Algorithm is a sequence of rules used to accomplish a task in a computerized device.

Algorithmic bias describes systematic and repeatable errors in a computer system that create unfair outcomes, such as privileging one arbitrary group of users over others. Bias can emerge from many factors, including the design of the algorithm or the unintended or unanticipated use or decisions relating to the way data is coded, collected, selected or used to train the algorithm. Algorithmic bias has been observed, for example, in search engine results and social media platforms.

Algorithmic journalism, see automated journalism

Android is a humanoid robot or synthetic organism designed to imitate a human.

Antropomorphism refers to the interpretation or perception of a non-human object in terms of human characteristics. It thus means humanization of an object that is not human, for example, naming software with human names.

Asimov's laws, also known as the three laws of robotics, are a set of principles devised by science fiction author Isaac Asimov in the 1940s. According to the First Law, "a robot may not injure a human being or, through inaction, allow a human being to come to harm." According to the Second Law, "a robot must obey the orders given it by human beings except where such orders would conflict with the First Law." According to the Third Law, "a robot must protect its own existence as long as such protection does not conflict with the First or Second Law." The original laws have been altered and elaborated by Asimov and other authors. The laws have widely pervaded science fiction and popular culture, as well as affected thoughts on the ethics of AI.

Automated journalism, also known as algorithmic or robot journalism, means the use of software to generate journalistic content without a human intervention.

Bad bot, see bot

Big data refers to data sets that are too large or complex to be dealt with by traditional data-processing application software.

Black box describes in science, computing and engineering a system that can be viewed in terms of its inputs and outputs without any knowledge of its internal workings, which are opaque ("black"). We can talk, for example, about black-box AI or algorithms, referring to the fact that we do not know how they are constructed.

Bot, also called an internet robot, is a software application that runs automated tasks on the internet, for example a chatbot that converses with a user. Some bots are non-harmful while there are also malicious bots

(bad bots) conducting attacks and frauds. The earliest chatbots were ELIZA (1966) and PARRY (1972). Bots are also referred to as “conversational AI”.

Clarke’s laws are a set of three principles formulated by the science-fiction author Arthur C. Clarke in the 1960s and 1970s describing the perception of technology. The First Law suggests: “When a distinguished but elderly scientist states that something is possible, he is almost certainly right. When he states that something is impossible, he is very probably wrong.” According to the Second Law, “the only way of discovering the limits of the possible is to venture a little way past them into the impossible”. The Third Law, which is the most frequently cited, reads as follows: “Any sufficiently advanced technology is indistinguishable from magic.” Isaac Asimov’s corollary to Clarke’s First Law is: “When, however, the lay public rallies round an idea that is denounced by distinguished but elderly scientists and supports that idea with great fervour and emotion – the distinguished but elderly scientists are then, after all, probably right.”

Computational journalism means the application of computation to journalistic practice. See also automated journalism.

Computer vision is an interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos. It seeks to understand and automate tasks that the human visual system can do, developing methods for acquiring, processing, analyzing and understanding digital images. Computer vision may deal with scene reconstruction, object or event detection, video tracking, object recognition, 3D pose estimation, motion estimation, 3D scene modeling, or image restoration.

Cyborg is a being (from *cybernetic* and *organism*) that has both organic and artificial body parts. The term was coined by Manfred Clynes and Nathan N. Kline in 1960.

Cyberbole is a neologism (from *cyber* and *hyperbole*) denoting exaggerated claims about what internet technologies can or will accomplish.

Data mining is an interdisciplinary field referring to the process of extracting and discovering patterns in large amounts of data. The extraction (“mining”) of patterns may be based on the identification of groups of data records (cluster analysis), deviating records (anomaly detection) or dependencies between records (association rule mining).

Datafication refers to the process of turning people’s lives into data and transferred into new form of value. The term was introduced by Kenneth Cukier and Victor Mayer-Schöenberger in 2013.

Deepfake means manipulated content (from *deep* learning and *fake*) where an existing image or video is altered by replacing some of its constituents with something else. Examples include modifying facial expressions of a person in video or cloning human voices.

Deep learning is a class of machine learning algorithms using multiple layers to progressively to extract features from the raw input. For example, in image processing, lower layers may identify edges, while higher layers may identify the concepts relevant to a human such as digits or letters or faces. The term was introduced by Rina Dechter in 1986.

Filter bubble, also known as an ideological frame, describes a situation of intellectual isolation online which is a result from separation from information that contrasts one’s beliefs and viewpoints. A filter bubble may occur based on a personalization algorithm that selectively guesses what information the user would like to see.

Friendly AI (FAI) refers to benign uses of AI, pursuing solutions to how to make AI technologies work for positive and good human purposes by using safe and useful agents. The term was coined and popularized by Eliezer Yudkowsky. A kinship concept is “good AI”, also used in the context of the concept of “good AI society”.

Humanoid is a non-human object with human form or characteristics.

Immersive journalism is a form of journalism that allows audiences to enter virtually recreated worlds by using immersive virtual realities such as computer graphic imagery (CGI) or 3D technologies.

Intelligent agent (IA) is anything that perceives and acts autonomously in an environment and is able to use accumulated knowledge to adjust or improve its behaviour. According to a common definition, if an agent acts to maximise the success of its performance based on past experience and knowledge it can be regarded as intelligent.

Internet of Things (IoT) describes physical objects that are embedded with technologies connecting and exchanging data with other device over the internet. For example, many household device such as remotely controllable washing machines, water heaters, refrigerators, lighting systems or cars are examples of IoT. The term emerged approximately in 2008.

Machine learning (ML) refers to the study of computer algorithms that can improve automatically through experience and by the use of data, the so called training data.

Natural language processing (NLP) is a subfield of linguistics and computer science dealing with AI, concerned with the interactions between computers and human language, in particular how to program computers to process and analyze large amounts of natural-language data. Natural language a language that has evolved naturally in humans through the use and repetition without any conscious planning, different from artificial or constructed languages such as computer programming languages.

Neural networks, also known as artificial neural networks (ANNs), are computing systems inspired by the biological neural networks that constitute animal brains. A neural network is a collection of connected nodes that imitate the structure of neurons in a biological brain.

Platformization refers to the increased role in communication by platforms, frameworks on which different applications can be run, typically owned by the major companies Google, Apple, Facebook (Meta), Amazon, Microsoft called GAFAM (GAMAM) or the Big Five.

Policy means a set of rules or guidelines that (especially in the meaning of “public policy”) is used as a basis for making decisions.

Robot is a task-performing machine that is capable of carrying out actions automatically. The term was first used by Karel Čapek to denote a fictional humanoid in a Czech-language play in 1920.

Robot journalism, see automated journalism

Singularity, also known as technological singularity, means a point of time where the intelligence of computing has grown to the extent that it surpasses the human intelligence, resulting in superintelligence, hyperintelligence, or superhuman intelligence. Superintelligence thus means a hypothetical agent that possesses intelligence that surpasses the human mind. The concept of technological singularity was developed by the futurist Raymond Kurzweil in his book in 2005.

Strong AI, also known as general AI, full AI or artificial general intelligence (AGI), describes machines with the ability to apply intelligence to any problem rather than just one specific problem.

Superintelligence, see singularity

Symbolic AI, also known as GOFAI, denotes a historical phase of AI that simulated the high-level conscious reasoning that people use when they solve puzzles, express legal reasoning and do mathematics. The term GOFAI stands for “Good Old-Fashioned Artificial Intelligence”, a word coined by John Haugeland in 1985.

Turing test is a test of a machine’s ability to exhibit intelligent behavior equivalent to, or indistinguishable from, that of a human. The test measures the “humanness” of an object and is based on an assumption that if humans take the actions of a non-human agent as human it can be regarded as intelligent. The idea was invented by Alan Turing in the 1950s, who called it an imitation game.

Uncanny valley describes the relationship between the human-like appearance of a robotic object and the human emotional response. It asserts that robots that are similar to human beings are positively perceived – until a certain point when they are perceived as becoming too real. At that point, acceptance drops significantly, and people feel a sense of unease in their response. “Valley” thus refers to a dip in the human observer’s affinity for the replica. The revulsion towards the humanoid robots increases along with the robots’ human likeness. The term and mechanism were first introduced in the 1970s by Masahiro Mori, a professor at the Tokyo Institute of Technology.

Weak AI, also known as narrow AI or applied AI, describes machines with the ability to apply intelligence to one narrow task, instead of any problem. This is the most commonly used form of AI in the public discussions.

Contributors

Experts

Arvind Narayanan is a professor of computer science at Princeton. He co-authored a textbook on fairness and machine learning and is currently co-authoring a book on AI snake oil. He led the Princeton Web Transparency and Accountability Project to uncover how companies collect and use our personal information. His work was among the first to show how machine learning reflects cultural stereotypes, and his doctoral research showed the fundamental limits of de-identification. Narayanan is a recipient of the Presidential Early Career Award for Scientists and Engineers (PECASE), twice a recipient of the Privacy Enhancing Technologies Award, and thrice a recipient of the Privacy Papers for Policy Makers Award.

Charlie Beckett is Professor of Practice in the Department of Media and Communications at the London School of Economics (LSE) in the UK. He is Director of Polis, an international journalism think-tank at the university. He was Lead Commissioner for the LSE Truth, Trust & Technology Commission into the information crisis in 2018. Currently, he is leading the Polis/LSE JournalismAI project, an international initiative to work with news organizations around the world to research, inform and innovate with artificial intelligence. He was the author of the 2019 report *New Powers and New Responsibilities*, on how news organizations around the world are using artificial intelligence technologies. Prior to joining LSE in 2006, Professor Beckett was a senior producer and editor at BBC and ITN's Channel 4 News.

Edward Finn is the Founding Director of the Center for Science and the Imagination at Arizona State University in the United States, where he is an associate professor in the School for the Future of Innovation in Society and in the School of Arts, Media and Engineering. His research and teaching explore the workings of imagination, computational culture and futures. He is the author of *What Algorithms Want: Imagination in the Age of Computing* (MIT Press, 2017) and co-editor of *Future Tense Fiction* (Unnamed Press, 2019), *Frankenstein: Annotated for Scientists, Engineers and Creators of All Kinds* (MIT Press, 2017) and *Hieroglyph: Stories and Visions for a Better Future* (William Morrow, 2014), among other books. Professor Finn completed his Ph.D. in English at Stanford University in 2011 and his bachelor's degree at Princeton University in 2002. Before graduate school, he worked as a journalist at *Time*, *Slate* and *Popular Science*.

Fredrik Heintz is Professor of Computer Science at Linköping University, Sweden, where he directs the Reasoning and Learning Lab. His research focuses on artificial intelligence, especially Trustworthy AI and the intersection between machine reasoning and machine learning. Fellow of the Royal Swedish Academy of Engineering Sciences, Professor Heintz is also Director of WASP-ED: The Wallenberg AI and Transformative Technologies Education Development Program, Director of the Graduate School of WASP: The Wallenberg AI, Autonomous Systems and Software Program, Coordinator of the TAILOR ICT-48 network developing the scientific foundations of Trustworthy AI, and President of the Swedish AI Society.

Frederic Heymans works as a research and project officer at the Knowledge Centre Data & Society (Kenniscentrum Data & Maatschappij), a centre for studying legal, ethical and societal aspects of artificial intelligence and data-driven applications at Vrije Universiteit Brussel, Belgium. Heymans works in the university's imec (a R&D hub for nano and digital technologies) research group SMIT (Studies in Media, Innovation and Technology) on various projects that

focus on the impact of AI systems on society, conducting policy analysis, drafting reports and publications, and creating practical tools tailored to the Knowledge Centre's target groups.

Santosh Kumar Biswal is an Associate Professor in the Department of Journalism and Mass Communication, Rama Devi Women's University in Bhubaneswar, India. Before joining Rama Devi Women's University, Dr. Biswal was an assistant professor at the Symbiosis Institute of Media & Communication, Symbiosis International (Deemed University) in Pune, India. Previously, he worked at Andhra Loyola College and for *Hindustan Times* (New Delhi). He was a university topper and gold medallist in Social Communication at Utkal University, Bhubaneswar, India. Holding an M.A., M.Phil., PGDJMC, UGC-NET and a Ph.D., his research interests include journalism, development communication, disability, digital media and media technology. In addition to many journal articles, he is the author (as co-editor) of the book *Holistic Approaches to Brand Culture and Communication across Industries* and co-editor of *Handbook of Research on Social and Cultural Dynamics in Indian Cinema*. He regularly writes opinion columns for English-language national dailies such as *The Indian Express*, *The Asian Age*, *The Telegraph*, *Hindustan Times*, *The Pioneer*, *The Statesman* and *Deccan Chronicle*.

Agnes Stenbom is Responsible Data & AI Specialist at Schibsted, the largest media group in the Nordics, leading various strategic efforts related to artificial intelligence. On the side, she pursues a doctoral degree in Industrial Economics and Management at KTH Royal Institute of Technology in Stockholm, Sweden, where her research centres on how legacy media organizations may leverage artificial intelligence to achieve journalistic goals. Ms Stenbom is a co-founder of the Nordic AI Journalism network that brings together industry peers from more than 40 media organizations to discuss artificial intelligence. She is also a columnist for *Dagens Media*, a futures analyst for the Sweden-based think tank Global Challenge and a regular conference speaker. In 2021, she received the International News Media Association INMA's *30 Under 30* award as one of the world's top 30 young talents in the media industry.

Sayash Kapoor is a computer science Ph.D. candidate at Princeton. His research focuses on the impact of AI in science, journalism, and the industry. He is currently co-authoring a book on AI Snake Oil. The book looks critically at what AI can and cannot do. He has worked on understanding the social impact of AI in the industry and academia at Facebook, Columbia University, and EPFL Switzerland. Kapoor is a recipient of a best paper award at ACM FAccT and an impact recognition award at ACM CSCW.

Suren Jayasuriya is an Assistant Professor at Arizona State University's School of Arts, Media and Engineering as well as its School of Electrical, Computer and Energy Engineering. He received his Ph.D. in Electrical and Computer Engineering from Cornell University in 2017 and a bachelor's in mathematics and in philosophy from the University of Pittsburgh in 2012. His research interests are in computer vision and computational imaging, specifically designing new computational cameras and projector systems with machine learning and artificial intelligence to better understand the visual world around us. Professor Jayasuriya also conducts research on artificial intelligence and STEAM (Science, Technology, Engineering, the Arts and Mathematics) education targeted for various students from middle/high school through undergraduate/graduate studies as well as outreach and public engagement.

Jenny Wiik is Associate Professor of Media and Communication Studies at the University of Gothenburg and Malmö University in Sweden. She is currently affiliated with the innovation programme Media & Democracy at Lindholmen Science Park in Gothenburg, Sweden, where she leads the research project "The Automation of Journalism: Innovation, Collaboration and Knowledge Management in the Implementation of AI in News Organizations". Jenny Wiik's research interests centre on journalism as a democratic institution, with a special focus on professional development and editorial organization. She has held various leadership positions within the academy, resulting, among other things, in the setting up of the University of Gothenburg's international Master's Programme on Investigative Journalism.

Editor

Maarit Jaakkola is Co-director of Nordicom, Centre for Nordic Media Research at the University of Gothenburg, Sweden. She is also Associate Professor at the Department of Journalism, Media and Communication at the University of Gothenburg and Adjunct Professor at the Faculty of Information Technology and Communication Sciences at Tampere University in Finland. Jaakkola's research interests are focused on journalism education and its signature pedagogies, cultural digital practices, cultural journalism and criticism, as well as on media literacy policies and pedagogies. Her most recent book is *Reviewing Culture Online: Post-Institutional Cultural Critique across Platforms* (Palgrave Macmillan, 2022). Jaakkola is, moreover, Editor-in-Chief of NordMedia Network, a digital platform dedicated to Nordic media research, and editorial board member of numerous academic journals. She leads several scholarly networks, among them the Collaboration Committee of the Nordic Journalism Education, which is a member association of the World Journalism Education Council, the UNESCO University Network on Media and Information Literacy and Intercultural Dialogue (MILID) focused on media literacy and related research in higher education, the Academic Forum for Media and Information Literacy Research in Sweden, and the NordMedia Network's research division on Media Literacy and Media Education.

References

- Acerbi, A. (2020). *Cultural evolution in the digital age*. Oxford University Press.
- Adgate, B. (2021, August 20). Newspapers have been struggling and then came the pandemic. *Forbes*. <https://www.forbes.com/sites/bradadgate/2021/08/20/newspapers-have-been-struggling-and-then-came-the-pandemic/?sh=b6d9a1212e64>
- Aliman, N.-M. & Kester, L. (2020). Malicious design in AIVR, falsehood and cybersecurity-oriented immersive defenses. Paper presented at the 2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR) in 2020, 130–137.
- Andi, S., Selva, M., & Nielsen, R. K. (2020). *Women and leadership in the news media 2020: Evidence from ten markets*. Reuters Institute for the Study of Journalism.
- Bawack, R. E., Fosso Wamba, S., & Carillo, K. (2019). Artificial intelligence in practice: Implications for IS research. Paper presented at the 25th American conference on information systems. Cancun, Mexico.
- BBC (2018). India WhatsApp 'child kidnap' rumours claim two more victims. *BBC*, June 11, 2018. <https://www.bbc.com/news/world-asia-india-44435127>
- Bebawi, S. (2019). Data journalism and investigative reporting in the Arab world: From emotive to evidence-based journalism. In: Mutsvairo, B., Bebawi, S., & Borges-Rey, E. (Eds.), *Data journalism in the Global South. Palgrave Studies in Journalism and the Global South*. Palgrave Macmillan, 193–204.
- Beckett, C. (2019). *New powers, new responsibilities: A global survey of journalism and artificial intelligence*. Report from POLIS journalism and Society, think tank at the London School of Economics (LSE).
- Biswal S. K., & Gouda N. K. (2020). Artificial intelligence in journalism: A boon or bane? In: Kulkarni A. & Satapathy S. (Eds.). *Optimization in machine learning and applications: Algorithms for intelligent systems*. Springer, 155–167.
- Bloomfield, B. P. (1987). *The question of artificial intelligence: Philosophical and sociological perspectives*. Routledge.
- Bloomfield, M. (1995). *The automated society: What the future will be and how we will get it that way*. Masefield Books.
- Bounegru, L., & Gray, J. (Eds.) (2021). *The data journalism handbook: Towards a critical data practice*. Amsterdam University Press.
- Brantner, C., & Saurwein, F. (2021). Covering technology risks and responsibility: Automation, artificial intelligence, robotics, and algorithms in the media. *International Journal of Communication*, 15(2021), 5074–5098.
- Brennen, J. S., Howard, P. N., & Nielsen, R. K. (2018). *An industry-led debate: How UK media cover artificial intelligence*. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-12/Brennen_UK_Media_Coverage_of_AI_FINAL.pdf
- Broersma, M., & Graham, T. (2015). Tipping the balance of power: Social media and the transformation of political journalism. In: *The Routledge companion to social media and politics*. Routledge, 89–103.
- Broussard, M., Diakopoulos, N., Guzman, A. L., Abebe, R., Dupagne, M., & Chian, C.-H. (2019). *Artificial intelligence and journalism. Journalism & Mass Communication Quarterly*, 96(3), 673–695.
- Bradshaw, P. (2014). Data journalism. In: Zion, K., & Craig, D. (Eds.) *Ethics for digital journalists: Emerging best practices*. Routledge, 202–219.
- Brundage, M., Shahar, A., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, G.C., Steinhardt, J., Flynn, C., hÉigeartaigh, S.Ó., Beard, S., Belfield, H., Farquhar, S., Lyle, C., Crootof, R., Evans, O., Page, M., Bryson, J., Yampolskiy, R., & Amodei, D. (2018). *The malicious use of artificial intelligence: Forecasting, prevention, and mitigation*. Future of Humanity Institute.
- Bruns, A. (2019). After the 'APocalypse': Social media platforms and their fight against critical scholarly research. *Information, Communication & Society*, 22(11), 1544–1566. <https://doi.org/10.1080/1369118X.2019.1637447>
- Brüggemann, M. (2017). Shifting roles of science journalists covering climate change. *Oxford Encyclopedia of Climate Science*, (2017, April 26).
- Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W.W. Norton & Company.
- Bunz, M., & Braghieri, M. (2021). The AI doctor will see you know: Assessing the framing of AI in news coverage. *AI & Society*, (2021, March 8).

References

- Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User Modeling and User-adapted Interaction*, 12(4), 331–370. <https://doi.org/10.1023/A:1021240730564>
- Carlson, M. (2015). The robotic reporter: Automated journalism and the redefinition of labor, compositional forms, and journalistic authority. *Digital Journalism*, 3(3), 416–431. <https://doi.org/10.1080/21670811.2014.976412>
- Carlson, M. (2018). Automating judgment? Algorithmic judgment, news knowledge, and journalistic professionalism. *New Media & Society*, 20(5), 1755–1772. <https://doi.org/10.1177/1461444817706684>
- Cassie, K. (2018). *Journey to AI*. https://www2.slideshare.net/ITARENA/cassie-kozyrkov-journey-to-ai?from_action=save
- Cater, L., & Heikkilä, M. (2021, May 28). Your boss is watching: How AI-powered surveillance rules the workplace. *Politico*. <https://www.politico.eu/article/ai-workplace-surveillance-facial-recognition-software-gdpr-privacy/>
- Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial intelligence and the “good society”: The US, EU, and UK approach. *Science and Engineering Ethics*, 24(2), 505–528. <https://doi.org/10.1007/s11948-017-9901-7>
- Cathcart, T. (2013). *The trolley problem or would you through the fat guy off the bridge*. Workman Publishing.
- CFI (n.d.). *Data journalism*. Available at <https://cfi.fr/en/search?search=data+journalism&type%5Bproject%5D=project>
- Chuan, C. H., Tsai, W.-H. S., & Cho, S. Y. (2019). Framing artificial intelligence in American newspapers. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. ACM Press, 330–244. <https://doi.org/10.1145/3306618.3314285>
- Chui, M., Harryson, M., Manyika, J., Roberts, R., Chung, R., van Heteren, A., & Nel, P. (2018). *Notes from the AI frontier: Applying AI for social good*. McKinsey Global Institute.
- Clerwall, C. (2014). Enter the robot journalist: Users’ perceptions of automated content. *Journalism Practice*, 8(5), 519–531. <https://doi.org/10.1080/17512786.2014.883116>
- Code for Africa (n.d.). *Code for Africa (CfA)*. <https://github.com/CodeForAfrica/>
- Coelho, F., Rodríguez Altube, F., Louzau, N., Arambillet, D., Bouret, G., Miño, G., Brom, R., Morón Escalante, L., Trigo Viera, M., Villazuela, C., Zenteno Frers, S., & Zuzel, M. (2021). *Computer vision to monitor elections transparency*. Sigma Awards. <https://sigmaawards.org/computer-vision-to-monitor-elections-transparency/>
- Collett, C., Neff, G., & Gouvea, L. (2022). *The effects of AI on the working lives of women*. Inter-American Development Bank.
- Dastin, J. (2018, October 11). Amazon scraps secret AI recruiting tool that showed bias against women. U.S. Available at <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>
- Deuze, M. (2006). Global journalism education: A conceptual approach. *Journalism Studies*, 7(1), 19–34.
- Deuze, M., & Beckett, C. (2022). Imagination, algorithms and news: Developing AI literacy for journalism. *Digital Journalism*, 10(10), 1913–1918. <https://doi.org/10.1080/21670811.2022.2119152>
- Diakopoulos, N. (2014). Algorithmic accountability. *Digital Journalism*, 3(3), 398–415.
- Diakopoulos, N., & Koliska, M. (2017). Algorithmic transparency in the news media. *Digital Journalism*, 5(7), 809–828. <https://doi.org/10.1080/21670811.2016.1208053>
- Diakopoulos, N. (2019). *Automating the news: How algorithms are rewriting the media*. Harvard University Press.
- Diakopoulos, N. (2021). The algorithms beat: Angles and methods for investigation. *The Data Journalism Handbook*, 219–229. Amsterdam University Press. https://doi.org/10.5117/9789462989511_ch31
- Doherty, S. (2016). NewsCubed: Journalism through design. *Journalism Practice*, 10(5), 569–588. <https://doi.org/10.1080/17512786.2015.1049645>
- Elbanna, A., & Engesmo, J. (2020, August 19). A-level results: Why algorithms get things so wrong – and what we can do to fix them. *The Conversation*. <https://theconversation.com/a-level-results-why-algorithms-get-things-so-wrong-and-what-we-can-do-to-fix-them-142879>
- Eldridge II, S. (2018). Repairing a fractured field: Dynamics of collaboration, normalization and appropriation at intersections of newswork. *The Journal of Applied Journalism & Media Studies*, 7(3), 541–559. https://doi.org/10.1386/ajms.7.3.541_1
- Elish, M. C., & boyd, d. (2018). Situating methods in the magic of Big Data and AI. *Communication Monographs*, 85(1), 57–80. <https://doi.org/10.1080/03637751.2017.1375130>
- European Commission (2018). *Artificial intelligence for Europe*. <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN>
- European Commission (2018). *Coordinated plan on artificial intelligence*. <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52018DC0795&rid=3>

- Fast, E., & Horvitz, E. (2017). Long-term trends in the public perception of artificial intelligence. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, 31(1), 963–969. <https://doi.org/10.1609/aaai.v31i1.10635>
- Fenn, J. (2010). *Hype cycle for emerging technologies*. <https://infota.siss.cl/concesiones/empresas/P05/ESSAL/13%20Estudio%20Final/Respaldo%20Gastos/Anexo%20N%203%20TIC/Gartner/Gartner/2009%20y%202010/Hype-Cycle-for-Emerging-Technologies-2010.pdf>
- Fengler, S., Lengauer, M., & Zappe, A.-C. (Eds.) (2021). *Reporting on migrants and refugees: Handbook for journalism educators*. UNESCO.
- Ferrer-Conill, R., & Tandoc, E.C. (2018). The audience-oriented editor. *Digital Journalism*, 6(4), 436–453.
- Finberg, H. I., & Klinger, L. (2014). *Core skills for the future of journalism*. The Poynter Institute of Media Studies.
- Finn, E. (2017, September 27). Art by algorithm. *Aeon*. <https://aeon.co/essays/how-algorithms-are-transforming-artistic-creativity>
- Finn, E. (2017). *What algorithms want: Imagination in the age of computing*. MIT Press.
- Flores-Saviaga, C., Savage, S., & Taroborelli, D. (2016). LeadWise: Using online bots to recruit and guide expert volunteers. *CSCW '16 Companion: Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion*, 257–260. <https://doi.org/10.1145/2818052.2869106>
- Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30(4), 681–694. <https://doi.org/10.1007/s11023-020-09548-1>
- Floridi, L., & Cowls, J. A. (2019). *Unified framework of five principles for AI in society*. <https://philarchive.org/archive/FLOAUF>.
- Fosso Wamba, S., Bawack, R. E., Guthrie, C., Queiroz, M. M., & Carillo, K. D. A. (2021). Are we preparing for a good AI society? A bibliometric review and research agenda. *Technological Forecasting and Social Change*, 164, 505–528. <https://doi.org/10.1016/j.techfore.2020.120482>
- Garvey, C., & Maskal, C. (2020). Sentiment analysis of the news media on artificial intelligence does not support claims of negative bias against artificial intelligence. *OMICS: A Journal of Interactive Biology*, 24(5), 286–299. <https://doi.org/10.1089/omi.2019.0078>
- Geiger, G. (2021, March 1). How a discriminatory algorithm wrongly accused thousands of families of fraud. *Vice*. <https://www.vice.com/en/article/jgq35d/how-a-discriminatory-algorithm-wrongly-accused-thousands-of-families-of-fraud>
- Gillespie, T. (2014). The relevance of algorithms. In: Tarleton, G., Boczkowski, P., & Foot, K. (Eds.) *Media technologies: Essays on communication, materiality, and society*. MIT Press, 167–194.
- Goodyer, J. (2021, October 14). How an AI finished Beethoven's last symphony and what that means for the future of music. *Science Focus*. <https://www.sciencefocus.com/news/ai-beethovens-symphony/>
- Graefe, A. (2016). *Guide to automated journalism*. Tow Center for Digital Journalism.
- Gutierrez-Lopez, M., Missaoui, S., Makri, S., Porlezza, C., Cooper, G., & Macfarlane, A. (2019). Journalists as design partners for AI. Paper presented at the *CHI 2019 ACM Conference on Human Factors in Computing Systems*, 4–9 May, 2019, in Glasgow.
- Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *Californian Management Review*, 61(4), 5–14.
- Hamon, R., Junklewitz, H., & Sanchez, I. (2020). *Robustness and explainability of artificial intelligence: From technical to policy solutions*. Technical report. Publications Office of the European Union.
- Hansen, M., Roca-Sales, M., Keegan, J., & King, G. (2017). *Artificial intelligence: Practice and implications for journalism*. Report from the policy exchange forum on June 13, 2017. Tow Center for Digital Journalism.
- Hanitzsch, T., Hanusch, V., Ramaprasad, J., & de Beer, A. S. (Eds.) (2019). *Worlds of journalism: Journalistic cultures around the globe*. Columbia University Press.
- Harcup, T., & O'Neill, D. (2017). What is news? News values revisited (again). *Journalism Studies*, 18(12), 1470–1488. <https://doi.org/10.1080/1461670X.2016.1150193>
- Hofkirchner, W., & Burgin, M. (Eds.) (2017). *The future information society: Social and technological problems*. New Jersey: World Scientific Publishing.
- Holton, A. E., & Belair-Gagnon, V. (2018). Strangers to the game? Interlopers, intralopers, and shifting news production. *Media and Communication*, 6(4), 70–78. <https://doi.org/10.17645/mac.v6i4.1490>
- Hu, C. (2018, December 11). In a world of “algorithmic culture,” music critics fight for relevance. *Columbia Journalism Review*. <https://www.cjr.org/analysis/algorithms-music.php>
- Hunter, M. L. (2011). *Story-based inquiry: A manual for investigative journalists*. UNESCO.
- Ideas, W. (2022, February 16). The case of the creepy algorithm that “predicted” teen pregnancy. *Wired*. <https://www.wired.com/story/argentina-algorithms-pregnancy-prediction/>
- Impe, A.-M. (2019). *Reporting on violence against women and girls: A handbook for journalists*. UNESCO.

References

- Ireton, C., & Posetti, J. (2018). *Journalism, fake news & disinformation: Handbook for journalism education and training*. UNESCO.
- Jaakkola, M. (2022). Journalists as media educators: Journalistic media education as inclusive boundary work. *Journalism Practice*, 16(6), 1265–1285. <https://doi.org/10.1080/17512786.2020.1844040>
- Jamil, S. (2019). Increasing accountability using data journalism: Challenges for the Pakistani journalists. *Journalism Practice*, 15(1), 19–40. <https://doi.org/10.1080/17512786.2019.1697956>
- Jemio, D., Hagerty, H. & Arandra, F. (2022, February 16). The case of the creepy algorithm that “predicted” teen pregnancy. *Wired*. <https://www.wired.com/story/argentina-algorithms-pregnancy-prediction/>
- Jia, P., & Xia, Q. (2008). Knowledge management, collaboration and innovation. 2008 International Seminar on Future Information Technology and Management Engineering, Leicestershire, UK, 2008, 228–232. <https://doi.org/10.1109/FITME.2008.104>
- Jung, J., Song, H., Kim, Y., Im, H., & Oh, S. (2017). Intrusion of software robots into journalism: The public’s and journalists’ perceptions of news written by algorithms and human journalists. *Computers in Human Behavior*, 71(2017), 291–298. <https://doi.org/10.1016/j.chb.2017.02.022>
- Kalamkar, V. B. S., Nouri, N., Ofek, E., Pahud, M., Dudley, J. J., Hu, J., & Grubert, J. (2022). Quantifying the effects of working in VR for one week. arXiv preprint, arXiv:2206.03189. <https://doi.org/10.48550/arXiv.2206.03189>
- Kantor, J. (2014, August 14). Working anything but 9 to 5. *The New York Times*. <https://www.nytimes.com/interactive/2014/08/13/us/starbucks-workers-scheduling-hours.html>
- Karlsen, J., & Stavelin, E. (2014). Computational journalism in Norwegian newsrooms. *Journalism Practice*, 8(1), 34–48. <https://doi.org/10.1080/17512786.2013.813190>
- Keefe, J., Zhou, Y., & Merrill, J. B. (2021). *The present and potential of AI in journalism*. The Knight Foundation. <https://knightfoundation.org/articles/the-present-and-potential-of-ai-in-journalism/>
- Keith, S., & Cozma, R. (2022, forthcoming). *Teaching journalism online*. UNESCO.
- Kim, H., Kim, S.-H., & Frear, C. (2016). News photos of scientists skew race but not gender. *Newspaper Research Journal*, 37(1), 261–274. <https://doi.org/10.1177/07395329166663210>
- Knowledge Center Data & Society. (n.d.). *Knowledge Center Data & Society*. <https://data-en-maatschappij.ai/journalisten>
- Kothari, A., & Cruikshank, S. A. (2022). Artificial intelligence and journalism: An agenda for journalism research in Africa. *African Journalism Studies*, 43(1), 17–33. <https://doi.org/10.1080/23743670.2021.1999840>
- Köuts-Klemm, R. (2019). Data literacy among journalists: A skills-assessment based approach. *Central European Journal of Communication*, 12(3), 299–315. [https://doi.org/10.19195/1899-5101.12.3\(24\).2](https://doi.org/10.19195/1899-5101.12.3(24).2)
- Laranjeira de Pereira, J., & Guimarães Moraes, T. (2022, February 14). Promoting irresponsible AI: Lessons from a Brazilian bill. *Heinrich Böll Stiftung*. <https://eu.boell.org/en/2022/02/14/promoting-irresponsible-ai-lessons-brazilian-bill>
- LeCompte, C. (2015). Automation in the newsroom: How algorithms are helping reporters expand coverage, engage audiences, and respond to breaking news. *Nieman reports*. <https://niemanreports.org/articles/automation-in-the-newsroom/>
- Leitner, M. (2018). *How media companies use data to sign up digital subscribers (and keep them)*. Reuters Institute Fellowship Paper, University of Oxford.
- Lewis, S., Guzman, A. L., & Schmidt, T. S. (2019). Automation, journalism, and human-machine communication: Rethinking roles and relationships of humans and machines in news. *Digital Journalism*, 7(4), 409–427. <https://doi.org/10.1080/21670811.2019.1577147>
- Lewis, S. C. (2012). The tension between professional control and open participation. *Information, Communication & Society*, 15(6), 836–866. <https://doi.org/10.1080/1369118X.2012.674150>
- Lewis, S. C., & Usher, N. (2014). Code, collaboration, and the future of journalism. *Digital Journalism*, 2(3), 383–393. <https://doi.org/10.1080/21670811.2014.895504>
- Lewis, S. C., & Simon, F. M. (2022). Why human-machine communication matters for the study of journalism and artificial intelligence. In: Guzman, A. L., McEwen, R., & S. Jones, S. (Eds.) *SAGE Handbook of Human-Machine Communication*.
- Lindén, C.-G. (2020). What makes a reporter human? A research agenda for augmented journalism. *Questions de communication*, 37(2020), 337–352 <https://doi.org/10.4000/questionsdecommunication.23301>
- Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151, 90–103. <https://doi.org/10.1016/j.obhdp.2018.12.005>
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–16. <https://doi.org/10.1145/3313831.3376727>
- LSE, London School of Economics and Political Science. (n.d.). *JournalismAI Starter Pack*. <https://www.lse.ac.uk/media-and-communications/polis/JournalismAI/Starter-Pack>

- Madiega, T. (2019). *EU guidelines on ethics in artificial intelligence: Context and implementation*. European Parliamentary Research Service (EPRS).
- Magnusson, M., Finnäs, J., & Wallentin, L. (2016). Finding the news lead in the data haystack: Automated local data journalism using crime data. Paper presented vid Computation & Journalism Symposium. Stanford University, Palo Alto, CA. [http://journalism.stanford.edu/cj2016/files/Finding the news lead in the data haystack.pdf](http://journalism.stanford.edu/cj2016/files/Finding%20the%20news%20lead%20in%20the%20data%20haystack.pdf)
- Malinverni, L., Valero, C., Schaper, M. M., & de la Cruz, I. G. (2021). Educational robotics as a boundary object: Towards a research agenda. *International Journal of Child-Computer Interaction*, 29. <https://doi.org/10.1016/j.ijcci.2021.100305>
- Marckmann Rasmussen, A. (2020). *Sådan forandrer automatiseringen medierna*. Dansk Journalistforbund.
- Marconi, F., & Siegman, A. (2017). Machine journalist. 2017. *The future of augmented journalism: A guide for newsrooms in the age of smart machines*. Associated Press.
- Mayor, A. (2018). *Gods and robots: Myths, machines, and ancient dreams of technology*. Princeton University Press.
- McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Magazine*, 27(4), 12–12. <https://doi.org/10.1609/aimag.v27i4.1904>
- McGregor, S. C. (2019). Social media as public opinion: How journalists use social media to represent public opinion. *Journalism*, 20(8), 1070–1086. <https://doi.org/10.1177/1464884919845458>
- Merrill, J. B. (2020, January 19). How Quartz used AI to sort through the Luanda Leaks. *Quartz*. <https://qz.com/1786896/ai-for-investigations-sorting-through-the-luanda-leaks/>
- Merten, M. (2022, March 10). Is the SA revenue service's risk algorithm the glitch in the tax collector's matrix? *Daily Maverick*. Available at <https://www.dailymaverick.co.za/article/2022-03-08-is-the-sa-revenue-services-risk-algorithm-the-glitch-in-the-tax-collectors-matrix/>
- Miao, F., Holmes, W., Ronghuai, H. & Hui, Z. (2021). *AI and education: Guidance for policy-makers*. UNESCO.
- Milosavljević, M., & Vobić, I. (2021). 'Our task is to demystify fears': Analysing newsroom management of automation in journalism. *Journalism*, 22(9), 2203–2221. <https://doi.org/10.1177/1464884919861598>
- Moran, R. E., & Shaikh, S. J. (2022). Robots in the news and newsrooms: Unpacking meta-journalistic discourse on the use of artificial intelligence in journalism. *Digital Journalism*, 10(10), 1756–1774. <https://doi.org/10.1080/21670811.2022.2085129>
- Muratova, N., Grizzle, A., & Mirzakhmedova, D. (2021). *Media and information literacy in journalism: A handbook for journalists and journalism educators*. UNESCO.
- Munoriyarwa, A., Chiumbu, S., & Motsaathebe, G. (2021). Artificial intelligence practices in everyday news production: The case of South Africa's mainstream newsrooms. *Journalism Practice*, 1(19). <https://doi.org/10.1080/17512786.2021.1984976>
- Nast, C. (n.d.). The artificial intelligence database. *Wired*. Available at <https://www.wired.com/category/artificial-intelligence/>
- Natale, S., & Ballatore, A. (2017). Imagining the thinking machine: Technological myths and the rise of artificial intelligence. *Convergence*, 16(1), 3–18. <https://doi.org/10.1177/1354856517715164>
- Nerone, J. (2013). The historical roots of the normative model of journalism. *Journalism*, 14(4), 446–458. <https://doi.org/10.1177/146488491246417>
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2. <https://doi.org/10.1016/j.caeai.2021.100041>
- Nordahl, L., Chavar, A. J., Porter, L. Z., Kim, M., & Blufarb, A. (2021, July 21). Using computer vision to create a more accurate digital archive. *The New York Times R&D*. <https://rd.nytimes.com/projects/using-computer-vision-to-create-a-more-accurate-digital-archive>
- North, L. (2016). The gender of "soft" and "hard" news. *Journalism Studies*, 17(3), 356–373. <https://doi.org/10.1080/1461670X.2014.987551>
- O'Neil, C. (2016). *Weapons of math destruction*. Crown.
- Ouchchy, L., Coin, A., & Dubljević, V. (2020). AI in the headlines: The portrayal of the ethical issues of artificial intelligence in the media. *AI & Society*, 35, 927–936. <https://doi.org/10.1007/s00146-020-00965-5>
- Nilsson, N. J. (2010). *The quest for artificial intelligence: A history of ideas and achievements*. Cambridge University Press.
- Paulussen, S. (2016). Innovation in the newsroom. In: Witschge, T., Anderson, C.W., Domingo, D., & Hermida, A. (Eds.) *The SAGE handbook of digital journalism*. Sage, 192–206.
- Pentzold, C., Brantner, C., & Fölsche, L. (2019). Imagining big data: Illustrations of "big data" in US news articles, 2010–2016. *New Media & Society*, 21(1), 139–167. <https://doi.org/10.1177/1461444818791326>
- Petre, C. (2018). Engineering consent: How the design and marketing of newsroom analytics tools rationalize journalists' labor. *Digital Journalism*, 6(4), 509–527. <https://doi.org/10.1080/21670811.2018.1444998>

References

- Reddy, E., Hoople, G., & Choi-Fitzpatrick, A. (2019). Interdisciplinarity in practice: Reflections on drones as a classroom boundary object. *Engineering Studies*, 11(1), 51–63. <https://doi.org/10.1080/19378629.2019.1614006>
- Romele, A. (2022). Images of artificial intelligence: A blind spot in AI ethics. *Philosophy & Technology*, 35(4). <https://doi.org/10.1007/s13347-022-00498-3>
- Roselli, D., Matthews, J., & Talagala, N. (2019). Managing bias in AI. *WWW'19 Companion proceedings of the 2019 World Wide Web Conference*, 539–544. <https://doi.org/10.1145/3308560.3317590>
- Ross Arguedas, A., Robertson, C., Fletcher, R., & Nielsen, R. (2022). *Echo chambers, filter bubbles, and polarisation: A literature review*. Reuters Institute for the Study of Journalism.
- Russell, S., & Norvig, P. (2010). *Artificial intelligence: A modern approach*. 3rd ed. Pearson.
- Samoili, S., Cobo, M. L., Gomez, E., De Prato, G., Martinez-Plumed, F., & Delipetrev, B. (2020). *AI watch – Defining artificial intelligence: Towards an operational definition and taxonomy of artificial intelligence*. Technical Report. Joint Research Centre.
- Schumpeter, J. A. (1994). *Capitalism, Socialism and Democracy*. Routledge.
- Sedola, S. (2021). *Artificial intelligence for Africa – blueprint*. <https://smartafrica.org/knowledge/artificial-intelligence-for-africa/>
- Shi, Y., & Herniman, J. (2023). The role of expectation in innovation evolution: Exploring hype cycles. *Technovation*, 119(C). <https://doi.org/10.1016/j.technovation.2022.102459>
- Siles, I., & Boczkowski, P. J. (2012). Making sense of the newspaper crisis: A critical assessment of existing research and an agenda for future work. *New Media & Society*, 14(8): 1375–1394. <https://doi.org/10.1177/1461444812455148>
- Simon, H. A. (1995). Artificial intelligence: An empirical science. *Artificial Intelligence*, 77(1), 95–127. [https://doi.org/10.1016/0004-3702\(95\)00039-H](https://doi.org/10.1016/0004-3702(95)00039-H)
- Simons, M., Tiffen, R., Hendrie, R., Carson, A., Sullivan H., Muller, D., & McNair, B. (2017). Understanding the civic impact of journalism. *Journalism Studies*, 18(11), 1400–1414.
- Sivadas, L., & Argoub, S. (2021, April 7). How to report on artificial intelligence. *Journalism.co.uk*. <https://www.journalism.co.uk/news/how-to-report-on-ai-effectively/s2/a811366/>
- Stalph, F. (2017). Classifying data journalism: A content analysis of daily data-driven stories. *Journalism Practice*, 12(10), 1332–1350. <https://doi.org/10.1080/17512786.2017.1386583>
- Stieglietz, S., Brachten, F., Ross, B., & Jung, A.-K. (2017). *Do social bots dream of electric sheep? A categorisation of social media bot accounts*. A paper presented at Australasian Conference on Information Systems, Hobart, Australia.
- Stray, J. (2019). Making artificial intelligence work for investigative journalism. *Digital Journalism*, 7(8), 1076–1097. <https://doi.org/10.1080/21670811.2019.1630289>
- Striphas, T. (2015). Algorithmic culture. *European Journal of Cultural Studies*, 18(4–5), 395–412. <https://doi.org/10.1177/1367549415577392>
- Sundar, S. S., Kang, J. & Oprean, D. (2017). Being there in the midst of the story: How immersive journalism affects our perceptions and cognitions. *Cyberpsychology, Behavior, and Social Networking*, 20(11), 672–682. <https://doi.org/10.1089/cyber.2017.0271>
- Thurman, N., Dörr, K., & Kunert, J. (2017). When reporters get hands-on with robo-writing. *Digital Journalism*, 5(10), 1240–1259. <https://doi.org/10.1080/21670811.2017.1289819>
- Turing, A. M. (2009). Computing machinery and intelligence. In: Epstein, R., Roberts, G., & Beber, G., (Eds.) *Parsing the Turing test: Philosophical and methodological issues in the quest for the thinking computer*. Springer, 23–65.
- UNESCO (2021). *Recommendation on the ethics of artificial intelligence*. <https://en.unesco.org/artificial-intelligence/ethics>
- United Robots (2021). *News automation and how to leverage it*. Whitepaper. <https://www.unitedrobots.ai/whitepaper-download?hsLang=en>
- University of Sheffield (2022). *Data literacy for journalists*. An online course. <https://www.futurelearn.com/courses/data-literacy-for-journalists/1>
- van Dalen, A. (2012). The algorithms behind the headlines: How machine-written news redefines the core skills of human journalists. *Journalism Practice*, 6(5–6), 648–658. <https://doi.org/10.1080/17512786.2012.667268>
- Venturini, T. (2019). From fake to junk news: The data politics of online virality. In: Bigo, D., Isin, E., & Ruppert, E. (Eds.), *Data politics: Worlds, subjects, rights*. Routledge, 123–144.
- Vergeer, M. (2020). Artificial intelligence in the Dutch Press: An analysis of topics and trends. *Communication Studies*, 71(3), 373–392. <https://doi.org/10.1080/10510974.2020.1733038>

- Viswanath, K., Blake, K. D., Meissner, H. I., Saiontz, N. G., Mull, C., Freeman, C. S., Hesse, B., & Croyle, R. T. (2008). Occupational practices and the making of health news: A national survey of US health and medical science journalists. *Journal of Health Communication*, 13(8), 759–777. <https://doi.org/10.1080/10810730802487430>
- Waldenström, A., Wiik, J., & Andersson, U. (2018). Conditional autonomy: Journalistic practice in the tension field between professionalism and managerialism. *Journalism Practice*, 1(16). <https://doi.org/10.1080/17512786.2018.1485510>
- Welbers, K., van Atteveldt, W., Kleinnijenhuis, J., Ruigrok, N., & Schaper, J. (2016). News selection criteria in the digital age: Professional norms versus online audience metrics. *Journalism*, 17(8), 1037–1053. <https://doi.org/10.1177/1464884915595474>
- Westlund, O., Krumsvik, A. H., & Lewis, S. C. (2021). Competition, change, and coordination and collaboration: Tracing news executives' perceptions about participation in media innovation. *Journalism Studies*, 22(1), 1–21. <https://doi.org/10.1080/1461670X.2020.1835526>
- Winfield, C., Sparkman-Key, N. M., & Vajda, A. (2017). Interprofessional collaboration among helping professions: Experiences with holistic client care. *Journal of Interprofessional Education & Practice*, 9(1), 66–73. <https://doi.org/10.1016/j.xjep.2017.08.004>
- Women's Media Center (2014). *The status of women in the U.S. wedia 2014*. https://wmc.3cdn.net/6dd3de8ca65852dbd4_fjm6yck9o.pdf
- Women's Media Center (2019). *The status of women in the U.S. wedia 2019*. <https://womensmediacenter.com/assets/site/from-bsd/WMCStatusofWomeninUSMedia2019.pdf>
- Woolgar, S. (1985). Why not a sociology of machines? The case of sociology and artificial intelligence. *Sociology*, 19(4), 557–572. <https://doi.org/10.1177%2F0038038585019004005>.
- Woolgar, S. (Ed.) (2003). *Virtual society? Technology, cyberbole, reality*. Oxford University Press.
- Wu, S., Tandoc, E.C. Jr., & Salmon, C. T. (2019). Journalism reconfigured: Assessing human–machine relations and the autonomous power of automation in news production. *Journalism Studies*, 20(10), 1440–1457. <https://doi.org/10.1080/1461670X.2018.1521299>
- Xiaomo, L., Nourbakhsh, A., Li, Q., Shah, S., Martin, R., & Duprey, J. (2017). *Reuters tracer: Toward automated news production using large scale social media data*. Research and Development. Thomson Reuters.
- Young, M. L., & Hermida, A. (2015). From Mr. and Mrs. Outlier to central tendencies: Computational journalism and crime reporting at the Los Angeles Times. *Digital Journalism*, 3(3), 381–397. <https://doi.org/10.1080/21670811.2014.976409>
- Yudkowsky, E. (2001). *Creating friendly AI 1.0: The analysis and design of benevolent goal architectures*. The Singularity Institute.
- Zachrisson, O. (2021, November 21). An algorithm to promote unique public service journalism. *EBU Tech-I Magazine*.
- Zelizer, B. (1993). Journalists as interpretive communities. *Critical Studies in Mass Communication*, 10(3), 219–237. <https://doi.org/10.1080/15295039309366865>
- Zelizer, B. (2004). *Taking journalism seriously: News and the academy*. Sage Publications.
- Zhai, Y., Yan, J., Zhang, H., & Lu, W. (2020). Tracing the evolution of AI: Conceptualization of artificial intelligence in mass media discourse. *Information Discovery and Delivery*, 48(3), 137–149.
- Zhang, S., & Feng, J. (2019). A step forward? Exploring the diffusion of data journalism as journalistic innovations in China. *Journalism Studies*, 20(9), 1281–1300. <https://doi.org/10.1080/1461670X.2018.1513814>
- Zhang, W. & Pérez Tornero, J. M. (2021). Introduction to AI journalism: Framework and ontology of the trans-domain field for integrating AI into journalism. *Journal of Applied Journalism & Media Studies*, published online September 3, 2021. https://doi.org/10.1386/ajms_00063_1

Table of figures

Figure 1: Journalistic competence areas related to reporting on AI	22
Figure 2: Key AI Subfields	32
Figure 3: Natural Language Processing (NLP) Hierarchy	33
Figure 4: Evolution of HAL	41
Figure 5: The Trolley Problem	43
Figure 6: AI Global Landscape	51
Figure 7: Ethics guidelines for trustworthy AI – an overview	54
Figure 8: Four ethical principles based on fundamental rights	55
Figure 9: The assessment list for trustworthy AI (ALTAI)	55
Figure 10: A risk-based approach	56
Figure 11: The OECD AI Principles and Recommendations	57
Figure 12: The UNESCO Recommendation on the Ethics of AI	57
Figure 13: Beyond 'Fake News'	71
Figure 14: Ian Goodfellow's tweet (14 January 2019)	75
Figure 15: Media attention dedicated to emerging technologies (Gartner's hype cycle)	84
Figure 16: A screenshot of a Google search for images of AI (February 2022)	86
Figure 17: Primary Purpose of AI in Journalism (2012-2021)	97
Figure 18: Where in the story pipeline is AI commonly implemented? (2012-2021)	98

The rise and control of artificial intelligence (AI) is impacting society as a whole, and journalistic coverage needs to keep up with the implications. This handbook covers:

- Understanding machine intelligence and identifying different types of AI
- Imagining diverse futures with AI by recognising pervasive popular narratives that inform public consciousness
- Understanding journalism's role in mediating and shaping AI discourse
- Finding ways of reporting about AI in a nuanced, realistic and accountable manner

Produced under the auspices of the World Journalism Education Council, the publication is supported by UNESCO's International Programme for the Development of Communication (IPDC).



unesco

The International
Programme for
the Development
of Communication

Researched and edited by



World Journalism Education Council (WJEC)
www.wjec.net/

