LUISS

Corso di laurea in Data Science and Management

Cattedra di Data Science in Action

JournAIsm: how AI is changing the field of journalism

Prof. Alessio Martino

RELATORE

Prof. Giuseppe Italiano

CORRELATORE

Ignazio Leonardo Scarpelli

CANDIDATO

Anno Accademico 2024/2025

ABSTRACT

This Master's thesis focus on the integration of Artificial Intelligence (AI) in the world of journalism. As AI is approaching and changing the landscape of the world as we know it, journalism does not represent an exception, with AI reshaping the way news are collected, produced and distributed. The implementation of algorithms such as Latent Dirichlet Allocation (LDA) or Transformer based architectures such as Large Language Models (LLMs) like GPT (Generative Pretrained Transformer) or BERT (Bidirectional Encoder Representations from Transformers) inside the newsrooms is transforming the journalist's figure and the newsroom's workflow.

This work highlights how artificial intelligence is enhancing the work of journalists through a combination of a detailed literature review on the algorithms and their specific use, case studies of integration into newsrooms, interviews with journalists, news professionals and intelligence officers. It underlines how the news professionals are taking advantages from the use of AI, especially in tasks such as data analysis, content creation and interviews transcription, and how the news industry is being disrupted by the technology while also addressing the risks that AI keeps with itself like easing disinformation, biases, and fear of job displacement. The thesis also contains a report on Italian journalists' sentiment and actual use of artificial intelligence, with data collected through a survey visible on Appendix B of this work.

In conclusion, this dissertation understands the advantages that artificial intelligence can bring to journalism, but at the same time emphasizes the need for ethical guidelines, journalists' education, and "man-machine" collaboration to preserve the integrity of the sector while embracing the full potential of the technology.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

This thesis contains the work of nine months made of researching, reading many articles, desperately collecting answers for my survey, scheduling interviews, meetings, and analyses. This section is here to thank all the people that helped me finishing a job that instead would have taken much longer.

First of all, I want to thank my thesis advisor, Prof. Alessio Martino, for giving me the opportunity to choose this topic, the freedom of developing it as I wished and for offering advices on how to improve it. I want to thank also the members of the Luiss DataLab and the Luiss School of Journalism that gave me the opportunity to be an intern at their organization. In particular, I would like to mention Lorenzo Federico, whose company and advice made my thesis way better, Gian Marco Passerini and Elisabetta Panico, who patiently put up with my endless requests for survey responses, Federica Urzo and Gianni Riotta, who accepted my application and made me a part of their team.

I am also deeply grateful to Lorenzo Canale, Beatrice Petrella, Michael Piemonte and Giulia Pozzi that took time off their schedule to be interviewed by me. Their contribution added a great value to this work and made it more effective.

This thesis also marks the end of five years as a student at Luiss University. I want to thank all the people that I've met in this long period of time, and this includes professors, faculty staff, friends and people I do not get along with. In particular I want to express my gratitude to the magnificent people and staff at the Luiss Language Cafè and at the Luiss Loft, where I've spent the best moments of this period.

Last, but probably the most important, I want to thank my family for the unwavering support during this period. Without them, the experiences I've had during these five years, together with this thesis, would have never existed.

INDEX OF FIGURES

Figure 8: Performance of noise filtering on benchmark data. **Source:** Research and Development Team @ Tomson Reuters, *Reuters Tracer: A Large Scale System of Detecting & Verifying Real-Time News Events from Twitter*, 2016.

Figure 9: Performance of noise filtering on streaming data. **Source:** Research and Development Team @ Tomson Reuters, *Reuters Tracer: A Large Scale System of Detecting & Verifying Real-Time News Events from Twitter*, 2016.

Figure 10: Performance of event clustering using LSH with various (k, h) & our algorithm on benchmark data. **Source:** Research and Development Team @ Tomson Reuters, *Reuters Tracer: A Large Scale System of Detecting & Verifying Real-Time News Events from Twitter*, 2016.

Figure 11: Scaled Dot-Product Attention (left) and Multi-Head Attention (right). **Source:** Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, I. Polosukhin, *Attention is all you need,* arXiv preprint, 2017.

Figure 12: The Transformer achieves better BLEU scores than previous state-of-the-art models on the English-to-German and English-to-French newstest2014 tests at a fraction of the training cost. **Source:** Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, I. Polosukhin, *Attention is all you need,* arXiv preprint, 2017.

Figure 13: BERT input representation. **Source:** Google AI Language Team (Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova), *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,* arXiv preprint, October 2018.

Figure 14: BERT scores compared to other models. **Source:** Google AI Language Team (Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova), *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,* arXiv preprint, October 2018.

INDEX OF TABLES

CHAPTER ONE

INTRODUCTION

In modern days, a widely discussed topic that fills newsrooms, politics and TV shows regards Artificial Intelligence (AI) and everything that slightly resembles to have something in common with it. Even if the origins of the field of Artificial Intelligence can be traced back to the mid-50s, it is only now that the topic is breaking into the general public, thanks to technological acceleration, data availability and mainstream roll-out of Large Language Models (LLMs), such as Mistral or ChatGPT, which have shown tangible potential of such technology even to non-experts in the field. Now we know for certain that AI will enter in almost every field of the society, and the development of LLM has just accelerated this process. In an oft-cited report[1] by Katja Grace published in August 2022, just four months before ChatGPT 3.5 was released to the public, 50% of the 352 interviewed experts predicted that human-level AI would exist before 2061, with this time now surely reduced after the recent technological advancements. A lot of fields have been affected by AI even before the arrival of ChatGPT, such as healthcare or construction, and journalism makes no exception. All the phases of a journalist's work are touched by AI, from information retrieval to content dissemination. One of the main areas in which journalist's work is more impacted by AI is content production, with more and more news organization that are automizing this phase[2], especially by leveraging LLMs for content production. This has caused many concerns about the possible atomization of the whole process of creating an article or a news story, leading to a loss of jobs for journalists. Some more moderate and less pessimistic comments state that AI will just be a tool to help journalists in doing the same job in less time, improving the quality of stories by making the journalists concentrate on looking for more data and more ways to improve their research and create a better content. Trying to better analyze these opinions will be the matter and the purpose of this thesis: understand how journalism is being affected by AI by analyzing the main models used and their purpose, analyze their results and their impact, understanding the opinion of journalists and experts, and try, with the means at my disposal and with some experiments, to predict how AI will affect the journalism sector

[1] https://aiimpacts.org/2022-expert-survey-on-progress-in-ai/#Chance_that_the_intelligence_explosion_argument_is_about_right

[2] https://edition.cnn.com/2023/01/26/media/buzzfeed-ai-content-creation/index.html

in the next years. The remainder of this work is organized as follows: after the Introduction, the second chapter will be purely technical, analyzing models that are currently use in every phase of a journalist's work.

The third chapter will be more 'journalistic'. Specifically, I will explain how AI is impacting journalism by creating new professions in the field and underlining both the challenges, such as the importance of the journalist's education on the topic, and the risks of a massive use of the technology in the field (fake news, public opinion's polarization, possible loss of jobs…)

Chapter four will be a report on how Italian journalists are approaching the new disruption represented by AI in their workflow, how they are implementing it and what are their thoughts about the future impact of this technology.

The fifth and final chapter will draw the conclusions of this work, underlining my final thoughts about the immediate impact of AI in journalism and its future consequences by balancing both risks and advantages of the technology.

After the final chapter there will be two appendices:

- Appendix A will contain all the interviews made by the candidate.
- Appendix B will contain the questionnaire that was proposed to Italian journalists to understand their actual use and sentiment on artificial intelligence.

CHAPTER TWO

A LITERATURE REVIEW OF JOURN(AI)SM

2.1 The Process of News

Journalists are using AI right now during their work to look for news, and ideas, and create content. Now, AI is highly integrated and widely used in newsrooms, and it is seen as a tool that helps journalists in their work, making them do the same amount of work in less time.

Being inspired by this video[3], I divided the process of a journalist's work into **six phases:**

- **Idea,** where journalists find arguments for news stories and contents.
- **Research,** in which journalists take a deep dive into the argument by acquiring and validating the information and scanning datasets.
- **Production,** where the content is effectively produced. This also includes the review of the content, which might consist of checking typos and the correctness of an article or checking for errors in video content.
- **Publication,** this is the phase where the content is put at the disposition of the public. Here usually, recommendations are made to disseminate the content, and more versions of the same content and subtitles are created.
- **Feedback,** here, the article receives comments and feedback from the public and journalists can interact with the audience. This is the phase where fact-checkers make their entry into the process, verifying if the information given in the content is misleading.
- **Archiving,** the last phase of the process is archiving the content, which means tagging it and creating relations with other content.

In this chapter, for every phase, we will explore the machine learning techniques that are and can be used by journalists. In this sense, we will analyze every algorithm, the papers that analyze them

[3] https://www.youtube.com/watch?v=qH9KWXHIM9U

and, by so, make a complete literature review of every algorithm presented. Later, when possible, we will provide practical examples of the usage of the algorithm in journalism.

2.2 Latent Dirichlet Allocation (LDA)

Latent Dirichlet Allocation (LDA) was described, in its introductory paper (David M. Blei et al., 2003) as 'a generative probabilistic model for the collections of discrete data such as text corpora'. It is a topic modeling algorithm, whose goal is to understand the topic of a certain document. LDA is a Bayesian model and, particularly, is part of the Latent Variable Models family, a set of statistical models that tries to explain complex phenomena by incorporating both observed and unobserved (latent) variables. LDA, in particular, differentiates itself from other similar models such as Unigram and Latent Semantic Indexing (LSI) because of its structure.

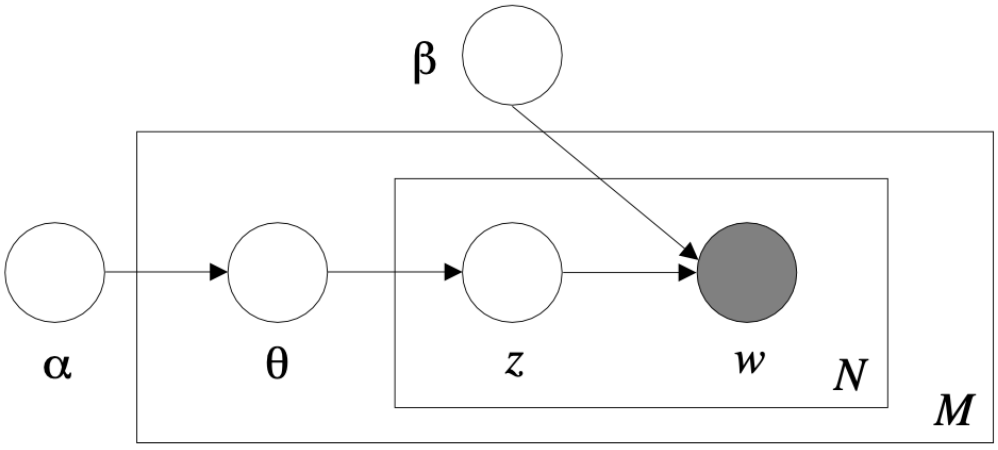


Figure 1: Graphical model representation of LDA. **Source:** David M. Blei, Andrew Y. Ng, Michael I. Jordan, "*Latent Dirichlet Allocation*", Journal of Machine Learning Research (993-1022), 1st March 2003.

Figure 1 depicts the structure of the LDA algorithm as figured in the original paper. To understand the blueprint, z represents the topic of a word, w are the words of such document, N is the sequence of words w in the document, and M is the number of documents in the *corpus,* denoted by D. α represents the parameters of the Dirichlet distribution before the per-document distributions, while

β represents a k x V matrix, where k is the dimensionality of the Dirichlet distribution, and so the dimensionality of the topic variable z, and V is a vector of all the words w, and so the vocabulary size. The basic idea behind LDA is that, quoted from Blei's paper, 'documents are represented as random mixtures over latent topics, where each topic is characterized by a distribution over words', and, for each document in M, the algorithm goes through a generative process that reads every unlabeled document in a corpus several times to catch patterns between some particular words and their repetition in a document to understand the topic of each argument. The Bayesian nature of LDA is clear also if we see its mathematical formula.

$$p(\theta, z, w|\alpha, \beta) = p(\theta|\alpha) \prod_{n=1}^{N} p(z_n|\theta) p(w_n|z_n, \beta).$$

In simple terms, this formula represents the probability of generating the exact same document as the one in exam from the *corpus*.

The final goal, regarding this formula, is to optimize α, β and θ in order to maximize the probability $p(\theta, \mathbf{z}, \mathbf{w}|\alpha, \beta)$ of replicating the exact same document as the one in exam. In order to do so, there are many techniques that can be used. The most famous ones are:

- **Gibbs Sampling,** a Markov Chain Monte Carlo (MCMC) method used to approximate the posterior distribution of the variables of interest by generating samples from the distribution after several iterations that, in LDA, maximize the probability function by iterating on the corpus in order to uniform as more as possible the assignment of a topic for each word.
- **Variational Inference (VI),** a deterministic alternative to MCMC methods like Gibbs Sampling. It approximates the true posterior distribution with a simpler, parameterized distribution by minimizing the Kullback-Leibler (KL) divergence between the two. In LDA, variational inference algorithms approximate the posterior distribution of the hidden variables with a factorized distribution and then optimize the parameters of this distribution to maximize the likelihood of the observed data.
- **Expectation-Maximization (EM) Algorithm,** in its general form, is used to find maximum likelihood estimates of parameters in statistical models with latent variables, like LDA.

In this work, we will explain the Gibbs Sampling method, which is the simplest and most used method in general for maximizing the probability in Latent Dirichlet Allocation. Before introducing Gibbs Sampling, I first need to explain what an MCMC (Markov Chain Monte Carlo) is.

A Markov Chain is a random process with the characteristic of not retaining memory of where it was in the past (J.R. Norris, 1998). This means that the next state of the process is only affected by its current state. In simpler terms, the probability of having a value X in a time t given the previous iterations is equal to the probability of X_t given only the value of the previous iteration X_{t-1}. This is called the Markov property, and it is mathematically represented in the equation below:

$$p(x_t|x_0, x_1, \dots, x_{t-1}) = p(x_t|x_{t-1}).$$

In simpler terms, the probability of every variable in the chain of taking a certain value is affected only by the preceding variable, and not by other variables' values in the chain.
Now, in an MCMC, the Monte Carlo simulation involves generating a large number of sample paths of the Markov Chain. The chain can start either with an initial state or a distribution over states and then can use sample paths to estimate various characteristics of the system. In a Markov Chain, this is important to arrive at the goal of reaching a stationary distribution, which means that the distribution remains unchanged after a certain number of iterations.

Gibbs Sampling is a specialized MCMC method that is used to generate a sequence of samples from the joint distribution of multiple variables. Gibbs Sampling's goal maximizes the probability function of LDA by reaching two important goals:

- It ensures that for every word in V, the majority of its appearances are assigned to one single topic.
- For every document in N, the majority of the words are assigned to a single topic.

LDA, as said before, is a topic modeling algorithm that automatically identifies topics from a set of documents named *corpus.* Journalists can use it to understand trends (i.e. perform a trend analysis) and, consequently, collect new ideas to produce interesting news stories. At the same time, newsrooms and news brands can use LDA to perform audience and sentiment analysis to build their editorial and content strategy. The generality of LDA makes it a very versatile algorithm that can be used not only in the phase of the idea but also in other phases of a news story's creation. In particular:

- LDA can be used in the last phase (i.e. archiving) in the form of automatic tagging of articles and the creation of content relations in the building of news recommendation systems.
- Latent Dirichlet Allocation can also be used to understand biases around media often associated with particular political views.
- Topic modeling can also be used in investigative journalism. In particular, it can help in identifying the relevant documents by uncovering the main topics they cover, thereby making the investigative process more efficient.

2.3 Web Scraping

As I will underline in the next chapter, one of the major causes of the disruption of traditional journalism's business model is the massive quantity of data generated and the velocity with which breaking news must be published. Especially in news agencies, this change of scenario must be met with new tools that speed up the process of collecting news. One of these tools is web scraping.

Web scraping is a series of techniques that allow to automatically retrieve and extract data from a web page in order to be then wrapped into a database. In order to extract data a scraper, or wrapper, is needed to act as a medium between the user and the website. A simplified visual scheme of how web scraping works can be found in Figure 2.

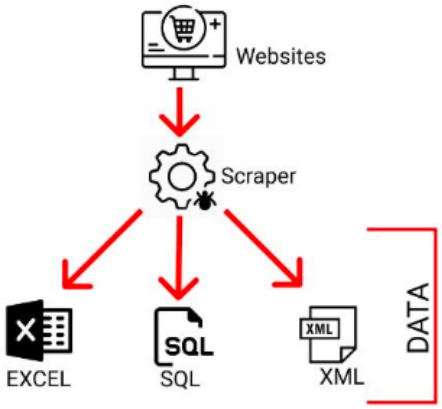


Figure 2: Simplified representation of Web Scraping's structure. **Source:** P. Matta, N. Sharma, D. Sharma, B. Pant, S. Sharma, "*Web Scraping: Applications and Scraping Tools*", International Journal of Advanced Trends in Computer Science and Engineering (8202-8206), September 2020.

In the figure above the format of the data extracted is either in spreadsheet, database, or markup language, but a lot of web scraping applications produce other file types, like JSON (very useful when extracting data from APIs) or PDF.

A scraper involves two different technologies (Lotfi et al., 2021):

- A crawler, and so a bot that extracts data from websites by visiting them. The key to a web crawler's work is the search for links inside web pages (Relative URL) to then going on searching for the URL's base (Absolute URL). There are many types of web crawlers:
    - Focused, a type of web crawler that only looks for web pages related to certain fields or subjects.
    - Incremental, that visit and access updated web pages.
    - Distributed, a crawler that assigns crawling to other crawlers.
    - Parallel, where multiple crawler processes are combined so that each process performs the process of filtering and retrieving the URLs.
    - Hidden, which filters the content of websites not accessible to general users.
- A parser, that is used by programmers to extract certain details from data.

There are many techniques used to perform web scraping, depending on the type of pages the data are from. Some of them are:

- **HTML Parsing** is a technique that analyses the HTML (HyperText Markup Language) code of a webpage and extracts relevant data.
- **DOM Parsing** interacts with the Document Object Model (DOM) of a webpage. This is useful for pages that load dynamic content using JavaScript after the initial page load. HTML and DOM parsing can also be used together in some occasions.
- Many websites offer APIs (Application Program Interfaces) that contain data in various types of formats, such as JSON and XML.
- Some libraries, such as Python's Selenium[4], automate browsing actions such as clicking buttons and filling forms, in order to access data that are not immediately available from the page source.

Web scraping can be performed in many programming languages (Python, R, etc.) with some specific libraries such as BeautifulSoap, useful to perform HTML parsing, and the just-cited Selenium, although there are also frameworks such as Scrapy. The 'request' module in Python can be used to parse data from APIs, specifically by sending HTTP requests to the website involved.

The web scraping techniques I have introduced, and the data extracted with them, can be enhanced in many ways, for example by creating immersive data visualization, or attractive User Interfaces (UI).

Web scraping can have many uses in the majority phases of journalism:

- One of the most immediate uses of the technique is in the phase of the idea. Web scraping can in fact be used to automatically explore websites or social media continuously, to extract data such as tweets or financial statements, or to monitor real-time events or financial statements. Web scraping can also be used to search for information as a base to build news stories. An example is the Reuters News Tracer[5], a tool that collects and analyzes tweets to

[4] https://www.browserstack.com/guide/python-selenium-to-run-web-automation-test
[5] https://www.reutersagency.com/en/reuters-community/reuters-news-tracer-filtering-through-the-noise-of-social-media/

capture important news and give Reuters' journalists a great advantage in gathering information and reporting.

- In the phase of research, web scraping can also be used to gather large datasets from public records, social media platforms, and other websites.
- In the production phase, scraping techniques can be used to automatically build content based on the data scraped. An interesting example of this is *Quakebot*[6], a software application developed by the Los Angeles Times that reviews earthquakes from the U.S. Geological Survey and automatically generates a report.
- Post-publication, web scraping can be used to constantly update web pages and parts of content, such as charts and other types of visual representations, both fixed and variable. It can also be used to update old articles with new information. An example of this is the Covid-19 tracker by The New York Times[7] [8]. In order to deal with a lot of data coming in at a very fast pace, the NYT digital team built a JavaScript web application (based on NodeJS) that was able to scrape data across all 50 US states. These data were then reviewed and made available to readers, citizens, and researchers[9] through immersive and interesting visualizations. Although scraping was an important part of the NYT Covid-19 tracker, some data had to still be reported by hand. The process is described in Figure 3.

[6] https://www.latimes.com/people/quakebot
[7] https://open.nytimes.com/tracking-covid-19-from-hundreds-of-sources-one-extracted-record-at-a-time-dd8cbd31f9b4
[8] https://www.nytimes.com/interactive/2023/us/covid-cases.html
[9] https://github.com/nytimes/covid-19-data

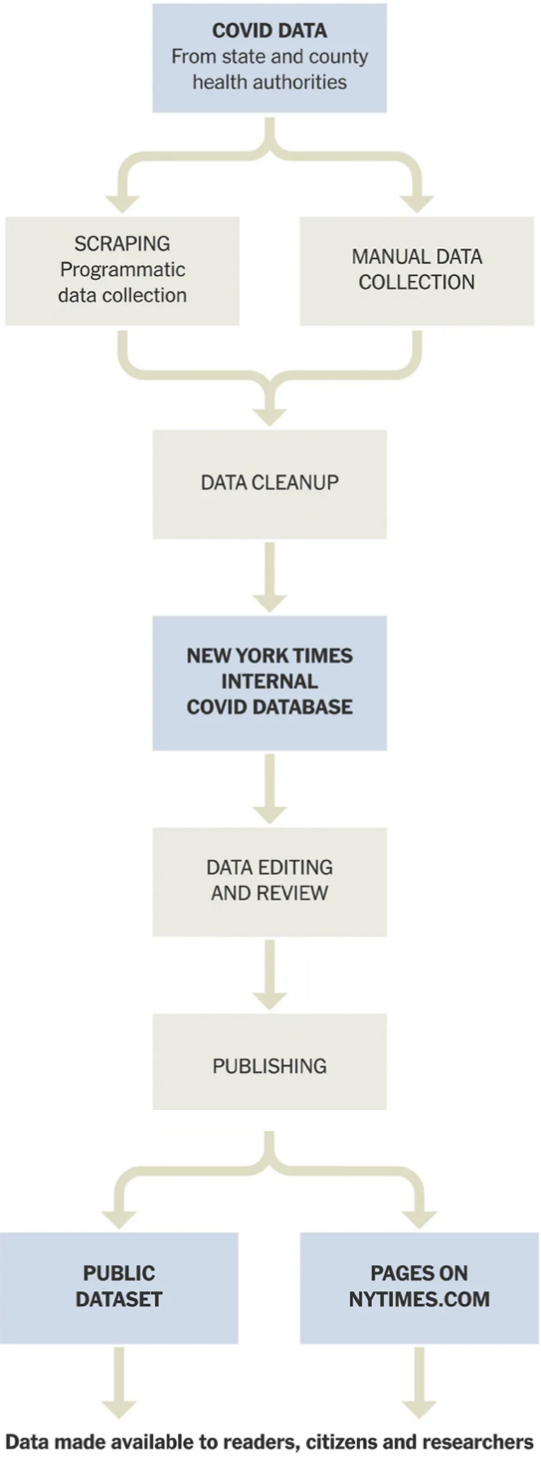


Figure 3: The functioning of the NYT Covid-19 Tracker. **Source**: The NYT Open Team, "*Tracking Covid-19 From Hundreds of Sources, One Extracted Record at a Time*": https://open.nytimes.com/tracking-covid-19-from-hundreds-of-sources-one-extracted-record-at-a-time-dd8cbd31f9b4, 17th June 2021.

The tracker, together with the whole coverage of the pandemic, was the reason why the NYT won the 2021 Pulitzer Prize for Public Service.

- Web scraping can also enhance fact-checking. Checkers can, in fact, use wrapping techniques to gather data and information from reliable sources to verify if a statement or a particular piece of news is either true or not.

2.3.1 Case Study: Reuters News Tracer

Nowadays, news professionals are tasked with the challenge of having to rely more and more on non-traditional sources of information. This is especially valid when we talk about social media, where everyone can, in real-time, publish updates on a particular event or situation. News events are now almost entirely broken in social media first, which makes them a crucial source in news discovery. In the later years, many tools have been produced to tackle this challenge, and one of them is the Reuters News Tracer, a powerful tool developed by Reuters in 2016 (Liu et al., 2016).

Reuters is a British news agency established in 1851 by Paul Julius Reuter, an Anglo-German journalist who, after moving to London, started a service of stock prices updating between the UK capital and Paris by using the newly built telegraphic cable under the 'La Manche' channel. Over the years, Reuters gained fame thanks to the fast pace with which breaking news was published. In particular, in its first century of existence, Reuters was the first news agency to report on important events, such as the outcome of the Battle of Solferino or the end of World War One. It was also the first European news agency to report on Abraham Lincoln's death. Reuters has maintained its competitive advantage through continue innovation and adaptation to technological breakthroughs, and AI made no exception.

The Tracer was developed with the primary goal of giving journalists a better tool than existing solutions to discover and verify news on social media, particularly on Twitter (now X). At the time, one of the major Twitter tools was the TweetDeck, whose role was to monitor live streams of tweets. At the same time, it had some advanced features that enabled the user to filter tweets by profiles and keywords. This tool, although useful, had its drawbacks:

- Firstly, news professionals had to screen a large volume of tweets before finding something relevant, even with all the filters.
- The second setback was delimited by human capacity, which is limited. Every journalist could only follow a finite set of events and news stories.
- Lastly, journalists still had to manually verify tweets published, which made the process even more tedious.

That's why the Reuters' newsroom decided to implement the News Tracer, an algorithm that was able to scrape tweets to detect news events and at the same time verify them. The Tracer needed to be designed with the goal of It to be fast in reporting news, but at the same time to be accurate, in order not to report on fake news and not newsworthy events. The algorithm should so have a low-latency pipeline, but at the same time it had to reduce noise as much as possible. The Tracer workflow can be seen in Figure 4.

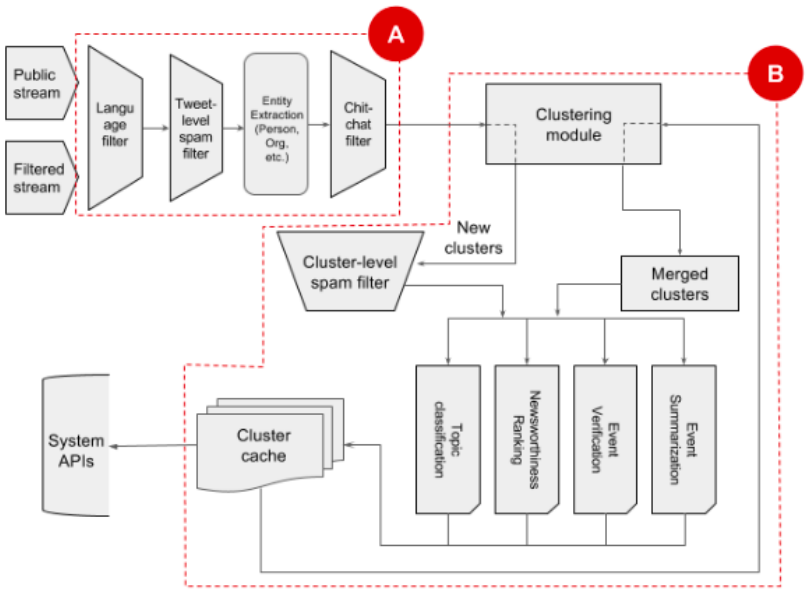


Figure 4: Machine learning system architecture for Tracer. **Source:** Research and Development Team @ Tomson Reuters, *Reuters Tracer: A Large Scale System of Detecting & Verifying Real-Time News Events from Twitter*, 2016.

It's possible to see how the model is divided into two parts, A and B:

- The A part is the tweet processing module, and so is the part in which tweets are acquired and processed. Here the tweets, after they are collected, pass a language filter in order to be processed (they have to be in English: the Tracer is not trained to understand more languages because it would have been too computationally expensive to train). After this process, the tweets are then controlled by a noise filter before being sent to clustering.
- The B part is the event detection process, where clusters are then merged and controlled. The events in them are then simplified and verified, before being sent in a cluster cache to be later sent in the System APIs.

The whole system scales through the use of durable and resilient Kafka message topics and the horizontal scalability of the Spark Streaming service.

- Kafka is a distributed messaging system which is commonly used for handling large volumes of data to allow different parts of an application to communicate by sending messages (data) through topics, named "channels"[10]. A characteristic of the Kafka system is its durability and resilience, meaning it can handle data persistently and maintain topic's availability even in adverse conditions, which ensures data won't be lost and, therefore, can be processed even after unfortunate events such as system failures, a characteristic that makes it suitable for an architecture, such as the Reuters Tracer, that needs to count on a strong architecture to operate on big real-time data feeds. Kafka also operates on a distributed system architecture, which enhances its tolerance and resilience and allows the Tracer to operate continuously and reliably 24/7.
- The Spark Streaming service is an extension of Apache Spark technology[11] that, together with the Kafka message topics system, enhances the Tracer's ability to quickly scale through the large quantity of data it must follow.

The integration of these two technologies supports Tracer's requirements to ingest, store, process, and analyze a large stream of Twitter data in real time and reliably.

[10] Kafka documentation: https://kafka.apache.org/documentation/
[11] Apache Spark documentation: https://spark.apache.org/docs/latest/

The process architecture is visible in Figure 5.

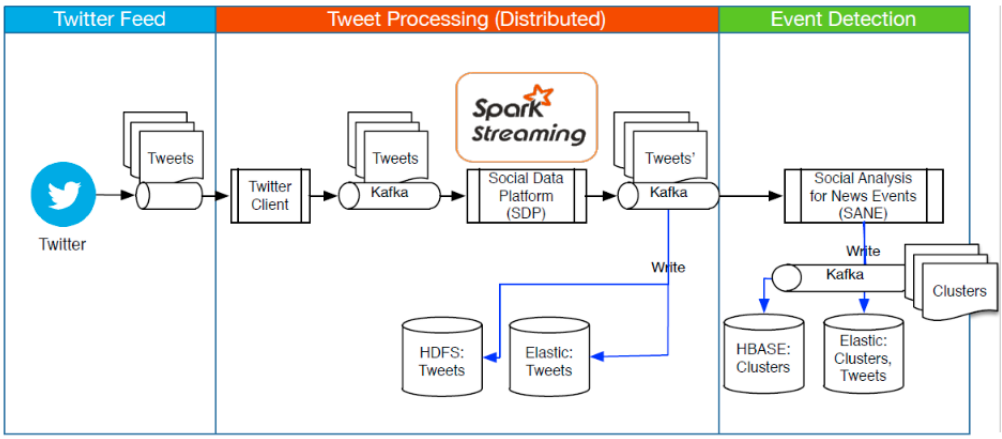


Figure 5: Data processing system architecture for Tracer. **Source:** Research and Development Team @ Tomson Reuters, Reuters Tracer: A Large Scale System of Detecting & Verifying Real-Time News Events from Twitter, 2016.

The Tracer's process can be divided into 4 major parts:

- After having collected tweets in real-time (more than 12 million per day) the Tracer preprocesses them before the analysis. This includes also, as said before, a language detection process to only include tweets in English. Because Twitter's language model didn't work well with messages with mixed languages, it was enhanced by concatenating it with Reuters' one.
- A critical step in the Tracer's pipeline was noise filtering: the author wrote that it was one of the most difficult parts of the whole project, together with the event clustering phase. In this context, noise was considered for all the tweets that didn't contain news. To tackle it, the team behind the Tracer adopted a hybrid approach of both rule-based filtering for specific noise types (spam, advertisements, everyday conversations, also called Chit-Chat) and machine learning algorithms for more general noise categories. The whole process was designed to have a high level of recall. The nose recognition process was divided between tweet-level and clustering-level filtering. The "tweet-level" phase followed three steps:
    - First, spam and advertisement were tackled with rules such as the number of daily tweets or the presence of particular usernames of firms or big corporations.

- Regarding conversations (or chit-chat), they were recognized with a classification algorithm trained on two online conversation corpora with 20,584 messages. A few heuristics, such as the presence of multiple first and second personal pronouns or emojis, were applied to enhance the model. If the tweets reached a certain threshold in the classification algorithm, they were automatically classified as chit-chat.
- As a last step, to further enhance the model, standard classification techniques are used to further separate noise tweets.

After the filtering, tweets are later distributed in clusters based on the event they were talking about. They need to go through cluster-level filtering before they are merged or emitted as event clusters. The algorithm that performs the task is similar to the tweet-level one, except the features are mapped to topic-level credibility features. The outline of the noise filtering algorithm, as written in the paper, is depicted in Figure 6.

Algorithm 1 Outline of our noise filtering algorithm

input: a stream of tweets $\mathbb{T} = \{t_1, ..., t_n\}$
output: a set of event clusters $\mathbb{C}_e = \{e_1, ..., e_m\}$

foreach tweet t in $\mathbb{T}$ **do**
  $s_l \leftarrow$ SPAM_TWEET(t)
  $a_l \leftarrow$ AD_TWEET(t)
  $c_l \leftarrow$ CHAT_TWEET(t)
  $n_l \leftarrow \max(s_l, a_l, c_l)$ ▷ maximize the noise level
  if $n_l \neq$ noise **then**
    $n_l \leftarrow$ NOISE_TWEET(T)
    if $n_l \neq$ noise **then**
      $\mathbb{T}_{\neg n} \leftarrow t$ ▷ obtain non-noise tweets
$\mathbb{C}_u =$ CLUSTERING($\mathbb{T}_{\neg n}$)
foreach unit cluster c in $\mathbb{C}_u$ **do**
  $n_l \leftarrow$ NOISE_CLUSTER(C)
  if $n_l \neq$ noise **then**
    $\mathbb{C}_e \leftarrow$ MERGING(c)

Figure 6: Outline of noise filtering algorithm for Tracer. **Source:** Research and Development Team @ Tomson Reuters, *Reuters Tracer*: *A Large Scale System of Detecting & Verifying Real-Time News Events from Twitter*, 2016.

- The next phase was event clustering, another critical phase. Unlike many other methods, the Tracer doesn't depend on hashtags because, in the authors' words, "if common hashtags

have emerged to refer to an event, it is very likely that this event is already trending and captured by news professionals." Instead, it relies on name entities (NE), identifiable proper nouns that represent real-world objects such as people, organizations, locations, or dates. Common TF-IDF (Term Frequency-Inverse Document Frequency) methods, and so statistical methodologies that are used to evaluate how important a word is to a document in a collection of corpus, weren't used, and the clustering process was divided into two stages: a unique clustering generation and a cluster merging one, to achieve real-time event detection, not update the UI each time a cluster has a new tweet, and to avoid excessive cluster veracity and continuous newsworthiness re-computation.

- During the unique clustering generation phase, every verified tweet was processed following three criteria: retweets, the similarity of links (tweets with the same link usually refer to the same event, and so belong to the same cluster), and event features like NE, nouns, verbs, and hashtags, which are then called 'markables' by the authors. A unit cluster contains an N number of tweets. This is the maximum value to form unit clusters, and it is set to 3 by the authors, whose goal is to form event clusters as soon as an event takes place. Also, it helps the Tracer to be more efficient by ensuring that the initial unit clusters do not expand indefinitely. Given a specific tweet, its similarity to a unit cluster is computed as:

$$S_i = a \cdot N_e + b \cdot N_n + c \cdot N_v + d \cdot N_h.$$

Where N_e, N_n, N_v, N_h represent the markables (respectively NEs, nouns, verbs, and hashtags) and a, b, c, d are parameters, whose values are learned after going through an event corpus of more than 500.000 tweets. Every cluster formed needs to go through cluster-level filtering, after which they go to the merging phase. Unit clusters would also be removed from the cache if they are outdated.

The unit clustering algorithm is depicted below in Figure 7.

Algorithm 2 Outline of unit cluster generation algorithm

input: a stream of tweets $\mathbb{T} = \{t_1, ..., t_n\}$
output: a set of unit clusters

```
1: foreach tweet t in T do
2:     if the id or source id of t is in the retweet map then
3:         place t in the corresponding cluster e
4:     if a link in t is in the link map then
5:         place t in the corresponding cluster e
6:     else
7:         foreach unit cluster e_i in cache do
8:             S_i = similarity between t and e_i
9:             if max(S_i) > similarity threshold then
10:                place t in cluster e_i with max(S_i)
11:    if number of tweets in e = N then
12:        remove e from cache
13:        send cluster e to cluster-level filter
14:        update maps for retweets, links and markables
15:    foreach e_i in cache do
16:        check inactivity time and do purge accordingly
```

Figure 7: Outline of unit cluster generation algorithm for Tracer. **Source:** Research and Development Team @ Tomson Reuters, *Reuters Tracer: A Large Scale System of Detecting & Verifying Real-Time News Events from Twitter*, 2016.

- Then, the unit clusters will be merged based on their similarity. In the Tracer's UI are represented all the clusters that have merged in the last Y hours (Y is set at 24 by default). All the merged clusters in the last Y hours are present in the cache. Here, the unit clusters are first merged with other clusters if similarities are found. Otherwise, it will form a cluster on its own. If newly formed clusters are not updated after X minutes with new tweets incoming or other clusters, they will be purged. All the newly formed clusters will affect both the cache and the Tracer's UI.

- Both the ML models for noise filtering and the event clustering were trained and validated on a Twitter event corpus of 51 million tweets, of which 3.034 of them belonged to 27 events. This base dataset was created by both crawling of event tweets and a sample of other tweets was adopted to create the benchmark dataset. The models were further validated on real-time streaming tweets in order to see how the model was working on real-world data.

- Regarding the noise filtering evaluation, the authors approached it by labeling as “positive” and “negative” classes respectively the noisy tweets and the non-noisy tweets. A cluster-level labeling effort was undertaken, mapping generated unit clusters to the 27 events to classify clusters as either noise or event-related. Out of 1.089-unit clusters, 224 were marked as event clusters. The models were selected based on high recall and low false negative rates. The best models at both tweet and cluster levels showed that a large fraction of noise could be removed without significantly impacting useful information. To assess the models' performance in a live environment, pilot evaluations were conducted with a sample of 4,000 tweets and 2,000 clusters annotated as "noise" or "event." These samples were reviewed by annotators with an inter-rater agreement kappa above 0.7, indicating strong agreement. Noise filtering’s evaluation can be seen below in Figures 8 and 9.

	Model	Prec.	Rec.	FP
tweet	Rule	0.999	0.332	0.0043
tweet	Rule+CS	0.999	0.557	0.0078
cluster	CS	0.991	0.627	0.0223

Figure 8: Performance of noise filtering on benchmark data. **Source:** Research and Development Team @ Tomson Reuters, *Reuters Tracer: A Large Scale System of Detecting & Verifying Real-Time News Events from Twitter*, 2016.

	Class	Sample	Kappa	Accuracy
tweet	Event	2000	0.81	79.4%
tweet	Noise	2000	0.85	67.9%
cluster	Event	1000	0.76	72.1%
cluster	Noise	1000	0.68	84.2%

Figure 9: Performance of noise filtering on streaming data. **Source:** Research and Development Team @ Tomson Reuters, *Reuters Tracer: A Large Scale System of Detecting & Verifying Real-Time News Events from Twitter*, 2016.

- Tracer's event clustering approach was benchmarked against the current state-of-the-art using Locality-Sensitive Hashing (LSH) for reducing candidate set sizes and enabling efficient clustering. The evaluation focused on the number of clusters generated and the recall of true events detected. LSH was found to produce suboptimal recall compared to Tracer's algorithm under various parameter settings. Tracer's event clustering algorithm outperformed LSH in terms of recall, indicating a superior ability to detect true events without missing significant occurrences. The recall of Tracer's clustering algorithm was also evaluated on real-time streaming data, further affirming its effectiveness in a production environment. The results indicated that Tracer's algorithm could maintain high recall in detecting true events from the Twitter stream, making it preferable over the LSH method. Specific recall figures and comparisons are provided in Figure 10, which shows the comparative analysis of Tracer's algorithm against LSH under different parameter settings.
-

Model	# of Cluster	Event Recall
LSH - (10, 60)	102	15/27
LSH - (13, 70)	94	16/27
LSH - (20, 80)	90	13/27
Our	441	21/27

Figure 10: Performance of event clustering using LSH with various (k, h) & the Tracer on benchmark data. **Source:** Research and Development Team @ Tomson Reuters, *Reuters Tracer: A Large Scale System of Detecting & Verifying Real-Time News Events from Twitter*, 2016.

- Lastly, before going into the cache, the clusters go into one last step of verification, summarization, and topic classification. At the end, a score is assigned to every merged cluster based on its newsworthiness.

In general, the use of the Tracer, which was made available to Reuters' journalists in 2015, sped up the work of the whole newsroom, making them gain an advantage in reporting news. In particular, on a set of 31 events between December 2015 and May 2016, the Tracer was faster in reporting than normal Reuters alerts on 27 of them. 3 out of 4 events in which the Tracer "lost" were political ones,

but just because Reuters journalists monitored them closely. The Tracer was faster also in 24 events against other global media, leading to a 47-minute lead on average. In particular, during the aftermath of an earthquake in Ecuador which caused 77 deaths, the Tracer gave Reuters' journalists 18 minutes of advantage on any other news outlet to gather more information. It also gave the organization's journal an eight-minute advantage over other news media over the Brussels bombing in March 2016 and a 15-minute lead in reporting on the Chelsea Bombing in October of the same year. As Reginald Chua, Executive Editor for Editorial Operations, Data & Innovation at Reuters quoted, "…Reuters News Tracer has beaten global news outlets in breaking over 50 major news stories. This has given our Reuters journalists anywhere from an 8- to 60-minute head start."[12]

2.4 Transformer Architecture based models

2.4.1 Attention is all you need

Since its foundation, a primary objective of Artificial Intelligence (AI) within the broader field of Computer Science has been to enable machines to understand and analyze human language, thereby allowing them to perform actions in response. This ambition led to the creation of Natural Language Processing (NLP), a branch of AI focused on enabling interactions between machines and humans by making computers comprehend human communications. Over the years, NLP has become more effective and more efficient, especially with the arrival of Deep Learning. In 2017, at the time the Transformer architecture was first introduced, the state-of-the-art regarding NLP was rapidly evolving, with Sequence-to-Sequence (Seq2Seq) models, based on Recurrent Neural Networks (RNNs), Long Short-Term Memory Networks (LSTMs) or Gated Recurrent Units (GRUs) being the most used for language modeling and machine translation. These models rely on an encoder-decoder structure, in which the encoder processes the input sequence and compresses the information into a context vector (or several vectors in more advanced models) and the decoder generates the output sequence. This solution, although seems pretty straightforward, can encounter several challenges when facing long sequences due to the sequential structure of the model, which processes data sequentially, one step at a time. This makes it very difficult to train, and computationally very expensive. A solution for this problem were the Convolutional Seq2Seq

[12] Source of the declaration and Tracer's real event performances: Thomson Reuters Fact Book 2017 (https://archive.annual-report.thomsonreuters.com/2016/downloads/thomson-reuters-fact-book-2017.pdf), page 74

(ConvSeq2Seq) models (Gehring et al., 2017), which were capable to process all the input sequence simultaneously and, at the same time, not suffer from vanishing or exploding gradients with the same extent as RNN, making them more stable during training. However, they were not able to reduce the computational complexity for long sequences, as each layer complexity is proportional to the sequence length.

In this context, the Transformer architecture was introduced as a competitor for the established models. In particular, as we will see in the architecture's analysis, it differentiates itself from the once-established Seq2Seq and ConvSeq2Seq models by its use of the attention mechanism. This concept was first made public in the context of the Neural Machine Transformation model's introduction in 2014 (Bahdanau et al., 2014). In the paper, attention was used to point the focus on different parts of an input sentence while translating it into another language, overcoming in this way the rigidity of the common encoder-decoder approach. While attention mechanism was first used in translating a text from one language to another, it sparked a wide interest in exploring and expanding the use of attention mechanism in other NLP tasks. Although it was used in successive Seq2Seq and ConvSeq2Seq models, the Transformer architecture was the first proper model to utilize a novel form of attention called 'self-attention', which eliminated the need for relying on a recurrent process for sentences computation.

The architecture of the first Transformer (Vaswani et al., 2017) follows, as the most competitive neural sequence transduction models at the time, an encoder-decoder structure. In particular:

- The encoder consists of a stack of N = 6 identical layers, with every layer having two sublayers, the first being a multi-head self-attention mechanism, while the second is a fully connected feed-forward network. Around each sublayer a residual connection is employed, followed by a layer normalization. In order to facilitate these connections, all the sublayers and the embedding layers produce outputs of dimension d = 512.

  The output of each sublayer is:

$$\mathrm{LayerNorm}\big(x + \mathrm{Sublayer}(x)\big)$$

Where x is the input of the sublayer and $\mathrm{Sublayer}(x)$ is the function implemented by the sublayer itself. The input embeddings are combined with positional encodings to provide information about the position of tokens in the sequence, as the model itself does not inherently capture sequential information.

- Like the encoder, also the decoder is composed of a stack of six identical layers but, differently from it, every layer has three sublayers, with the first being a multi-head self-attention mechanism, the second being a multi-head attention mechanism and the third one is a fully connected feed-forward network, similar to the one we talked about before. Around each sublayer a residual connection is employed, followed by a layer normalization. Differently from the encoder, the self-attention layer is modified in order to prevent positions from attending to subsequent positions. This masking, together with output embeddings positions are offset by one, ensures that the predictions for position i can depend only on the known outputs at positions less than i.

The central component of the Transformer, as said before, is the use of the attention mechanism. An attention function, quoted by Vaswani's paper, "can be described as mapping a query and a set of key-value pairs to an output, where the query, keys, values, and output are all vectors." The output is computed with the values, where each of these values is assigned. Specifically, the paper introduces a specific form of attention called "Scaled Dot-Product Attention", and then extends it into "Multi-Head Attention". Both of these two mechanisms can be seen below, in Figure 11.

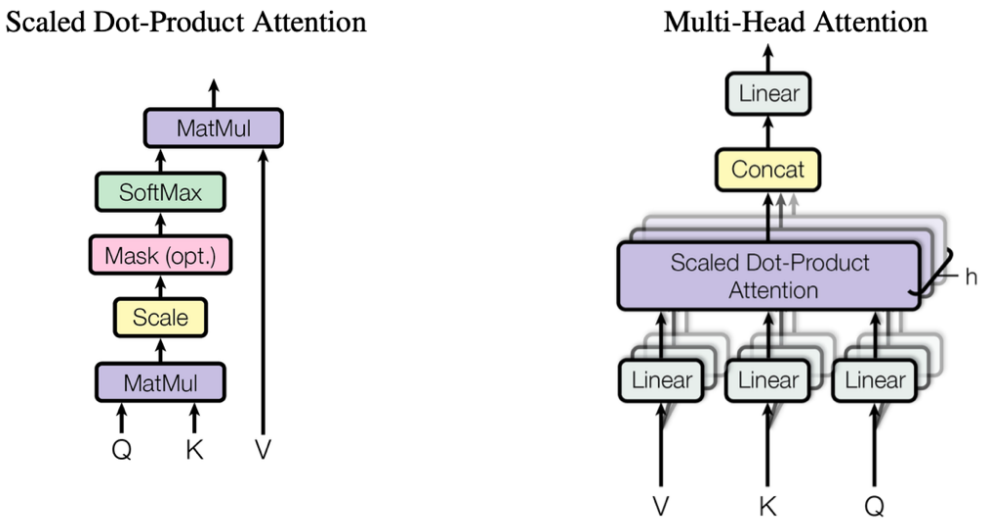


Figure 11: Scaled Dot-Product Attention (left) and Multi-Head Attention (right). **Source:** Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, I. Polosukhin, *Attention is all you need*, arXiv preprint, 2017.

Each of the layers contains a fully connected feed-forward network, which consists of two linear transformations with a non-linear activation, which is a ReLU, between them.

The first Transformer was trained on a standard WMT 2014 English-German dataset, that contains about 4.5 million sentence pairs, and a WMT 2014 English-French dataset containing 36 million sentences. Every sentence was converted into tokens using Byte-Pair Encoding (BPE), a form of subword tokenization that helps to effectively deal with rare words in languages. Specifically, the target vocabulary for English-German was made of about 37000 tokens, while for English French it consisted of 32000 word-piece vocabulary. Every sentence pair was batched together by approximate sequence length, In order to ensure uniformity with each batch.
The model obtained outstanding results, which surpassed what were the current state-of-the-art models.

- For English to German translation, the base model of the Transformer obtained a BLEU score (Bilingual Evaluation Understudy, a metric used to evaluate the quality of a machine-translated text)[14]. of 27.3, a big improvement from the previous benchmark which was set

[14] https://medium.com/@priyankads/evaluation-metrics-in-natural-language-processing-bleu-dc3cfa8faaa5

by a ConvSeq2Seq Ensemble model at 26.4. The big model improved the score to 28.4, setting a new benchmark for translation models.

- For English to French translation, the Transformer again surpassed the ConvSeq2Seq Ensemble model by achieving a BLEU score of 41.8, although the base model was far less accurate than all the other models with which was compared.
- One of the biggest achievements of the Transformer model was its efficiency. In fact, it was able to obtain better results than other models by being less computationally expensive, as visible in Figure 12. The authors also pointed out the fact that the training was significantly fast on the hardware (eight NVIDIA P100 GPUs): it took only 12 hours to train the base model, and up to 3.5 days for the bigger one, thanks to the ability of the model of parallelize the processing of data.

Model	BLEU		Training Cost (FLOPs)	
	EN-DE	EN-FR	EN-DE	EN-FR
ByteNet [18]	23.75			
Deep-Att + PosUnk [39]		39.2		$1.0 \cdot 10^{20}$
GNMT + RL [38]	24.6	39.92	$2.3 \cdot 10^{19}$	$1.4 \cdot 10^{20}$
ConvS2S [9]	25.16	40.46	$9.6 \cdot 10^{18}$	$1.5 \cdot 10^{20}$
MoE [32]	26.03	40.56	$2.0 \cdot 10^{19}$	$1.2 \cdot 10^{20}$
Deep-Att + PosUnk Ensemble [39]		40.4		$8.0 \cdot 10^{20}$
GNMT + RL Ensemble [38]	26.30	41.16	$1.8 \cdot 10^{20}$	$1.1 \cdot 10^{21}$
ConvS2S Ensemble [9]	26.36	**41.29**	$7.7 \cdot 10^{19}$	$1.2 \cdot 10^{21}$
Transformer (base model)	27.3	38.1	**$3.3 \cdot 10^{18}$**	
Transformer (big)	**28.4**	**41.8**	$2.3 \cdot 10^{19}$	

Figure 12: The Transformer achieves better BLEU scores than previous state-of-the-art models on the English-to-German and English-to-French newstest2014 tests at a fraction of the training cost. **Source:** Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, I. Polosukhin, *Attention is all you need*, arXiv preprint, 2017.

2.4.2 Transformer's Improvements and Adaptions: BERT & GPT

Although Transformer was used in its first model for machine translation, it was clear that its range of use could be much bigger. In 2018, research by OpenAI and Google AI department created two

improvements of Transformer: Generative Pre-traned Transformer (GPT) and Bidirectional Encoder Representations from Transformers (BERT).

- GPT was announced by OpenAI researchers (Radford et al., 2018) in June 2018, and it represented a significant advancement in the field of NLP, specifically for generating content and coherent text by starting from a given prompt. GPT can so be considered an advancement and a deepening of Transformer's field of use. GPT's architecture uses only a decoder stack of modified Transformer decoder blocks. In particular, the decoder followed a structure with 12 layers, with each decoder layer using a masked self-attention mechanism, where the prediction for a given token can only consider earlier tokens in the sequence. Also, GPT employs, like Transformer, Multi-Head Self Attention in its decoder layers with a ReLU activation between them and, as the Transformer, includes Position-Wise Feed-Forward Networks of 3072 dimensional inner states.

- Despite being a great advancement in NLP and being a pioneer in using a Transformer based architecture for NLP purposes, GPT was, like its predecessors, a unidirectional model. These types of models were either trained to predict the next or the previous word in a sequence, but not both simultaneously. During the same year, in October 2018, four people from Google AI (Devlin et al., 2018) developed BERT (Bidirectional Encoder Representations from Transformers), whose novelty was in both its "deep" bidirectional training and in its training approach, the Masked Language Model, which randomly masks some percentage of the input tokens, with the objective of predict the masked word based on its context.

    - BERT's architecture consists only of multiple layers based on Transformer's encoder. The choice allows BERT to be fit for tasks that require a full understanding of the context of sentences or paragraphs, which can be optimally achieved by an encoder. Moreover, this allows both left and right directionality, which is the base of BERT's model. The structure in the base model, 12 layers of Transformer blocks, 768 hidden units, and 12 attention heads (ca. 110 million parameters), is the same as GPT, a deliberate choice for comparison reasons, while the large model has 24 layers, 16 attention heads, and 1024 hidden units, totaling about 340 million parameters. BERT is trained using an Adam optimizer with specific learning rate schedules and

hyperparameters. For example, a learning rate of 1e-4, L2 weight decay of 0.01, and learning rate warm-up over the first 10,000 steps.

- The input representation in BERT, in order to be flexible in handling a variety of NLP tasks, is a continuous text stream that includes both singles and pairs of sentences. The input uses WordPiece embeddings, a 30,000 token vocabulary, in order to tokenize the input text into subword units and convert them into embeddings. To handle pairs of sentences, BERT uses segment embeddings to differentiate between two sentences and help the model understand which sentence a particular token belongs to (for example, each token of the first sentence will be assigned to segment embedding A, while a token from the second sentence will be assigned to segment embedding B). Another difference between BERT and Transformer is that Google AI's model uses positional embedding in order to provide information to the model over the position of each token in the sequence. This is important for BERT's bidirectionality, a feature that normal Transformer doesn't have. In the end, the final input representation, which is then put into the encoder, is the combined sum of token, segment, and positional embedding. This is clear by looking at Figure 13 below.

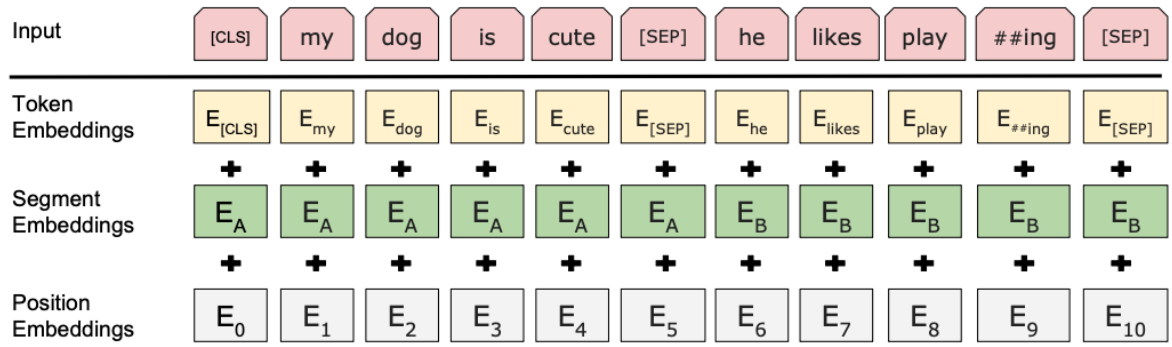


Figure 13: BERT input representation. **Source:** Google AI Language Team (Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova), *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,* arXiv preprint, October 2018.

- The output representation, although varies depending on the task the model is fine-tuned for, is generally represented by a vector representation for each input token that passes through one or more additional layers, depending on the specific task.

The results for BERT were way better than any other state-of-the-art NLP model at the time. Specifically, BERT was evaluated with the GLUE score, and outperformed OpenAI's GPT across all tasks by an average of 4.5% with its base model and 7% with its larger model. BERT's comparison with other models can easily be seen in Figure 14 below. It's important to recall that BERT's base model has the same architecture of OpenAI's GPT and outperformed it in several NLP tasks.

System	MNLI-(m/mm) 392k	QQP 363k	QNLI 108k	SST-2 67k	CoLA 8.5k	STS-B 5.7k	MRPC 3.5k	RTE 2.5k	**Average** -
Pre-OpenAI SOTA	80.6/80.1	66.1	82.3	93.2	35.0	81.0	86.0	61.7	74.0
BiLSTM+ELMo+Attn	76.4/76.1	64.8	79.8	90.4	36.0	73.3	84.9	56.8	71.0
OpenAI GPT	82.1/81.4	70.3	87.4	91.3	45.4	80.0	82.3	56.0	75.1
$BERT_{BASE}$	84.6/83.4	71.2	90.5	93.5	52.1	85.8	88.9	66.4	79.6
$BERT_{LARGE}$	**86.7/85.9**	**72.1**	**92.7**	**94.9**	**60.5**	**86.5**	**89.3**	**70.1**	**82.1**

Figure 14: BERT scores compared to other models. **Source:** Google AI Language Team (Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova), *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,* arXiv preprint, October 2018.

Both Transformer-based models we talked about represent a great advancement in the field of Natural Language Processing, and both of them have had strong advancements over the years. We previously mentioned GPT-2, GPT-3, and GPT-4 models, enhancement of the first GPT model which has a much larger and robust architecture, and therefore more parameters and a boost in results[15] [16] [17]. BERT too has had improvements, such as DistillBERT[18], that have enhanced its performance and made it less computationally expensive. Both GPT and BERT have their advantages and disadvantages.

- Due to its sequential output generation, GPT excels in tasks that require coherent text generation, such as content creation or article creation. On the other hand, it does not understand the context of a sentence, although GPT-4 has improved the capabilities of the model in this type of task, making it able to have better performances on understanding

[15] GPT-2 paper summarization: https://openai.com/research/better-language-models
[16] GPT-3 paper: https://arxiv.org/abs/2005.14165
[17] GPT-4 paper summarization: https://openai.com/research/gpt-4
[18] https://arxiv.org/abs/1910.01108

longer prompts and documents. Speaking about GPT-4, it has also improved the versatility of the model. GPT, like BERT, is very expensive both computationally and economically.

- Bert is way better than GPT-4 in understanding the context of sentences, thanks to its bidirectional nature, and it's also easier to fine-tune for other tasks way more effectively. This makes it way better at tasks that require deep understanding and reasoning over texts, such as language inference, sentiment analysis, and document summarization. However, it's less effective than GPT in creative text generation. Although some successive models, like RoBERTa and ALBERT, have improved its efficiency, it's very important to understand that BERT is still very resource intensive. Also, deploying BERT and its subsequent models still requires a non-trivial setup, especially when handling real-time applications, due to their computational and memory requirements.

2.4.3 Transformer and Transformer-based models' application in Journalism

There are many uses of Transformer and the models based on its architecture can be used in journalism.

- The most obvious one is, of course, content generation. In recent years GPT's most famous commercial software, ChatGPT by OpenAI, has demonstrated how simple is to create a coherent text from a given prompt. Many repetitive types of news, such as sport results or financial reporting, can be automated with the use of GPT. The Associated Press (AP) uses GPT to automatically deliver news reports on repetitive types of news. This allowed the AP to significantly improve the number of companies covered in their financial reporting articles, making it 10 times larger[19]. Since we are still at a point in which a human eye can, in most cases, understand if a content is generated by an AI or not, single journalists may use it to get ideas on how to start a written piece: Beatrice Petrella, one of the journalists I've interviewed for this thesis' purposes, uses ChatGPT to automatically produce a draft from which taking ideas on how to start an article.

[19] https://www.ap.org/solutions/artificial-intelligence/

- Algorithms like BERT can be very useful in extracting specific information from large documents. This can be used by journalists to quickly scan very large datasets and official documents in order to gather information in a very short amount of time.
- BERT it's also very useful and effective in summarizing large texts such as articles, scientific papers, or reports, allowing journalists to quickly understand the context of these types of texts. BERT was probably by the journalists of the International Consortium of Investigative Journalists (ICIJ) in order to quickly scan leaked documents and datasets from offshore companies during one of their most famous reportages, *Pandora Papers*[20]. These summarization capabilities can assist journalists also in tasks such as interview preparation, by helping the journalists understand the background of the person they are going to interview. If we push forward, we can even think about these algorithms being applied to speech-to-text ones, analyzing answers in real-time and even suggesting to the journalist what question ask to the interviewed person.
- Transformer can be used also in translating languages, as it was its first proper use in the 'Attention is all you need' paper.
- Transformer can even be applied in computer vision. The Vision Transformer (ViT) by Google Brain[21] uses Transformer to effectively be applied to tasks such as object detection in images and image classification. How it works is basically dividing the images into fixed-size patches which are then flattened and linearly imbedded like tokens. Then, positional encoding is applied to each patch in order for it to retain its positional context before processing them into Transformer encoder layers. After this, the images are classified thanks to a classification head (typically a simple linear layer) and then produce the final output. The use of this type of algorithms can be very important for investigative journalism. La Naciòn, one of the main newspapers of Argentina, used computer vision to analyze ballots and verify the regularity of the 2021 elections[22].
- BERT is particularly useful to journalists that need to conduct sentiment analysis, in order to understand better what stories might be interesting to his/her audience.
- Transformer, and in particular GPT, are very useful to write different types of the same articles, some more complex and some more simple. This might seem useless at first glance

[20] https://www.icij.org/investigations/pandora-papers/global-investigation-tax-havens-offshore/
[21] ArXiv link to the paper: https://arxiv.org/abs/2010.11929
[22] https://www.lanacion.com.ar/sociedad/telegramas-electorales-bajo-la-lupa-nid22102021/

but may be very important to better disseminate an article to different kinds of audience, and therefore reach a higher audience.

- Lastly, Transformer can enhance fact-checking in a similar way of web scraping, by processing large amounts of data from verified sources to either confirm or deny the accuracy of a statement or of an information. Some models can be trained to detect patterns typical of fake news and disinformation, therefore providing tools to both journalists and readers to verify the credibility of information.

CHAPTER THREE

EVOLUTION OF JOURN(AI)SM AND FUTURE TRENDS

3.1 Technology Adoption in Journalism

3.1.1 Internet and the Advent of Digitization

The end of the 20th century marked an important time in human progress. The 90s brought a new type of world with them, a more open and "little" one. Two events were the major causes of this new world's birth: the fall of the Berlin Wall and the consequential end of the Cold War, and the birth of the World Wide Web. Although these two events seem to be quite detached from each other, the fact that they happened together marked a significant disruption in the distances and the openness of the world. The Internet (as the WWW is most known) would "remain in an incubator" for most of the 20th century's last decade, but many people and organizations, from the most to the least famous ones, started to experiment by creating simple websites. Newspapers and news outlets also took the decade's opportunity to kick off their journey but, because most of them were large organizations, they had time and money to wait and study how to better use the technology. The BBC, one of the most technologically enthusiastic and advanced news outlets, the first in the world to transmit a live event on television (King George VI's coronation, 12th May 1937) and the first to broadcast a live event on color television (the Wimbledon Tennis Championships in 1967) opened its website only in 1997, and although it would look unreadable and ugly today, as can be seen in Figure 15, at the time it was a nice piece of advancement.

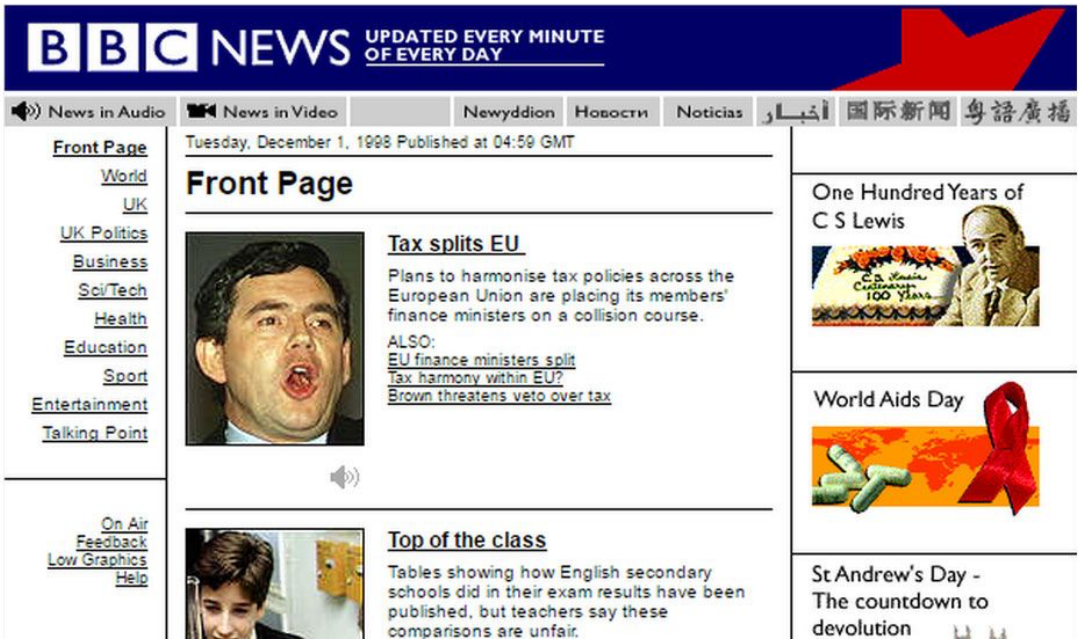


Figure 15: BBC website on December 1st, 1998. **Source:** *How the BBC News website has changed over the past 20 years,* an article from the BBC (British Broadcasting Corporation) website, 9th November 2017.

It is important to understand that few people had a computer during the end of the 20th century, and many of those who possessed one did not have access to the internet. The BBC's website only reached 1 billion monthly views on July 7th, 2005, seven years after its opening, during the London bombings[23].

In general, the advent of the World Wide Web influenced the journalism world in many ways:

- It democratized information dissemination and transformed traditional news cycles. News websites allowed continuous news updates and significantly faster dissemination of breaking news compared to traditional print cycles. This instant access shifted the attention of news consumers on real-time news. This extreme democratization of news also brought challenges with itself, linked to concerns about the verification of news.
- Many outlets also started their tech journey by scanning their newspaper and making it available in a PDF format.
- Also, digitization enabled newspapers and broadcasters to digitize their archives, making vast amounts of information more accessible and searchable.

[23] https://www.bbc.com/news/uk-41890165

- Finally, it gave the news networks in successive years an urge to change their business model to a more digital one, in order to contrast the contraction on the revenues from newspapers. This included starting to rely less on newspapers' sales and more on online advertisements.

3.1.2 The Arrival of Automation, RSS Feeds and CMS in the New Millennium

The passage from a slow news cycle (newspapers in the morning and TV news in the evening) to a continuous one brought, consequently, not only the need for news professionals to produce more articles and content but at the same time to make it more interactive and informative with the help, for example, of visualization tools. The growth of digital content and, at the end of the first decade of the new millennium, social media also made real-time data availability and communication crucial in the industry: these two platforms became at the same time sources of data and news, and channels where to publish instant news. In this context, journalists found suddenly themselves in the position of producing a way larger number of articles and new types of content, such as live feeds on important events like sport matches, elections, or, unfortunately sometimes, suicide attacks, and they had to start using real-time data feeds to track developing stories.

Therefore, automation became a must-do for the news outlets that digitalized during the 90s. At the start of the millennium, AI in journalism was largely experimental, and the first examples of application in the industry were in the automation of the most tedious and routine tasks in the newsroom. Earliest examples of it were simple algorithms designed to assist in sorting and categorizing large volumes of data. While technology advanced, some programs were also able to quickly scan and analyze data to generate basic reports on, for example, sport results or stock market fluctuations.

This decade also accelerated the cycle of news. One of the first accelerators was the introduction of RSS feeds in journalism, introduced by Netscape in 1999[24]. RSS (which stands for Really Simple Syndication) feeds were described in 2004 as "a simple XML syntax for describing a channel or feed of recent additions to a website" (Wusteman, 2004). It is, essentially, a way for websites to summarise their webpage with brief descriptions and link to news articles. Users can subscribe to a website's RSS feed by right clicking on the dedicated icon of the website and pasting the link they

[24] https://www.rssboard.org/rss-0-9-0

obtained in their RSS readers. In this way, users have the possibility to personalize their feeds and save time by directly accessing articles and news without going directly on the network's website. In particular, RSS feeds were particularly useful to launch breaking news or to link articles and deep analysis on news stories.

This increased automation freed up journalists from the most repetitive tasks but, eyebrows apart, it also raised concerns about the standardization and reduction of depth in reporting, which could result in errors by the machine, and led to a strong debate on the quality of reliability of automated news and increased the importance of fact-checking and analysis. The increasing availability of data online allowed the establishment of the first steps of what later would be called big data journalism: journalists could now analyze datasets at scale, and therefore produce more accurate and compelling pieces and investigations.

Another important evolution in the decade was represented by Content Management Systems (CMS). CMSs are types of software platforms designed to help users create, manage, and modify content on a website without the need for specialized technical knowledge. An example of a modern CMS is WordPress, which enables its users to create websites without having to rely on HTML skills. These types of platforms allowed newsrooms to be faster in updating their content and more efficient in managing it. CMS evolution during the decade brought these platforms to slowly be able also to make more complex tasks, integrating tools for analytics and targeted content delivery.

3.1.3 AI and the Big Data Revolution of the 2010s

During the second decade of the 2000s, we saw an increase in the use of AI and big data in journalism due to the advancements of technology in these fields. In particular, there was a change in perception of AI in journalism. During the previous decade, AI tools and big data were seen almost exclusively as something to automate repetitive and boring tasks, such as categorizing articles or transcribing interviews although, as we said before, some pioneers started enhancing their reports with the use of analytics and data analysis. At the start of the 2010s journalists started using AI more seriously to quickly analyze big data at a scale that was impossible to reach before, making them able to discover patterns and connections although unnoticeable with manual research. The availability of softwares able to perform complex data analysis tasks, such as visualizations, meant

that even journalists without a technical background or programming skills could perform complex analyses on large datasets. These factors, together with the increasing availability of large public and private datasets, made it possible for journalists to deepen their reports, contents, and stories, making them more accurate, detailed, and informative, even by putting the same time and effort as before. This new tendency was especially important for investigative journalism, with the series of investigations by ICIJ (International Consortium of Investigative Journalism) named *Panama Papers* (2015)[25]*, Paradise Papers* (2017)[26]*,* and *Pandora Papers* (2021)[27] being a famous example.

The development of AI algorithms in the field of both NLG (Natural Language Generation) and NLP (Natural Language Processing) led to an enlargement in the use of these algorithms. In the 2000s, automation of content creation started to take off seriously, but entering into the new decade newsrooms began going a little bit more far from experimenting, since now algorithms were more precise. As the years passed, bigger and more important newsrooms started relying much more on automated content. To enhance its coverage of the 2016 Olympics, the Washington Post introduced *Heliograph*[28]*,* an in-house built tool for automated storytelling that would write automatic reports, game stories, and Twitter updates (the Washington Post created an account, @WPOlympicsbot, where *Heliograph* would publish its content) on matches outcome and medal leaderboard. After the Olympics, *Heliograph* was then used to monitor the 2016 and the 2020 US election race[29], allowing the Washington Post newsroom to cover more than 500 races on the 2016 election night, and to enhance local coverage[30].

The increase in the use of big data in the 2010s, not only in journalism, raised significant ethical challenges regarding privacy issues and data misinterpretation. Although not directly linked to the sector we are talking about, the Facebook-Cambridge Analytica scandal of 2018[31] showed the dangers related to privacy and how data, sometimes very sensible, can be collected by corporations even without the needed permissions.

---

[25] https://www.icij.org/investigations/panama-papers/
[26] https://www.icij.org/investigations/paradise-papers/
[27] https://www.icij.org/investigations/pandora-papers/
[28] https://www.washingtonpost.com/pr/wp/2016/08/05/the-washington-post-experiments-with-automated-storytelling-to-help-power-2016-rio-olympics-coverage/
[29] https://www.washingtonpost.com/pr/2020/10/13/washington-post-debut-ai-powered-audio-updates-2020-election-results/
[30] https://www.washingtonpost.com/pr/wp/2017/09/01/the-washington-post-leverages-heliograf-to-cover-high-school-football/
[31] https://www.vox.com/policy-and-politics/2018/3/23/17151916/facebook-cambridge-analytica-trump-diagram

3.1.4 A Pandemic Causes the Acceleration of AI Adoption in Journalism

In the 2020s, journalism saw an acceleration in the adoption of AI and the use of big data, with the cause of this being the COVID-19 pandemic. During the lockdown period newsrooms had the significant role of informing the population on the evolution of the virus and infection counts, all of this with journalists being confined into their houses. AI was leveraged by many news networks and newsrooms in various ways during this period. In the previous chapter, we discussed the NYT COVID Tracker, an interactive real-time visualization of the evolution of cases in the US. Some similar trackers were done by The Atlantic[32] and academic institutions such as the John Hopkins University[33]. At this stage, AI technologies have progressed from basic automation and content generation to more sophisticated applications. The release of tools, such as ChatGPT-3.5 and ChatGPT-4, and user interfaces made possible even for smaller newsrooms to use AI in their work.

This large availability of AI for content creation raised concerns recently about their possible use to spread and disseminate disinformation. These worries were enhanced by the recent release of new tools like image and video generators such as Midjourney and OpenAI's DALL-E and Sora, which could be used by malicious people to manipulate reality and spread fake news. The Russo-Ukrainian War is (unfortunately) a ground of proof of disinformation spreading, and it's not secret both sides tried to use AI techniques to their advantage[34] [35]. In this context, AI can and will be used by fact-checkers to improve the quality, quantity and speed of the process.

3.1.5 From Automated to Augmented to Iterative Journalism

In the first part of this paragraph, we saw the development of technology in the last thirty years, from the World Wide Web to Artificial Intelligence and Big Data, and how it was adopted in journalism. In parallel with the development of technology, journalism also developed itself around it. We have seen two phases of this process in the paragraph above:

[32] https://covidtracking.com
[33] https://coronavirus.jhu.edu/map.html
[34] https://www.npr.org/2022/03/16/1087062648/deepfake-video-zelenskyy-experts-war-manipulation-ukraine-russia
[35] https://www.snopes.com/fact-check/putin-deepfake-russian-surrender/

- The first phase is **news automation.** Particularly, when newsrooms started to experiment with AI the first thing, they started to automate repetitive types of news stories, such as sport results or reports on key findings in balance sheets. This allowed journalists to not waste their time in redundant activities and save it to investigate more deeply and produce better news stories. The first experiments involved the so called branch writing, a particular technique which consists in generating a certain word or sentence under particular conditions, most of the times using if-else statements.
- The second phase was **news augmentation.** In this phase the journalists can use tools that augment their own reporting. This includes the use of software for big data analysis, speech-to-text translation and visualization.
- Last phase is **iterative journalism,** in which newsrooms, with the integration of data analysis and algorithms in the process of reporting and by mixing editorial insights with audience feedbacks, share minimal details of stories to understand the needs of the readers to then adjust coverage to deepen only the most important ones.

3.2 AI in Newsrooms

3.2.1 A Switch in Journalism Workflow

The introduction of AI into newsrooms has completely disrupted many aspects of the work inside them. The journalist's workflow, as we saw in the previous chapter, is composed of six phases: idea, research, production, distribution, feedback and archiving. In the majority of reports, and in Francesco Marconi's book, *Newsmakers* (2020), the stages are three, without counting the phases of feedback and archiving:

- **Newsgathering.** This phase includes both the idea and the research, so it's the stage where journalists collect information about trends and topics.
- **Production i**s basically the stage in which the content is produced.
- **Distribution,** that is, as understandable, the phase in which the content or article is distributed through various platforms.

Before AI and big data, the journalist's work in these three phases was linear and rigid:

- The journalist would usually discover scoops and news by contacting the network of sources around him/her. The journalists would then try to take as much information as possible through the just-mentioned network by interviewing the people inside it. Another way of collecting information was with official printed documents. Still, processing documents and interviews was a work which would have been done manually, and so it would take a lot of time for the journalist to
- The production consisted of writing an article on the topic that included all the information collected by the journalist.
- The distribution was very rigid, too. Before digitalization, the only channel through which the journalist could share his/her work was the newspaper.

Today the advent of digitalization and artificial intelligence in journalism made this workflow more dynamic and flexible:

- Thanks to the World Wide Web, now the journalist has at his/her disposal terabytes of data in the form of datasets from which extract information and find relevant patterns. As we will see later, the journalist can easily use IoT devices too to leverage their ability of collecting data.
- Content production has seen a diversification in the type of content published. The birth of formats such as podcasts and video platforms have in fact given journalists a wider choice regarding the channel they can use to disseminate news.
- Together with the increase in content production formats, there's also been a rise in distribution platforms for journalists.

The journalist model has become so flexible that phases can overlap. Examples are dashboards being updated in real-time as soon as new information enters the system, with the NYT COVID Tracker being an example.

The increasing number of sources from which information can be extracted has profoundly impacted how journalists collect, analyze data for news stories, and how they distribute it:

- The *South Florida Sun Sentinel* conducted an investigation on speeding police officers using GPS devices in 2013. In particular, the journalists obtained a dataset from SunPass, Florida's toll system, which contained 1.1 million instances representing tool transactions for 3,900 South Florida police transponders, with features such as the date, location and time in which a car passed through a toll booth. To measure the distance between toll booths the *Sentinel's* journalists used a very accurate Garmin cyclo-computer and registered it by driving from toll system to toll system. They then put the recorded distance into an Excel spreadsheet, where they can easily calculate the speed having the time officers took to go from toll to toll. They caught nearly 800 police officers driving above the speed limit (90 mph) without reason, with most of them being off-duty or commuting either from or to work. They also discovered special treatment for the offenders. This resulted in 23 officers from three different departments being disciplined, the installation of speed monitoring systems on the police cars in one city and the exploration of GPS technologies to catch speeding cops in three other centers. The investigation won the Pulitzer Prize for Public Service in 2013[36].

[36] https://www.ire.org/how-the-sun-sentinel-reported-its-pulitzer-prize-winning-coverage-of-off-duty-cops/

- *Radiolab,* a US public radio program and podcast, developed a dashboard to predict the arrival of cicadas by leveraging temperature sensors[37].
- The *Berliner Morgenpost* created an interactive 3D dashboard to visualize places on Earth that will become uninhabitable in the year 2100 due to global warming. They collected data from the Intergovernmental Panel on Climate Change (IPCC) reports, scientific papers and models that provide scenarios for future climate conditions based on different levels of greenhouse gas emissions. Although there are many scenarios, the project assumes a temperature rising of 2.5-3°C by the end of 2100, with the major climate change-related factors considered in the visualization being excess heat, water stress, sea level rise and tropical cyclones. The *Morgenpost* also used current population data from national census or international institutions such as the UN to estimate the number of people that will be affected by temperature rising. Data are plotted on a virtual globe divided in hexagons with different heights to represent the number of inhabitants of that area that will be affected by climate change[38]. An image is represented below in Figure 16.

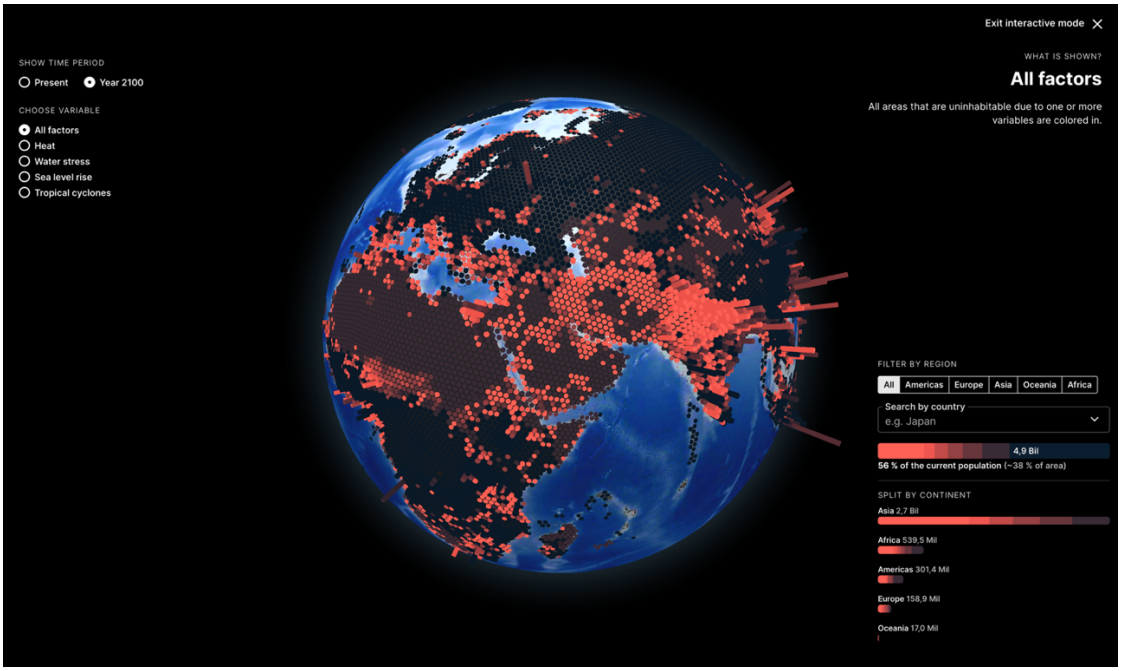


Figure 16: Berliner Morgenpost's dashboard. **Source:** *Mapping where the earth will become uninhabitable* (https://interaktiv.morgenpost.de/klimawandel-hitze-meeresspiegel-wassermangel-stuerme-unbewohnbar/en.html), The Berliner Morgenpost.

[37] https://project.wnyc.org/cicadas/
[38] https://interaktiv.morgenpost.de/klimawandel-hitze-meeresspiegel-wassermangel-stuerme-unbewohnbar/en.html

- As we previously wrote, the International Consortium of Investigative Journalists (ICIJ) used machine learning and AI tools in all their most famous investigations. When the ICIJ analyzed their first major leak of documents in the *Panama Papers* (2015) they used various tools to scrape 2.6 terabytes of information and documents. In particular, tools such as Apache Solr and Apache Tika for document processing and indexing, Tesseract, an optical character recognition (OCR) software, to transform scanned documents in searchable text, and Linkurious, to visualize complex networks and relationships uncovered in the data, were used to make sense of the vast amount of data[39].
- In 2016 *Quartz,* an information website, used computer vision to better understand the moods of the two presidential candidates of that year's US election. Specifically, the 'robot' was trained around the Emotions API by Microsoft, a set of images of faces labeled with a particular mood (happy, sad, angry etc.). The algorithm was then fed with one still frame from every five seconds of the debates between Donald Trump and Hillary Clinton, and would then label it with one between six emotions: happiness, surprise, sadness, contempt, disgust, anger. Although the algorithm found out that Clinton was happier than Trump during debates it also made mistakes, for example labeling raised eyebrows as surprise. *Quartz* stated that the methodology was new and prone to errors, but it also said that it was the best alternative "if you can't stand talking-head analysis"[40].

3.2.2 How to Implement AI in the Newsroom

Following Marconi's *Newsmakers,* a newsroom that implements AI needs first to understand its needs. In particular, it needs to decide whether 'why' and 'where' they want to introduce it.

- Regarding the 'why', a newsroom can either focus on automation or augmentation of their content. We've seen what automation and augmentation are in the first paragraph of this chapter, but news networks have to understand if their goal is to automate tasks that require repetitive work and doesn't produce a differentiated outcome or to improve a human task

[39] https://www.icij.org/investigations/panama-papers/data-tech-team-icij/
[40] https://www.facebook.com/quartznews/videos/what-happens-when-a-computer-watches-the-debates/1316110158422683/

with the help of a machine. In the first case they will opt for automation, in the second the newsroom will choose augmentation.

- There are multiple areas 'where' the newsroom can implement artificial intelligence. We've spoken about the automation of reports or processes such as news and source gathering, news selection etc. Particularly, Marconi states that content should be automated when "the news organization need to serve a big audience or an audience with very specific interest" but warns about automation exaggeration, that can devalue the overall value of the news organization. Processes, he continues, must be instead automated whenever possible, because "applying AI to certain activities can dramatically reduce human error and improve the overall uniformity of how content is labeled."

After having understood needs and purpose, a newsroom has to build, train the model, and then implement it into its workflow.

The outlet has usually two ways to implement the technology: they can either build the tools internally or they can choose to partner with external parties to build them. Usually, the choice between these two is based on four criteria:

- **Costs.** Building an in-house tool requires costs such as the hiring of new specialized personnel, the management of web servers, data storage, maintenance etc. Costs of partnering with third parties can include the payment of a monthly fee to use the service for undirected partnership. For bigger and more ambitious projects news networks can build a stronger partnership with external entities, which can therefore delve into the split of risks and costs. Littler newsrooms may not afford to build an internal tool, and so they would be obliged to rely on external companies.
- **Stability.** Building a tool internally means a better control on it, but it also means that system errors and maintenance must be addressed directly by the outlet. On the other hand, a partnership with an external entity will result in delegating some of the duties, including maintenance, management and errors' addressing.
- **Customization.** An in-house built tool can be specifically tailored for the needs of the newsroom, while external partner usually offers general solutions to the needs of the newsroom.

- **Privacy.** Journalists may be more comfortable in analyzing data or confidential documents in a tool that is directly controlled by the network than by external companies which may have access to the same data.

News outlets and networks look at all these criteria to decide whether to build tools internally or partner with external companies.

- We've spoken before about *Heliograph,* an NLG tool used for the first time to automate statical news, such as medal count or schedules, during the 2016 Olympics and built in house by the Washington Post.
- To automate the sport stories, the Associated Press decided to partner with tech company Automated Insights.

The last phase of the AI adoption inside the newsroom is the evaluation, a task that Marconi in his book develop by using five metrics:

- **Accuracy**
- **Speed**
- **Scale**
- **Integration**
- **Price Performance.**

3.2.3 Research Labs in the Newsroom

To accelerate the process of AI integration inside newsrooms, many important news outlets have started to develop research labs inside them.

Research labs, also called innovation labs, are R&D units within the newsroom specialized in various activities to foster innovation in journalistic practices (Cools et al., 2022). The first labs were established around a decade ago, with most of them representing a first attempt to foster innovation inside a traditional environment such as journalism. They were usually isolated from the newsrooms and were focusing more on the development of new technologies and methodologies inside it. The

COVID-19 pandemic and the technological advancements forced an evolution in innovation labs, which are now more blend into the newsrooms and aim to enhance how news are gathered, produced, and distributed by developing tools that are closely aligned with the newsroom's needs. This major integration resulted in a more Cools, in his paper, also define three types of innovation labs:

- **Static labs** are research labs with an established and fixed team dedicated to one or more projects. They usually follow a hierarchical structure, with a project leader at the top of it, and have to adhere to a clear set of methodologies or practices for developing tools, like strict deadlines. These labs, with their structured and hierarchical nature, allow for an improved efficiency.
- **Dynamic labs,** differently from static labs, are teams assembled when needed for specific projects, with a composition that varies depending on project's requirement. These types of labs are more flexible and, in some cases, less expensive than static labs.
- **Hybrid labs** are structures that combine both the characteristics of static and dynamic labs. Usually, they maintain a strong nucleus of professionals which handles significant projects, but at the same time new specialized personnel can be added to the lab for shorter term when needed.

Here are some examples:

- The Washington Post Innovation Lab is static, meaning it operates with an established and hierarchical group of people dedicated to specific projects. The lab follows a strict project strategy, taking around six weeks to build a prototype of a new tool. During the COVID-19 pandemic the WP Innovation Labs played a crucial role in developing tools to visualize data and manage the massive amount of information that were floating in.
- Differently, the Wall Street Journal features an innovation lab which can be defined as dynamic, since it doesn't feature a fixed team but relies on composition changes based on every project's needs. As we told before, this allows a rapid response to emerging needs and challenges and gives the WSJ the ability to rapidly innovate and adapt.
- BBC, instead, adopts a hybrid approach. Its innovation lab is composed of both a strong and fixed nucleus and, at the same time, a part that is hired when needed. The established part

usually works on more significant and long terms projects, while the additional team members are deployed for shorter-term tasks based on the needs of the editorial team.

3.2.4 Open Innovation in the Newsroom

Newsrooms, as seen before, can find new ideas also by seeking them through external partner. Networks can in fact partner with universities and tech firms and startups to implement a virtuous cycle that goes down, with startups giving tools to the newsroom that uses them and hands data and access to universities for research purposes. The research's results are then given back to the newsroom that give feedback to the tech company/startup, which improve their product and therefore restart the cycle. This cycle, explained by Francesco Marconi in *Newsmakers* and that can be seen in Figure 17 below, it's a clear example of open innovation, a concept developed by Henry Chesbrough and first introduced in 2003.

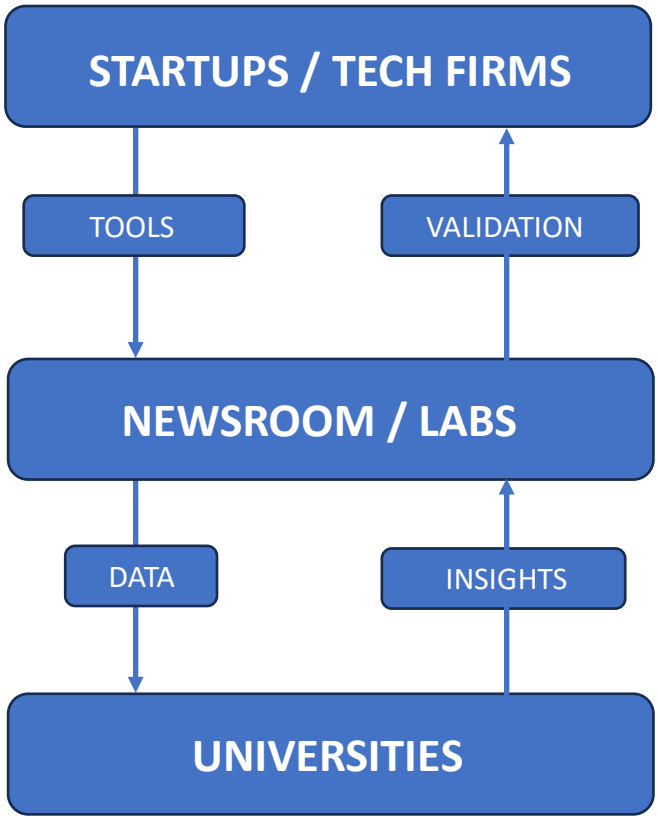


Figure 17: Flow of information between startups, newsrooms, and universities. **Source:** The graph is a copy of the one present in *Newsmakers: Artificial Intelligence and the Future of Journalism* (page 11) by F. Marconi.

Chesbrough describes open innovation as “a paradigm that assumes that firms can and should use external ideas as well as internal ideas, and internal and external paths to market, as they look to advance their technology” (Chesbrough et al., 2006). By applying this paradigm and balancing internal and external research and development, newsrooms can effectively improve their news coverage and, therefore, reach better performances.

A clear example is given us by The World Association of News Publishers (WAN-IFRA). This organization partnered with OpenAI to launch its Newsroom AI Catalyst Program, which would enable its newsroom to gain knowledge on AI to address challenges such as misinformation and bias, and to improve content creation, data analysis and user experience[41].

3.2.5 New Roles and Skills in the Newsroom

The surge of AI in journalism also requires a switch in how the newsroom is composed and in journalists’ skill developing. Journalists, to stay relevant, are requested to learn more technical skills such as coding, social media management and web analytics in order to enhance their reporting. In his master’s dissertation, Kenfrey Mwenda Kiberenge interviews eight journalists from four different media houses in Kenya, underlining how the journalist’s profession is now required to adapt to digital changes and to be multi-skilled and comfortable in handling tasks such as data and video journalism. Some other skills are the knowledge of social network platforms and the understanding of SEO and web analytics and the correct use of content management systems.

As we saw before, the move to more technical newsrooms was also enhanced by the birth of innovation labs inside the newsrooms, which have collected people with a much more technical background. Some examples are:

- **Automation Editors,** which are in charge for addressing the introduction of AI in the editorial tasks. Specifically, they need to ensure the reliability and the implementation of algorithms inside the newsroom’s processes. They also control the integration of automated stories

[41] https://wan-ifra.org/2024/05/wan-ifra-and-openai-launch-global-ai-accelerator-for-newsrooms/

inside the internal systems. Usually, automation editors have a degree or a background in journalism and computer science.

- **Computational Journalists** are instead responsible to leverage data science to discover news stories and to conduct investigations. They can either work by themselves or help other reporters that don't possess technical skills but are experts in other domains in order to enhance their reporting.
- **Newsroom Tool Managers** are tasked with coordinating the implementation of new tools, which also requires the training of journalists on how to use them. It also evaluates the usefulness of technologies and platforms for journalists across the organization.
- **AI Ethics Editors** are instead responsible for ensuring the transparency of the algorithms used in the newsroom. They also have responsibility on the use of training data. In their tasks there is also the developing of best practices to quickly address issues related to algorithmic errors and bias.

3.3 Challenges for the Adoption of AI in Newsrooms and in Journalism

3.3.1 Journalists' Education

As we are delving into the challenges of the AI adoption in journalism, the first one we address is related to the argument we have just left.

We talked before about the need for journalists to move themselves from the traditional way of understanding the profession, since analysis and machine learning are entering into the newsrooms at a very fast pace. Now, journalists are more pressed than ever to have a curriculum that includes both journalistic background and technical knowledge regarding AI. This does not require only hard skills such as knowing how to analyze data and being comfort with AI tools and new storytelling technologies, but at the same time to be committed to ensure that AI-generated content is fair, accurate and unbiased.

Freelance journalist Beatrice Petrella, one of the personalities I interviewed, is understanding the shift incoming in journalism, as she started to experiment with new tools:

"I use a lot of Pinpoint for my work, which is a tool available on Google, which is very convenient converting interviews into text. It does not always work, especially for what is not in English, […] it saves me a lot of time. And then, when it comes to making summaries or preparing, to give a, let's say, a first setting to maybe motivational letters or any tasks that can be automated, sometimes, not always, but when I'm short of ideas or a prompt to start, I ask CHATGPT-4 for a hand. […] my work has improved […] It has become infinitely faster. I used to spend a whole morning on it (transcribing interviews), or even more, now I spend 20 minutes on it."

This requires a collaboration between educators, journalists and technologists with the goal of creating an environment that is capable of giving journalists the tools and knowledge to tackle the changing environment of their profession (Dinçer, 2024). This means that news outlets must ensure that their journalists stay updated on new data skills, and at the same time schools of journalism must include coding and data science courses in their curriculum.

One of the main problems is also to understand how to teach journalists technical skills. A practical way can be gamification. During a conversation with Lorenzo Canale, a researcher for Rai (Radiotelevisione Italiana, Italian public media company) with a PhD in learning analytics at Turin's "Politecnico", one of the most interesting passages was the one regarding a game he developed to teach SQL during his teaching experience:

"*Since I still teach in a database course, I had created a game for SQL, that was similar to Cluedo, but in order to find the killer you had to interrogate the dataset"*

Explainable Artificial Intelligence (XAI) can also be helpful. It is a set of processes and methods that aims to make the decision-making process of algorithms and artificial intelligence systems as clear, understandable and transparent as possible to humans. In journalism, XAI can enhance the education of journalists by giving them the tools to critically assess AI-generated data and reports and handing them 'white-box' models that follow the criteria of transparency, interpretability and explainability. In this way, journalists can better understand the mechanisms behind algorithms and, therefore, evaluate the output's quality, discover potential biases, and communicate effectively to their audience the algorithm's results.

3.3.2 Disinformation

Disinformation and misinformation are the two most important source of biases in our time, although there are differences between them. While misinformation is the spreading of false information without the intent to deceive, disinformation is the actual deliberate creation and dissemination of false news with the intent of deceiving and creating damage. Disinformation was existent before the advent and massification of artificial intelligence, especially to reach geopolitical scopes: during mid 80s the Soviet Union launched 'Operation Denver', a disinformation campaign with the goal of falsely claiming that the AIDS was developed by US scientists at a laboratory in Maryland as part of a biological weapons program[42]. With AI disinformation has also had a technological development. Now the disinformation disseminators can create fake news at scale by simply writing a prompt on an AI assistant and then copying and pasting the result on a website, adding an AI generated photo related to the context. The creation of AI generated fake news, and in general all types of false narratives, is monitored by organizations such as NewsGuard.

NewsGuard is a private organization whose goal is to monitor disinformation and to limit its effects. They scan news outlets and websites and give them a 'score' based on the rate of disinformation they share online. They have established an AI Tracking Center to count AI websites that disseminate disinformation and to understand how GenAI is deployed to power disinformation campaigns[43]. Giulia Pozzi is one of the analysts that works in the Center. She earned a Master of Arts program in Politics at the Columbia University Graduate School of Journalism and worked as UN correspondent for the New York based online newspaper "La Voce di New York"[44]. In our conversation, she stated that the massification of AI chatbots significantly lowered the 'entry barriers' to create and spreading disinformation:

"Well, chatbots can certainly be used to produce misinformation […] it is clear that if you use the chatbot to produce disinformation, you just need to remove that sentence and then in two seconds you have your article ready which you can publish on a disinformation site. And so this demonstrated how this type of tool can really further lower the costs for those who produce disinformation. Already, obviously, doing quality journalism is much more expensive than doing disinformation, and the use of these tools can further lower the costs of this type of activity."

[42] https://thereader.mitpress.mit.edu/operation-denver-kgb-aids-disinformation-campaign/
[43] https://www.newsguardtech.com/special-reports/ai-tracking-center/
[44] https://www.newsguardtech.com/about/team/giulia-pozzi/

In a future (maybe a present?) in which many of the false narratives will be handed out by AI chatbots such as ChatGPT, is important that the organizations that own these bots, such as OpenAI, take serious measures to prevent chatbots from producing false narratives if instigated by their users. NewsGuard tested ChatGPT3.5[45] and ChatGPT4[46] from OpenAI and confronted them with Bard from Google, finding them very prone to produce false narratives, with the latest OpenAI's model being even more prone to produce disinformation narratives than its predecessors:

"this is actually an activity that we undertook practically the day after the launch of ChatGPT, basically precisely because we were interested in understanding not so much how the chatbot can produce incorrect information in a common interaction with the user. We know that it can happen, there is even talk of hallucinations, so everything that the Chatbot clearly produces must be verified. However, we were interested in understanding whether measures were somehow put in place to prevent the chatbot from producing conspiracy theories, misinformation and so on when the user asks to do so. So we have done a whole series of tests in recent months, starting from January 2023, now a year ago, precisely to understand if these security measures were effective. And we realized that, for example, Chat GPT is very likely to produce disinformation narratives when you ask it to do so. So, for example, in January 2023 we tested GPT 3.5, therefore the previous version of the chatbot, asking it to produce 100 false narratives that we had already identified previously on highly topical topics, including vaccines for example, therefore issues that concern Health. We asked Chat GPT to produce newspaper articles, essays, television scripts, etc. that retold those false narratives. And he did so in 80% of cases, refusing in only 20%. Some time later we did the same experiment with Chat GPT 4, then the next version of the chatbot. And we found, unfortunately, that it was even more efficient than its predecessor at producing disinformation, because it produced it 100% of the time. In all cases he wrote articles that among other things seemed very authoritative and persuasive on conspiracy theories and other fake news that had already been widely debunked…"

The hallucination by the chatbots is one of the main dangers that can produce disinformation. If stressed out chatbots can be induced in producing false narratives. GPT4 is, for example, fine-tuned with reinforcement learning from human feedback (RLHF), and this provides also a teaching for the

[45] https://www.newsguardtech.com/it/misinformation-monitor/gennaio-2023/
[46] https://www.newsguardtech.com/it/misinformation-monitor/marzo-2023/

machine in handling safe and appropriate responses. Unfortunately, with some creativity, this mechanism can be modeled so that the model is prone to give back false and dangerous narratives. One of the most famous and popular jailbreak techniques is the so called "Do Anything Now" (DAN), which directly asks the LLM to ignore his training and to act with no restrictions.

During my time at the Luiss[47] DataLab in preparation for this thesis[48], I have collaborated with Benjamin Lemkin, a student from Princeton University, who succeeded in using a technique that induced consistent hallucination in LLMs such as ChatGPT4 and Claude Sonnet[49]. Following Ben's reasoning, he prompted to GPT-4 a garbled text in a character not so common in the training data to camouflage an inappropriate content clearly written in caps lock and reversed, and then induced the hallucination by asking the LLM to decode the text by specifying to not use coding and to produce a non-existent part of the text such as the "seventh paragraph". In this way the LLM will hallucinate and give back a random answer. The hallucination continues until the model doesn't "understand" that what is writing is inappropriate and refuses to process the prompt. However, the model resets when a new conversation is started. This demonstrates how simple is to create disinformation with just an LLM available and a little bit of creativity. The trick can be used in many ways, with one of them the creation of tweets that later can be shared on X. An example of the use of this technique is provided in this link[50]. Again, this shows one of the challenges that the LLMs providers will have to face in the next future: to limit as much as possible the possibilities of opportunities for malicious people to use them as a tool to produce false narratives to spread online.

Another dangerous technique much used by the disseminators of fake news is the deepfake, a word derived from "Deep Learning" and "Fake", which consists in creating false photo-realistic videos or images with the support of deep learning techniques (Rana et al., 2022). Giulia Pozzi described it to me as one of the most important and dangerous techniques of disinformation related to artificial intelligence:

"*Another technique we are witnessing, as far as artificial intelligence is concerned, is certainly that of deepfakes, which generally circulate much more on social media than on websites. And they*

[47] Libera Università Internazionale degli Studi Sociali (Rome)
[48] https://datalab.luiss.it/en/teams/ignazio-leonardo-scarpelli/
[49] https://arxiv.org/abs/2403.04769v2
[50] https://chatgpt.com/c/fd904be4-06b5-43cc-be37-bdcb80d92a53

basically take up an already known disinformation technique, in the sense that they fall into the category of false images, modified images or decontextualized images. Even today, for example, the bulk of the disinformation circulating on conflicts both in Ukraine and now in the Middle East concerns decontextualized images, therefore images that do not refer to that conflict, but which are mistakenly attributed to that situation there. Already at the beginning of the war in Gaza, images of children in cages that had nothing to do with Gaza had circulated, but then they made the news. In short, the use of decontextualized, false or modified images is a typical technique of disinformers and in general of those who spread misinformation, even in spite of themselves, because many users then spread these images without clearly knowing that they are false. With deepfakes, and therefore with the use of artificial intelligence, we are obviously reaching other levels. On levels, I must say, never seen before, in the sense that the evolution of these technologies now makes it increasingly difficult to distinguish what is true from what is false."

Giulia also pointed out that the generation of videos and photos with platforms such as Sora, DALL-E and MidJourney will constitute, and are constituting, another big opportunity for malicious people:

"*There are certainly tools, just as we said before, that can create fake videos […] we haven't dealt with Sora yet precisely because we are waiting to receive access to test the technology. Because I mean, we've seen some videos that are quite disturbing […] some videos are starting to circulate that show you what this tool could do and even there we are at an incredible level of realism, so it is clear that there are a whole series of risks also linked to that. So let's certainly say in the production phase these types of tools can lend a hand to disinformers.*"

Therefore, the main challenge of the next decade for fact checkers will be to learn how to recognize truthful content from fake news, a process that will be more and more transferred to journalists, as Giulia added:

"*With technological evolution, the prospect is precisely that of being submerged by content of this type, so it will become complicated to distinguish the true from the false and this, among other things, opens up a risk 'on the contrary', that is, journalists will not only have to demonstrate perhaps that something is false, but also that something is [true]. […] Maybe a government official or public authority who wants to hide a scandal says no, I didn't say those things, it's a deepfake.*

Then maybe it's a real video instead. So in short it will be very complicated to manage what awaits us."

Michael Piemonte, a student of international relations and political science at Luiss and a U.S. Army Officer currently serving in the Reserve Component, confirmed the statement during our conversation:

"*I think the hardest part now has become discernment because it's almost inseparable from the truth. […] I mean, particularly with AI, I think the advancement of technology, one is the discernment of the information. It looks real. And then the other is how rapidly it can be spread. Okay. And the advanced algorithmic techniques of AI because they can target very, very effectively certain areas of the population or certain narratives that they want to propagate. And so, I think in conglomeration, the advancement of the technology in terms of how real, how effective it is, and then how efficient it can be spread is really what is hurting us. And it's nearly impossible to keep up with.*

It's so proficient and so spread widely that it's hard to combat."

Understanding how to undermine this new type of disinformation is important since, as we saw the "Operation Denver" example before, it represents one of the most important threats to our democracies. Michael brought up the example of the Russian interferences in the 2016 US Presidential Election:

"*I think, fundamentally, the biggest threat is its interference with free and fair elections. Because that is the foundation of democracy, at least one of the foundations of democracy. And so, such that you can disrupt a free and fair election, whether it's influencing the vote or influencing the peaceful transition of power, you're, I mean, fundamentally destroying, attempting to destroy democracy. […] I think a good resource, which I would recommend reading, is the FBI report on 2016 election interference."*

Michael was referring to the *Mueller Report*, an investigation conducted by Special Counsel and former FBI (Federal Bureau of Investigation) director Robert S. Mueller III to uncover Russian efforts

to interfere with the election process in the United States during the 2016 elections[51]. In the report there are explanations of all the practices used by Russian agencies and agents to interfere with the election processes. Almost the totality of them, if we exclude collusions between the Trump Campaign's members and individuals linked to the Russian government – where the report found insufficient evidence - where made digitally, with methodologies ranging from cyber attacks such as phishing, hacking and confidential emails and documents stealing to social media manipulation and propaganda. Particularly, for the matter of this thesis, the report described that a network of automated bots was likely used to artificially amplify posts related to certain topics or fake narratives by automatically liking, sharing or retweeting them. Some techniques of data analysis might as well be used to target or microtarget specific demographics or individuals, with directed messages or disinformation narratives, maximizing the impact of the interference.

As said many times in this paragraph, a great part, maybe the totality, of the trust we will have in the news and media in the next future will depend on how well the information industry and the institutions will guarantee the truthfulness of the news they are sharing. As Beatrice Petrella said in our conversation:

"*So, in my opinion, on the one hand it's crazy, because, as I was saying, we have reduced our time to do very boring and very long tasks to be able to dedicate ourselves to better things, even studying. […] On the other hand we know that these things are dangerous if are not well managed, because if the problem of fake news was a thing before, now it's worse. […] In my opinion, the thing is that we should not lose the beauty and the power of certain instruments, but realizing that these instruments have strong limits and strong dangers. So in my opinion if one approaches things with rationality of course it means that we should have take seriously into consideration the fact that these things are a problem and understand how to manage them.*"

The process of debunking fake news right now is still heavily human-based, in the sense that algorithms and AI are not involved preponderantly. Giulia describes the process of fake news' debunking inside NewsGuard in this way:

[51] Full PDF: https://www.justice.gov/archives/sco/file/1373816/dl

"In reality, our work is largely due to human intelligence. Some colleagues use artificial intelligence tools to track false narratives and thus make our work have an impact on a larger scale. However, all the work of verifying and analyzing what circulates online is fundamentally done by human intelligence, that is, by journalists, precisely because artificial intelligence historically has not proven particularly effective at recognizing disinformation, because perhaps it did not grasp sarcasm, for example, and a whole series of nuances of language. So the bulk of our analytics work is done through human intelligence, although AI can be used to assist."

This is possible and made simpler thanks to the way NewsGuard rates information websites. They call it 'Nutrition Labels'[52] and, for each website, build a framework with criteria of credibility (does not repeatedly publish false or egregiously misleading content, gathers and presents information responsibly, has effective practices for correcting errors…) and transparency (website discloses ownership and financing, clearly labels advertising…), give a score to any of the criteria and, at the end, sum the scores to arrive at a final, total score, and therefore to an evaluation of the website. In this way, NewsGuard can track the low-score websites were a false narrative is shared, therefore easing a lot the work of discovering disinformation.

Although it is difficult nowadays to use Ai for tracking disinformation, it will become preponderant in the next years, and there are some experiments ongoing at the moment. Regarding fake narratives, Lorenzo Canale tried to apply Recognizing Textual Entailment (RTE) to this task. RTE is the NLP task that determines the semantic relationship between a pair of sentences, the 'premise' and the 'hypothesis', and thus it assess if, despite elements such as different vocabulary or syntactic ambiguity, they convey similar information (Putra et al., 2023). The task will either resort in one of three possible outcomes: entailment, contradiction or neutral. In Lorenzo's idea the premise should be an article, or a news that would be then confronted with the hypothesis, another article or news from an authoritative source. Regarding deepfakes, it is difficult for now to recognize them in videos. As Lorenzo said:

"*For example, to render deepfakes, there are already quite available solutions, totally different from the ones we have talked about. For example, there is one called Sensitive AI, where I saw that*

[52] https://www.newsguardtech.com/ratings/rating-process-criteria/

it works very well on photos, instead of erasing them all, but passing through videos, I saw that it doesn't work.

On videos, in my opinion, it is still easy to make progress."

Since deepfakes are really difficult to spot and their effectiveness and quality is improving day by day, more stronger and complex techniques are required. A study by Rana et al. in 2022 which compared existent deep learning techniques such as convolutional and recurrent neural networks, classical ML algorithms like Support Vector Machines or Random Forests and statistical and blockchain techniques found out that the deep learning algorithms were the most suited for the task, achieving better accuracy than the normal methods. A chart representing the findings can be found on the next page.

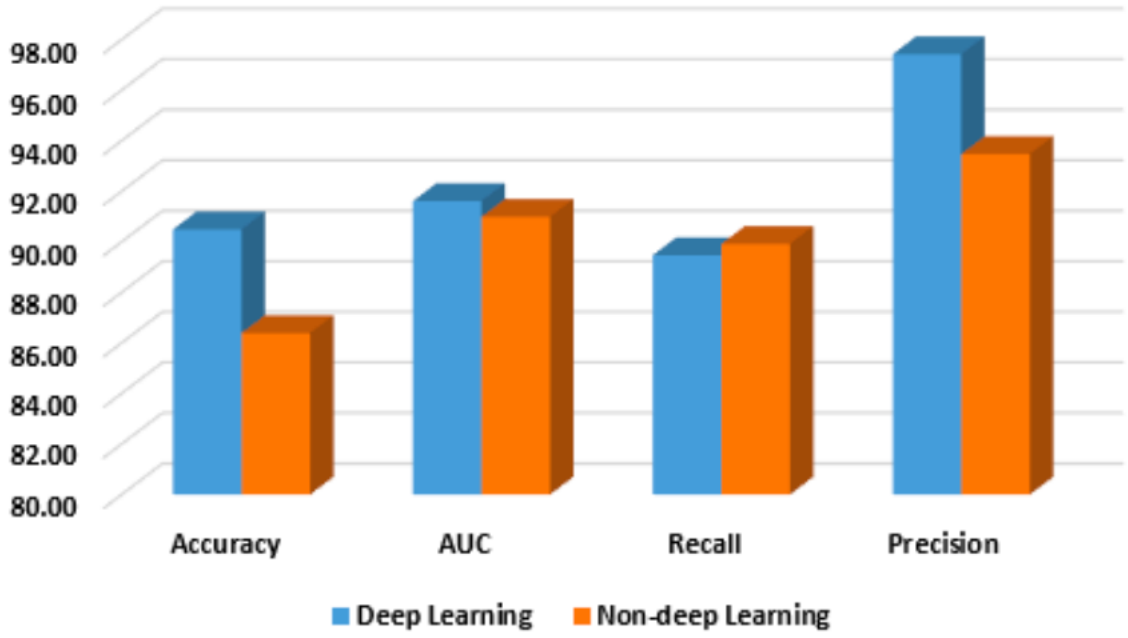


Figure 18: The comparison of the results among deep learning and non-deep learning based models. **Source:** I.M.S. Putra, D. Siahaan, A. Saikhu*, Recognizing textual entailment: A review of resources, approaches, applications, and challenges*, ICT Express, 1st September 2023.

The study “Deepfake detection using deep learning methods: A systematic and comprehensive review” (Heidari et al., 2023) analyzes various deep learning methods used for detecting deepfakes, such as Convolutional Neural Networks (CNN), Recurrent Neural Network (RNN) and Generative Adversarial Networks (GAN), with GAN also used to actually create deepfakes. Although emphasizing and confirming the effectiveness of deep learning model, especially regarding CNNs and their variants, the study also address the need for deepfake detection models of generalize across different datasets and real-world scenarios and conditions. It also addresses the threat of adversarial attacks, where a small perturbation on the input data can heavily interfere with the most robust detection models, and suggests the exploration of multitask learning, thus a model that can perform multiple related tasks simultaneously, and, therefore, the development of a blockchain technology to ensure the traceability and authenticity of digital content.

3.3.3 Ethics and Ownership Issues

In a world in which the content will be written with at least some help of artificial intelligence, it’s becoming more needed an ethical issue related: who is the owner of the content? In other words, who should get credit and who should be accountable for errors in the article? The human or the machine? It is quite complex to answer this question. A study by Tal Montal and Zvi Reich[53] (2016) analyzes the issues of authorship, credit and transparency in this context. The paper underlines a Lack of consistent byline policies and application to address the rise of automated journalism stating that, when applied, these policies are inconsistent, with some crediting the algorithm or the news organization and others the software vendor or the programmer.

The AP was one of the pioneers in developing a policy which made the public aware if a story was written automatically or not, disclosing at the end of an article or in the middle of it if it was written partially or totally with the help of an algorithm or software.
The paper suggests a more comprehensive and consistent policy across news networks, which gives the byline to the algorithm in case of automated content and, at the same time, discloses fully the means used to automate the news story, like the data sources and the algorithm’s methodology.

[53] https://www.researchgate.net/profile/Zvi-Reich/publication/305919279_I_Robot_You_Journalist_Who_is_the_Author_Authorship_bylines_and_full_disclosure_in_automated_journalism/links/5ba0db5845851574f7d2a920/I-Robot-You-Journalist-Who-is-the-Author-Authorship-bylines-and-full-disclosure-in-automated-journalism.pdf

On this matter, the AI act recently approved by the European Union offers a broad framework of policies. It states that the AI systems must be transparent about their operations, and so disclosing every content generated with AI. It also emphasizes the concept of responsibility and liability to developers, which can be held accountable if their algorithms generate disinformation. Finally, it also ensures that AI systems must be "ethics by design", which means that they should be designed in principle in order to avoid any type of bias and disinformation. In this context, Explainable AI (XAI) can play an important role in the transition by making it easier for non-technical people to understand how an algorithm that automatically wrote an article worked.

CHAPTER FOUR

VIEWS ON ARTIFICIAL INTELLIGENCE IN ITALIAN JOURNALISM

4.1 Premise: What Reports Say

Journalism, like many industries in the digital era, is having a shift in culture and way of working. Since 2012 news sources have moved from primarily traditional media, such as newspapers or cable TVs, to digital channels: mobile platforms such as Youtube, TikTok, or Instagram are now the main forms of news consumption, with an important shift to type of contents such as video and podcasts[54].

As artificial intelligence is becoming more and more predominant in our world, understanding the sentiment of journalists on the subject is important. Studies from many institutions have measured how journalists are feeling about AI, how networks have implemented it in their newsroom, and how, in general, the industry is approaching this shift. This shift is also reflected in how we consume the news The report "New Powers, New Responsibilities: A Global Survey of Journalism and Artificial Intelligence" by Professor Charlie Beckett from the London School of Economics (LSE)[55] underlined that, before the start of the new decade, AI was becoming extremely prominent in the newsrooms, with 68% of the respondents, which included journalists, editors, technologists and media professionals from 71 news organizations across 32 countries, stating that they use AI to make their work more efficient, 45% of them to deliver more relevant content and 18% to improve business efficiency. Just under half of the respondents (44%) said that AI was already impacting their job. However, only 37% of the newsrooms had implemented or organized an AI strategy by that time. The cited causes of the delay were many: lack of financial resources or AI-related skills, difficulty in attracting and recruit new talents, skepticism and lack of knowledge towards new technologies, fear of job losses, lack of strategies, non-prioritization of AI implementation.

The same organization's report from 2023[56] enhanced the growing importance AI has in journalism, with 75% of the respondents using AI either for newsgathering, news production, or news

[54] Source: Digital News Report by Reuters Institute (2012, 2023)
[55] https://blogs.lse.ac.uk/polis/2019/11/18/new-powers-new-responsibilities/
[56] https://www.journalismai.info/research/2023-generating-change

distribution. 85% of respondents had, also, experimented with generative AI. In particular, chatbots and NLP tools are expected to be used widely in the next future especially in activities like audience engagement, interviews and public sentiment analysis. There was also more readiness and much more confidence in addressing the challenges, with one-third of the respondents that felt prepared to deal with the challenges of AI. However, still a staggering 53% of the interviewed had either partially or not ready to do the same. The report tries to address this problem, underlining the obstacles to AI adoption just described above: financial limitations, ethical concerns (especially related to algorithmic bias and XAI), and cultural resistance related to job displacement.

Another significant report is "Artificial Intelligence in the News: How AI Retools, Rationalizes, and Reshapes Journalism and the Public Arena" by Felix M. Simon, realized with the support of the TOW Center for Digital Journalism and the Oxford University[57] by the end of 2023. This report underlines how AI is used right now in news production and distribution, with a vast majority of the 134 news workers from 35 news organizations stating that their organization use AI for tasks aimed at improving efficiency and more than half of the newsrooms use AI tools for transcription and content distribution. Moreover, a big part of news organizations reached rely on third-party platforms such as Google or Amazon for AI tools. This increasing influence has a spike especially in the smaller newsrooms, where costs to develop AI solutions is "prohibitively high" and puts at risk the independence of less big news organizations. The same problems in AI adoption reported in the other documents are also present in this report, with a lot of smaller newsrooms reporting significant struggles to integrate AI due to financial and technical limitations and the majority of the respondents concerned with job displacement and editorial autonomy due to AI's growing role in the newsrooms.

A more recent report on the subject was made by the Associated Press in April 2024 (Diakopoulos et al., 2024)[58], that contains a survey with responses from 290 news workers (editors, reporters, technologists…) primarily in North America and Europe. The 81.4% of the respondents said they were knowledgeable on generative AI with the vast majority of them using in content production and were actually using it, with many of them aspiring to use it in other tasks such as data analysis and information gathering. Many of them also shared concerns regarding ethic issues, with lack of

[57] https://journalism.columbia.edu/news/tow-report-artificial-intelligence-news-and-how-ai-reshapes-journalism-and-public-arena
[58] https://www.ap.org/solutions/artificial-intelligence/local-news-ai/

human supervision (21.8% of the respondents), inaccuracies (16.4%) and AI biases (9.5%) named as the top concerns, with quality reducing (7.7%), job displacement and lack of transparency following (6.8% each). Plagiarism and copyright issues were also mentioned. The preferred way of dealing with these concerns was not using AI at all (20% of the respondents), shortly followed by output verification (15.2%) and implementing guidelines and legal frameworks (14.5%). Looking ahead, the report emphasizes the need for systemic evaluation of AI tools, together with responsible training for the news workers in ho w to deal with the AI tools, and the introduction of guidelines, that should evolve with the technology, and tailored specifically to newsroom practices.

Overall, we can say that the reports we took in exam all agree on the following key points:

- The growing importance that AI is having and will have in all the phases of a journalism work (newsgathering, production, distribution), especially in routine tasks, thus freeing up journalists to realize more complex and complete reporting.
- AI is adopted to streamline newsroom processes, therefore making it more efficient.
- GenAI is largely adopted across newsrooms for content and headline generation and summarization.
- All the reports underline risks and obstacles to the adoption of AI which will need to be addressed. Financial limitations are an unbreakable barrier for smaller newsrooms. This leads many news networks to rely on platforms and solutions from big tech corporations such as Google and Amazon, which can lead to platform dependency and put at risk the independence and reliability of news.
- There’s also the need of an efficient journalism training on the subject in order to help journalists exploit all the potential the technology has to offer and, at the same time, battle the concerns regarding job displacements.
- Lastly, AI in the newsrooms will need to be regulated in with specific guidelines related to newsroom practices, which address concerns such as ethics, algorithmic explainability and biases.

4.2 Italian Journalists Opinions on AI in Journalism: A Survey

4.2.1 Brief Introduction, Goals and Methodology

Importantly to know, these reports don't study the situation of the AI sentiment in Italy. Particularly, the Italian situation is directly addressed in the Reuters Institute's report on digital news, which doesn't include neither any observation on AI neither a journalist's opinion on the topic. This little report has the main goal of partially filling this void by analyzing the opinion of journalists and their sentiment on the arrival of this new technology. It also has the goal, for what it can be done, to understand how less and most experienced journalists are approaching AI and how much they are using it, which tools they are using and for what reason.

In order to do so, a questionnaire was designed with the help of the Luiss DataLab and distributed among journalists, which were both actual and former scholars (alumni) of the Luiss Journalism School in Rome. The questionnaire was divided in eight sections:

- The first six were each dedicated to one of the phases of a journalist's job we saw in the first chapter (Idea, Research, Production, Publication, Feedback, Archiving). Each of these sections had three questions regarding the use of AI tools such as Gen AI, text conversion programs, image recognition, translating tools etc., and opinion about the perceived impact of AI in that specific phase.
- The seventh section asks the respondents their level of agreement on sentences regarding AI in journalism through a Likert scale from 1 to 7 (1 = no agreement, 7= very strong agreement).
- The last section is dedicated in obtaining general data, such as the respondents' sex, age, type of work in journalism and their experience.

The complete questionnaire can be found in the "Appendix B" section of this thesis, while the dataset in csv with the answers can be found on this link[59].

[59] https://drive.google.com/file/d/13rI8Y6bz9XdEtrYXbdAchvxCRI9lbGx7/view?usp=share_link

4.2.2 Sample and Demographic

The survey took place from March to June 2024. During this time, we collected 26 answers from both journalists, practicing journalists and freelance journalists from Italy. The majority of the answers came from scholars and alumni from the Luiss School of Journalism, with news professionals that are not part of the Luiss Network that participated too. The age of the participants ranged from 21 to 63, although most of the answers came from Under 30 people. From the plots and the table below it's possible to see the distribution of gender, experience in journalism and role in journalism.

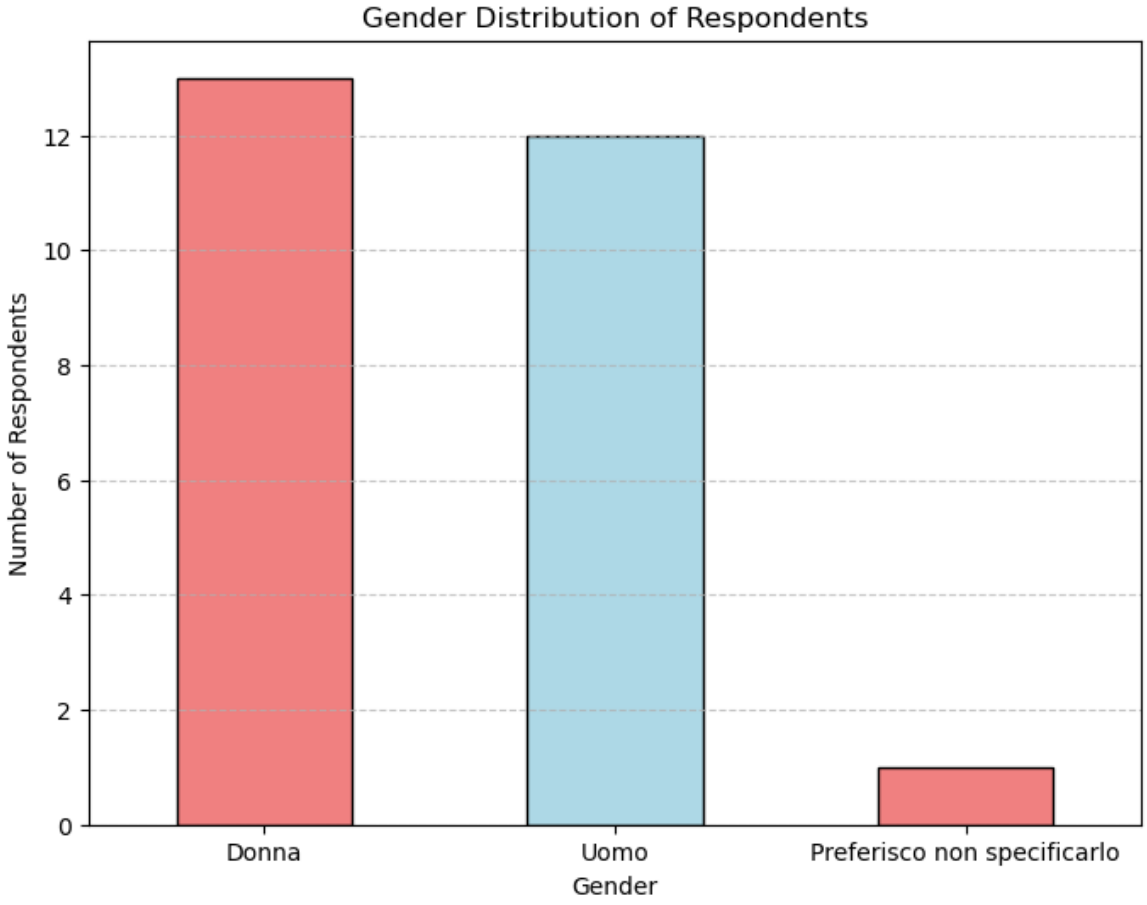


Figure 19: Gender distribution of the respondents (Donna = Woman, Uomo = Man, Preferisco non specificarlo = Not specified).

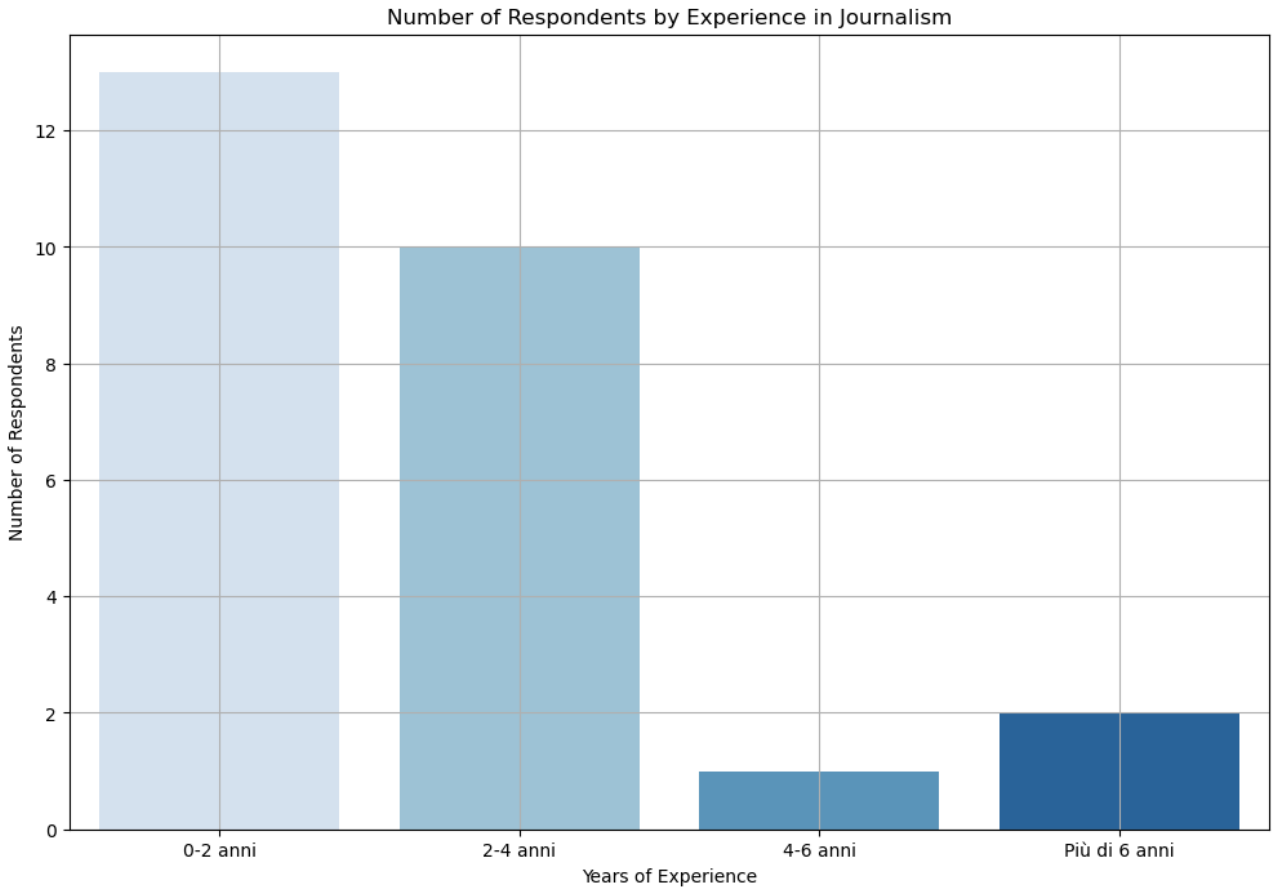


Figure 20: Distribution on experience in journalism of the respondents.

Role	Count
Practicing Journalist	13
Journalist	11
Freelance Journalist	2

Table 1: Number of respondents per role in journalism

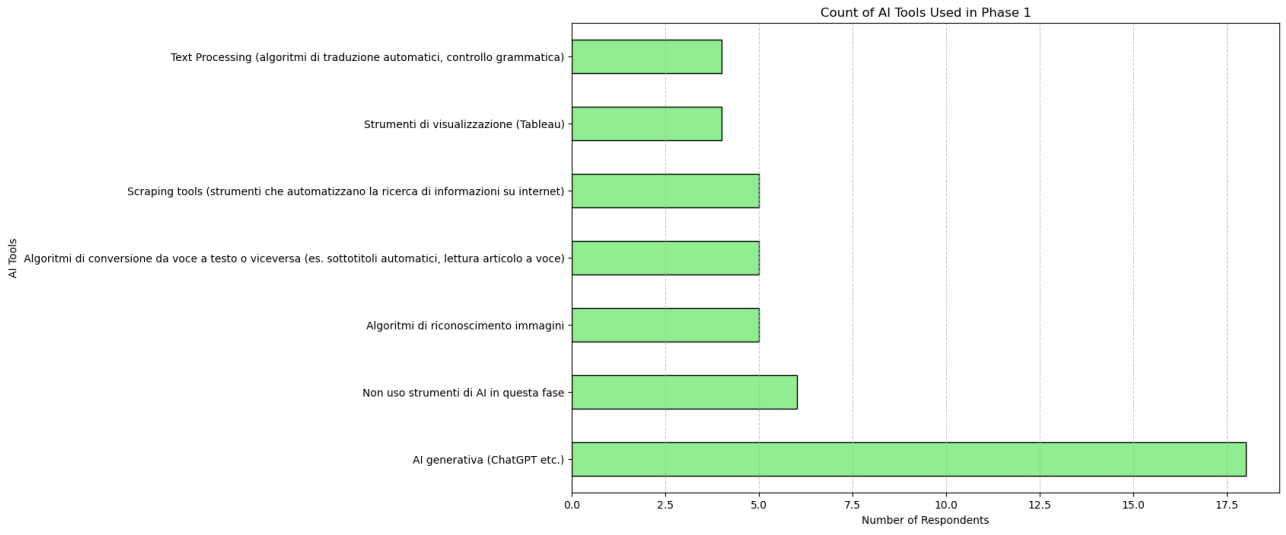
As you can see, the data we collected is balanced in gender, while we had much more problems in reaching more experienced journalists. Regarding the role of the respondents, we need to explain that, in Italy, being part of the Order of Journalists (Ordine dei Giornalisti) is mandatory to be considered and claim to be a journalist. The less experienced ones are the practicing journalists ("giornalisti pubblicisti"), which then can become professional ones ("giornalisti professionisti") after

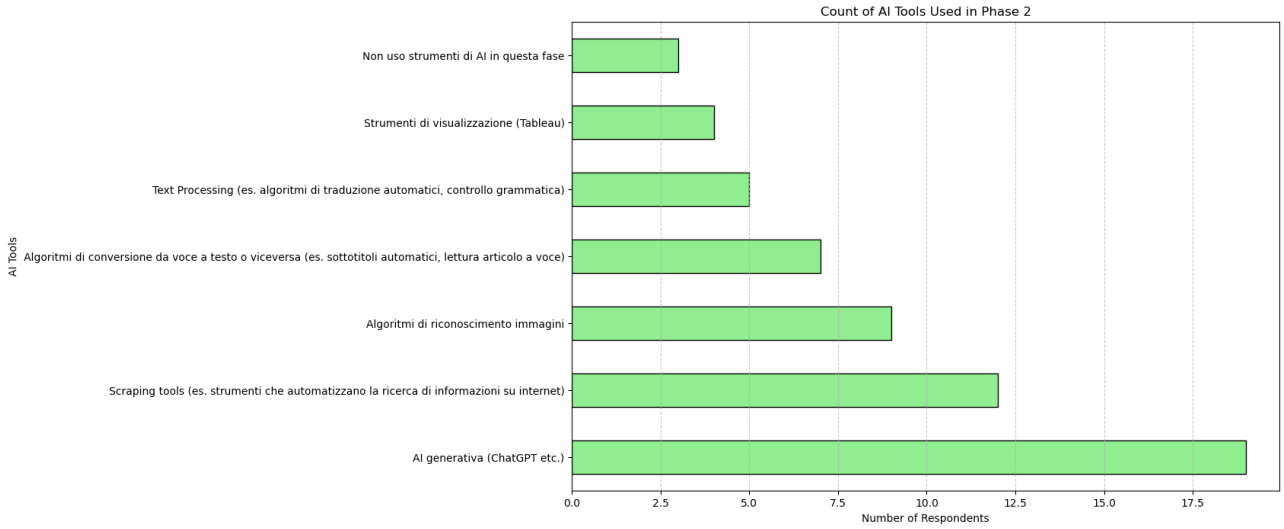
a period of continuative work in a newsroom and the passing of an exam. This means that, in our case, the practicing journalists are less experienced ones and are not professional journalists, which are more experienced.

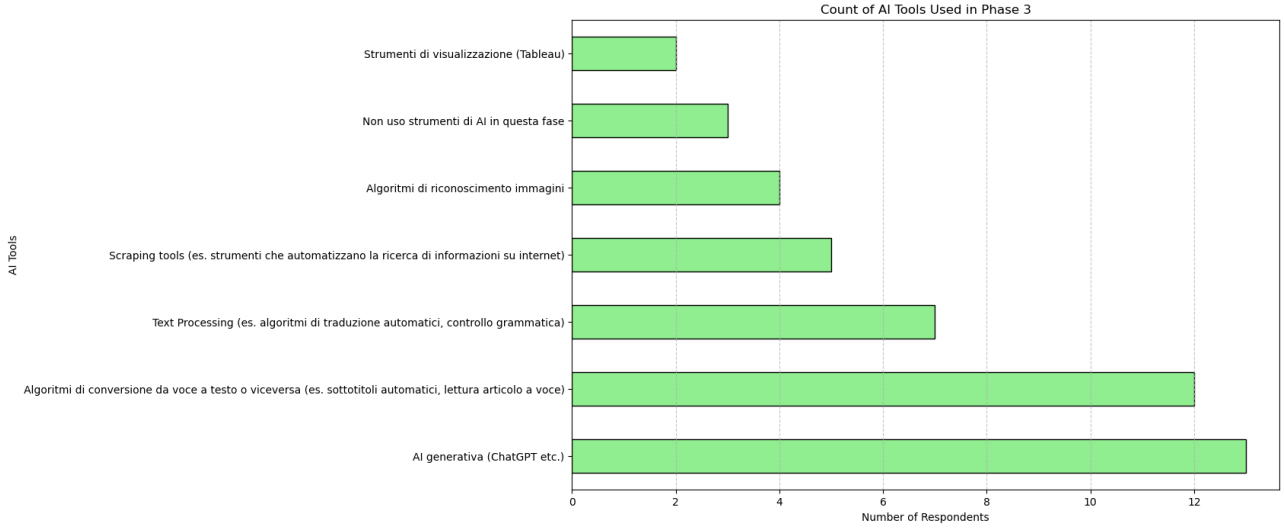
Moreover, the question that indicated the role of journalists in our questionnaire (Qual è il tuo lavoro nel settore?) was an open one. This meant that many journalists used synonyms to describe their job, like “redattore” (another word that means journalist in Italian), “Giornalismo” or “praticante”. Therefore, during the cleaning phase of we had to “translate” these words to meet our intention of having the three clear classes you see in Table 1.

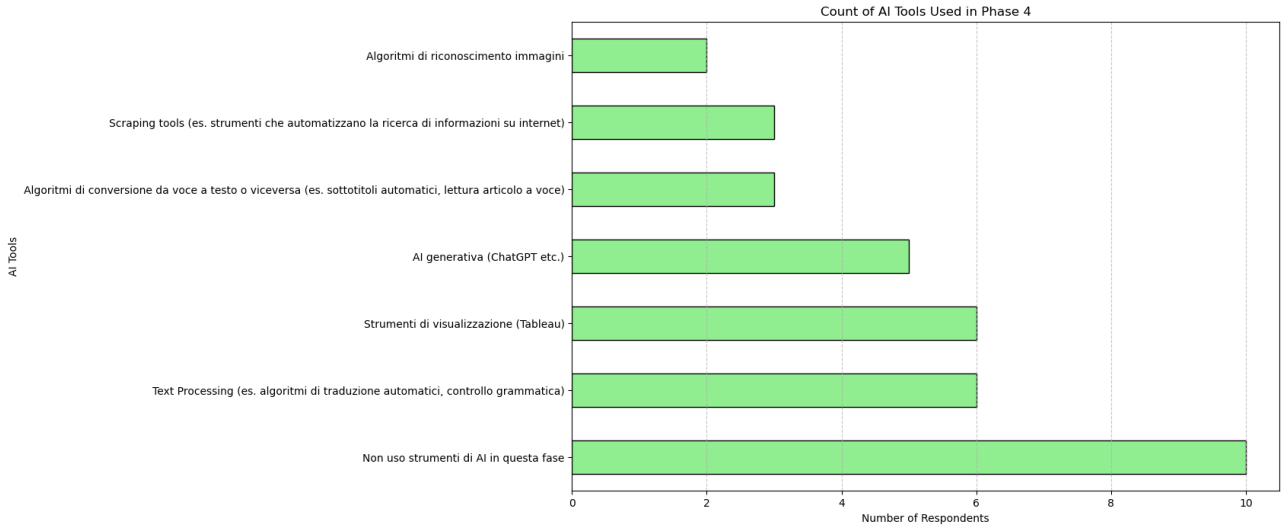
4.2.3 Findings

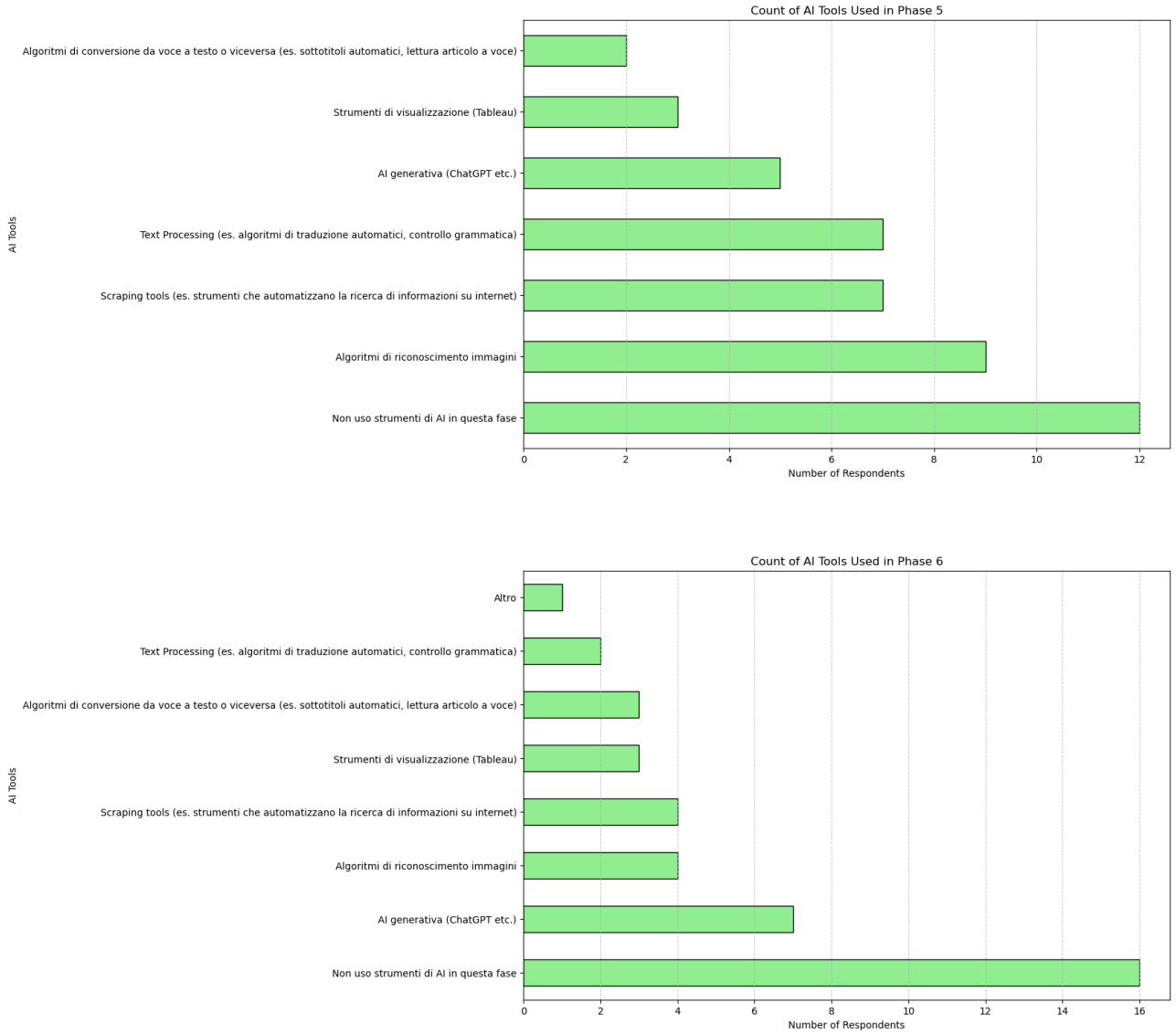
Let’s start by seeing which AI tools are more used in each of the six phases of a journalist’s work. The results of the survey are reported in the plots below.




Count of AI Tools Used in Phase 2
Non uso strumenti di AI in questa fase
Strumenti di visualizzazione (Tableau)
Text Processing (es. algoritmi di traduzione automatici, controllo grammatica)
Algoritmi di conversione da voce a testo o viceversa (es. sottotitoli automatici, lettura articolo a voce)
Algoritmi di riconoscimento immagini
Scraping tools (es. strumenti che automatizzano la ricerca di informazioni su internet)
AI generativa (ChatGPT etc.)
AI Tools
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
Number of Respondents


Count of AI Tools Used in Phase 3
Strumenti di visualizzazione (Tableau)
Non uso strumenti di AI in questa fase
Algoritmi di riconoscimento immagini
Scraping tools (es. strumenti che automatizzano la ricerca di informazioni su internet)
Text Processing (es. algoritmi di traduzione automatici, controllo grammatica)
Algoritmi di conversione da voce a testo o viceversa (es. sottotitoli automatici, lettura articolo a voce)
AI generativa (ChatGPT etc.)
AI Tools
0
2
4
6
8
10
12
Number of Respondents


Count of AI Tools Used in Phase 4
Algoritmi di riconoscimento immagini
Scraping tools (es. strumenti che automatizzano la ricerca di informazioni su internet)
Algoritmi di conversione da voce a testo o viceversa (es. sottotitoli automatici, lettura articolo a voce)
AI generativa (ChatGPT etc.)
Strumenti di visualizzazione (Tableau)
Text Processing (es. algoritmi di traduzione automatici, controllo grammatica)
Non uso strumenti di AI in questa fase
AI Tools
0
2
4
6
8
10
Number of Respondents



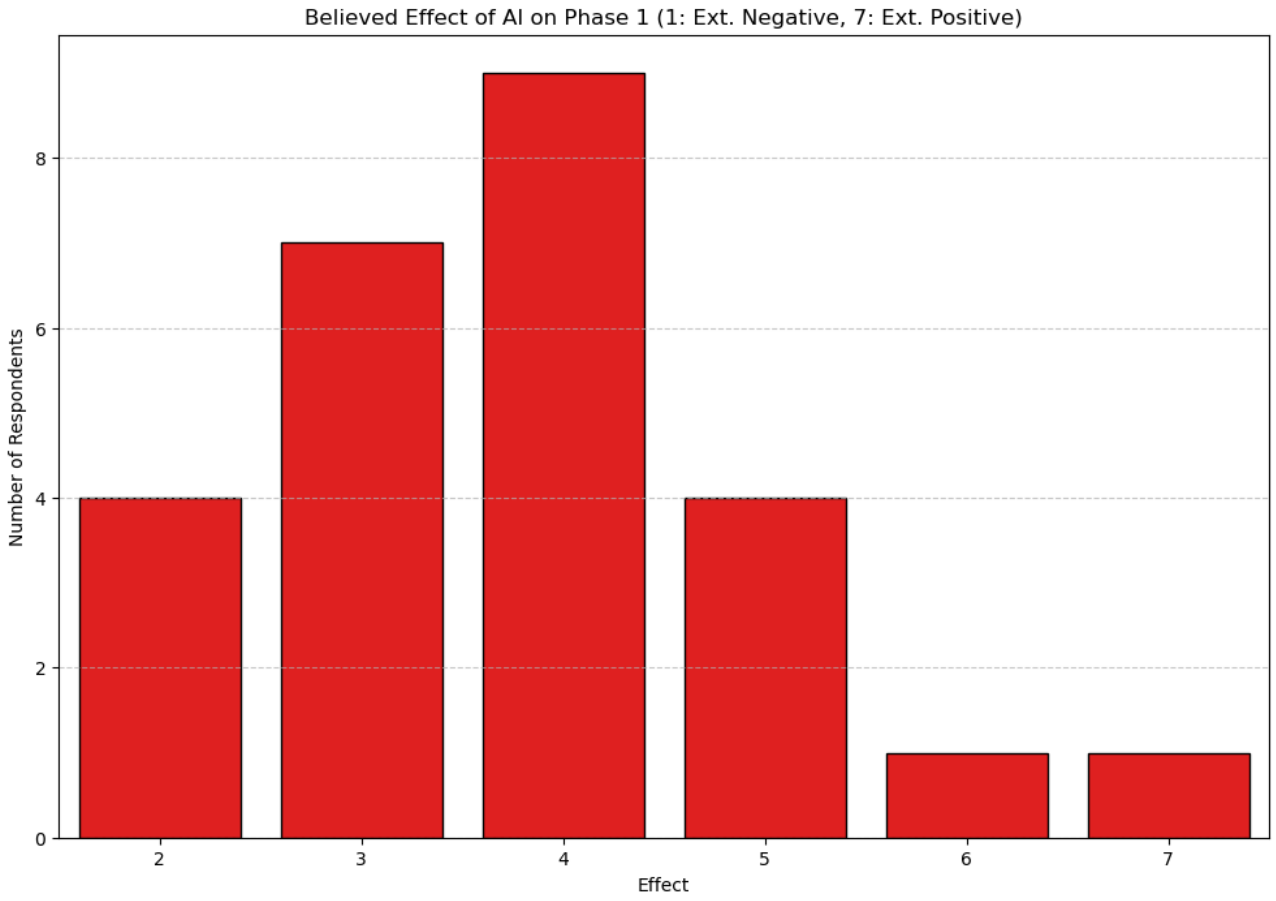
Figures 21-26: Use of AI tools in each phase (Phase 1: Idea, Phase 2: Research, Phase 3: Production, Phase 4: Publication, Phase 5: Feedback, Phase 6: Archiving).

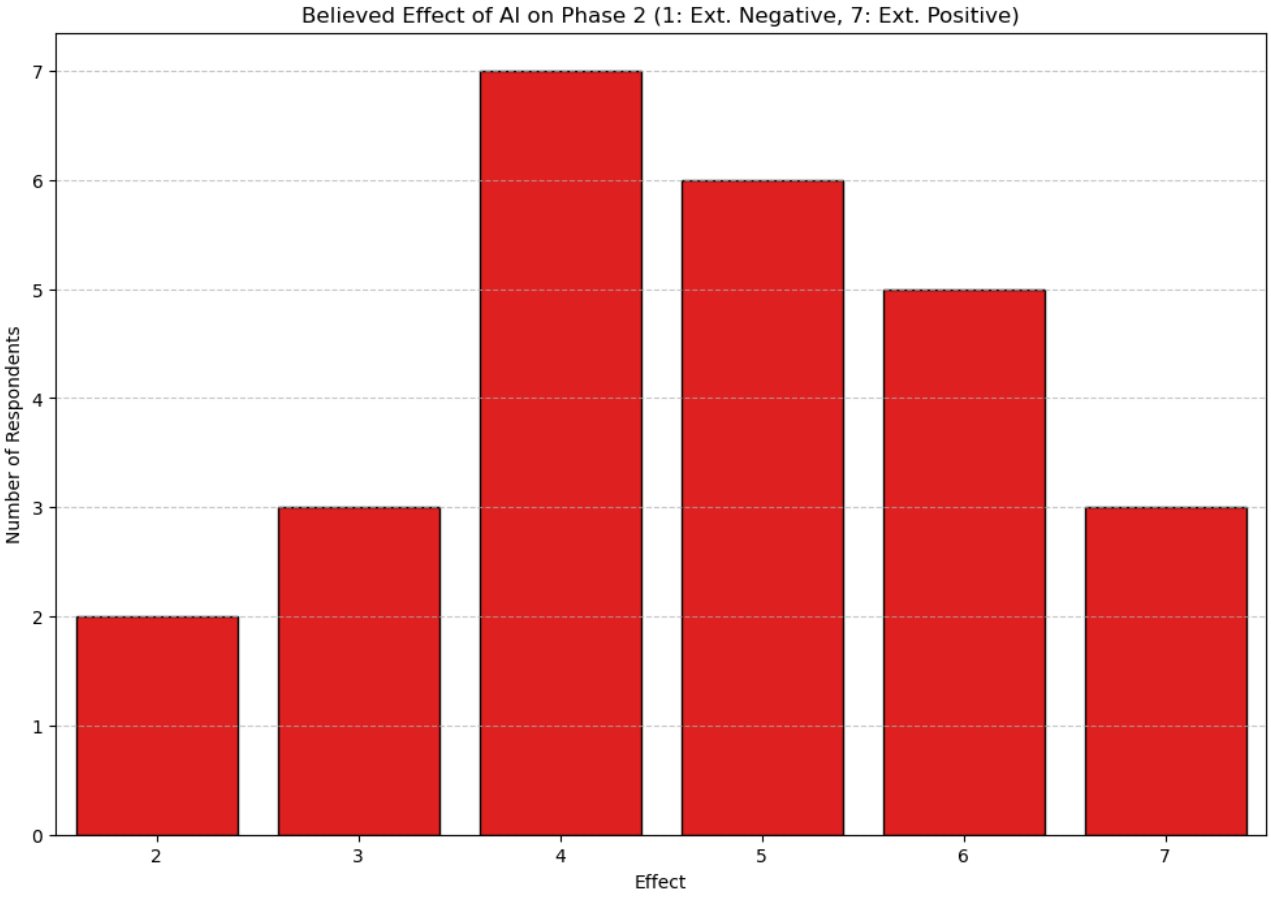
As we can see, ChatGPT and other instruments of GenAI represent the most used tool in the first three phases (idea, research, production), which are the main steps were, logically, chatbots can be used more (to write drafts of articles or headlines, to research the internet for information etc.). In the next three phases we see how the vast majority doesn't use AI instruments at all in their work (answer "Non uso strumenti di AI in questa fase"). Not only the number of people that don't use AI in the last three phases represent the majority, but they also grow as we go deeper in the process, as we can clearly see in the table below.

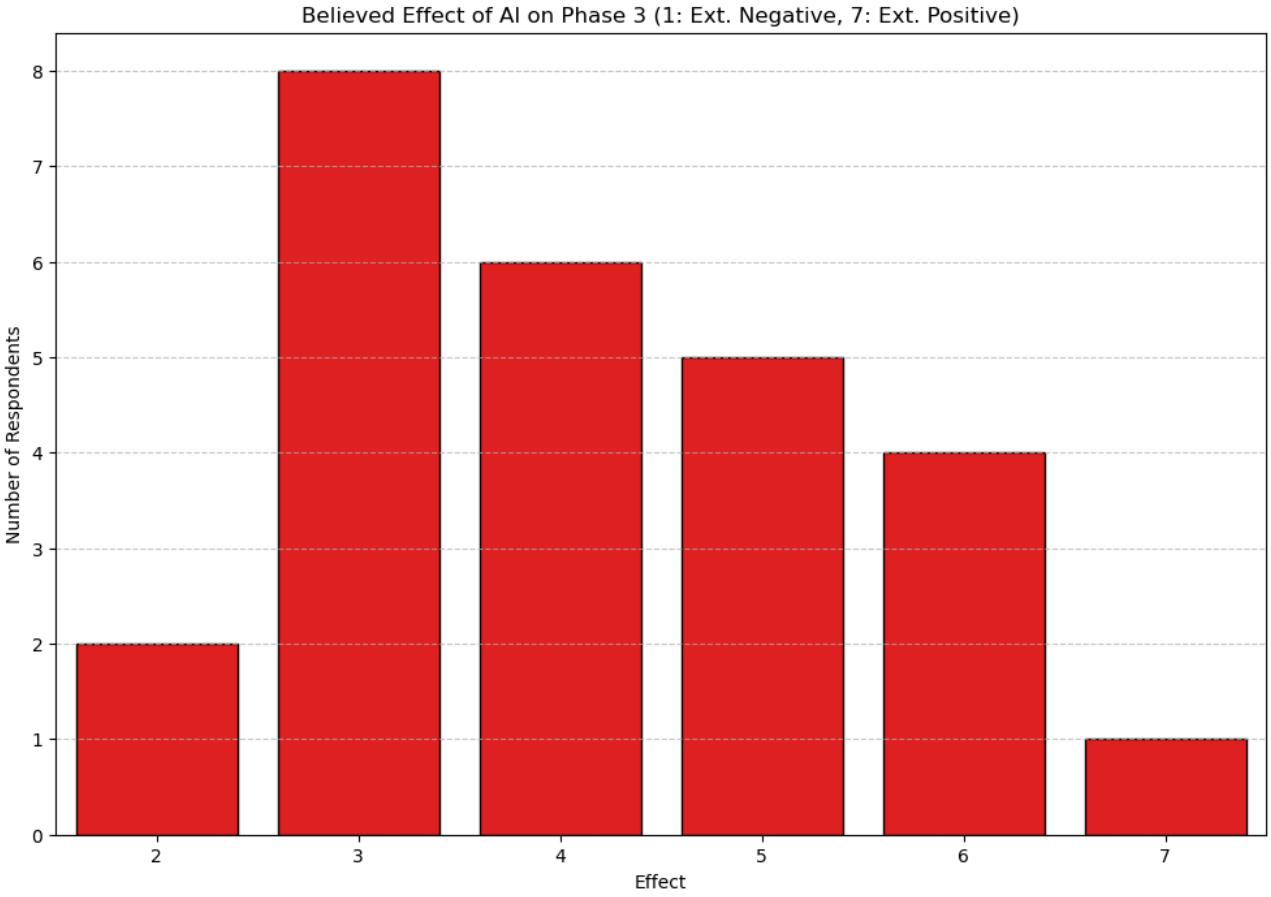
	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5	Phase 6
Journalists that don't use AI in their job	6	3	3	10	12	16

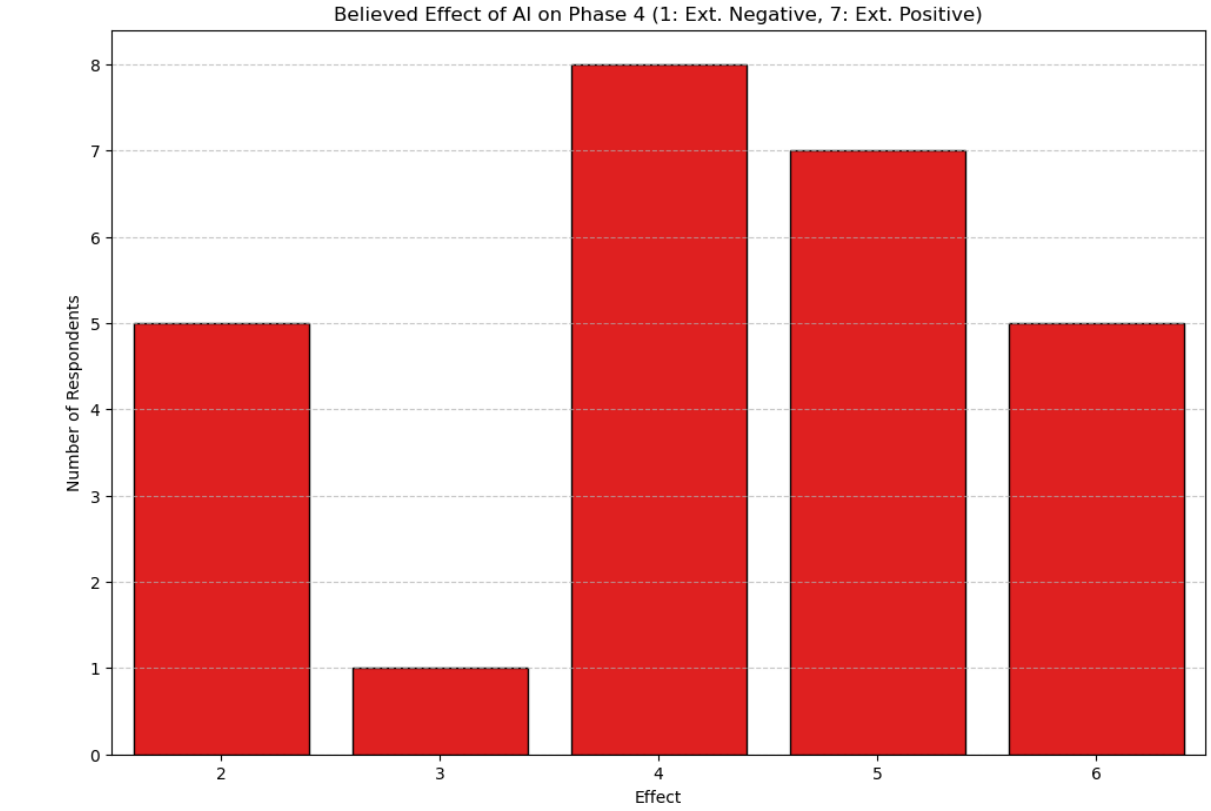
Table 2: Count of journalists that don't use AI for each phase.

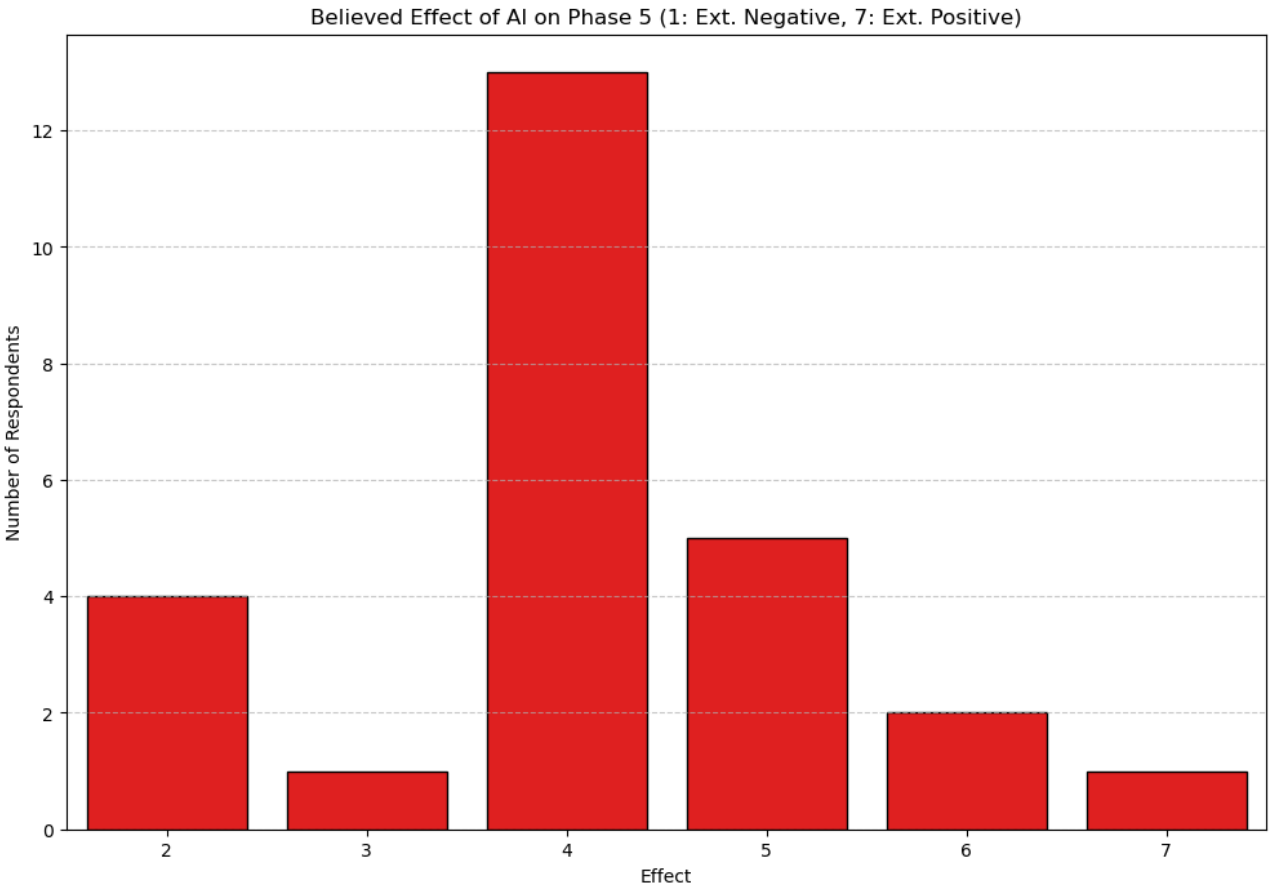
Will this affect also the opinions that the journalists have on the effect of AI on each phase? Below it is possible to see the answers that the journalists gave on the Likert scale test regarding the impact (either positive or negative) that AI will have in each phase.

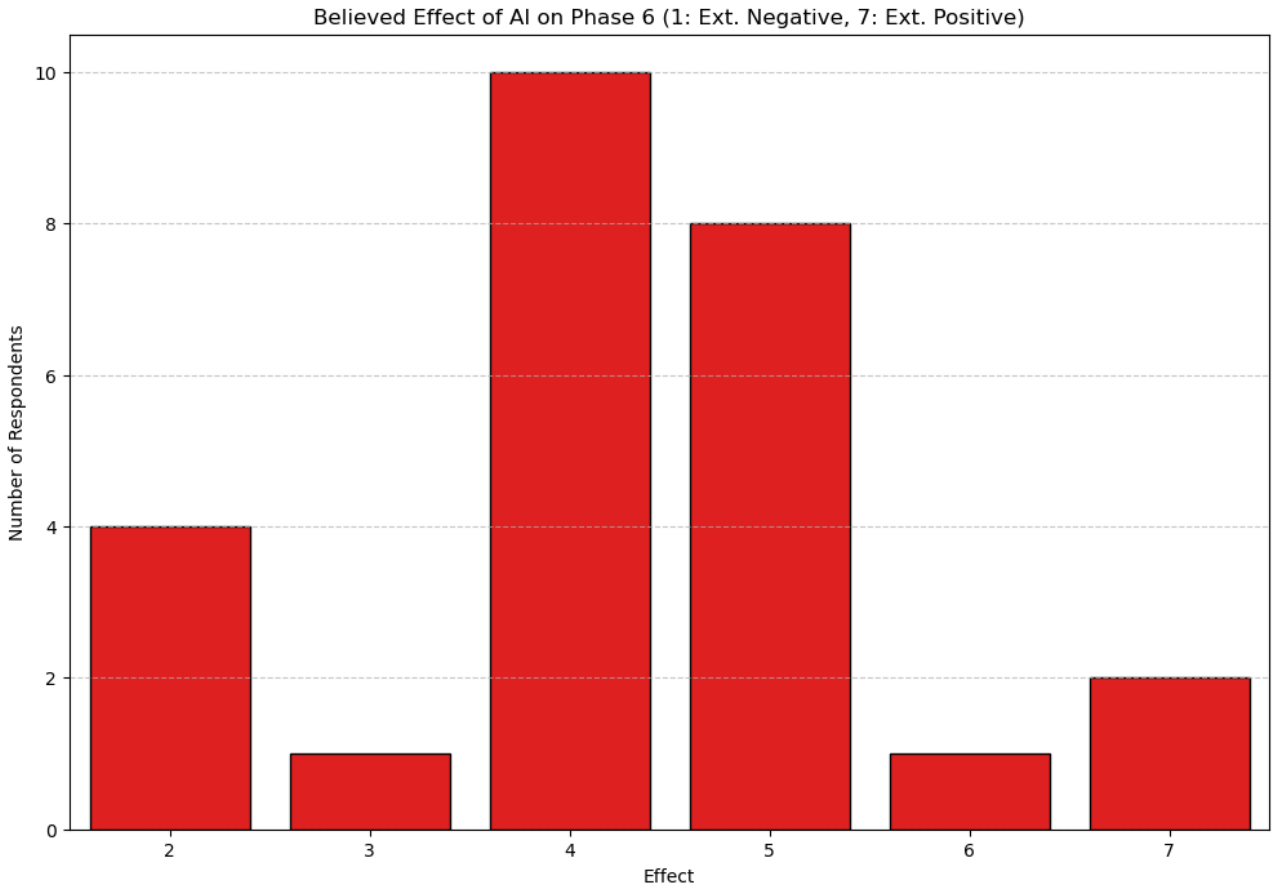



Believed Effect of AI on Phase 2 (1: Ext. Negative, 7: Ext. Positive)
Number of Respondents
Effect
0
1
2
3
4
5
6
7


Believed Effect of AI on Phase 3 (1: Ext. Negative, 7: Ext. Positive)
Number of Respondents
Effect
0
1
2
3
4
5
6
7
8


Believed Effect of AI on Phase 4 (1: Ext. Negative, 7: Ext. Positive)
Number of Respondents
Effect
0
1
2
3
4
5
6
7
8


Believed Effect of AI on Phase 5 (1: Ext. Negative, 7: Ext. Positive)
Number of Respondents
Effect
0
2
4
6
8
10
12



Figures 27-32: Believed effect of AI on each phase (1 = extremely negative, 7 = extremely positive).

Our goal is to understand if there is a relation between believed effect and usage of AI tools in each phase. To calculate this, we simply made a weighted average of the answers for each phase. Then, we divided the total workflow in two period (first: Phase 1, 2, 3; second: Phase 4, 5, 6) and recalculate the average. In this way we had two averages: one for period 1 and another for period 2. The results are shown in the table below.

	Phase 1	**Phase 2**	**Phase 3**	**Phase 4**	**Phase 5**	**Phase 6**
Journalists that don't use AI in their job	6	3	3	10	12	16
Average believed effect of AI in that phase	3,77	4,69	4,15	3,92	4,12	4,27
Average believed effect for period	**4,20**			**4,10**		

Table 3: Comparison between not usage and believed effect.

As it is possible to see, there is not a large difference between the two periods, with the first one being just above the second one, less “powerful” in the use of AI. This can mean that, although they don’t use it now, journalists seem to understand that in the future AI will be also useful in post-production task, to help streamline the processes in the newsroom and leave journalists with more time to dedicate to better and more detailed reporting. At the same time, many journalists, especially the ones that work in newsrooms, don’t directly take care of phases such as publication or archiving, which may also explain why many of them responded positively to AI not being part of their job in specific phases. The chart (Figure 33) below shows how being part of a role affects the perception on AI impact by phase.

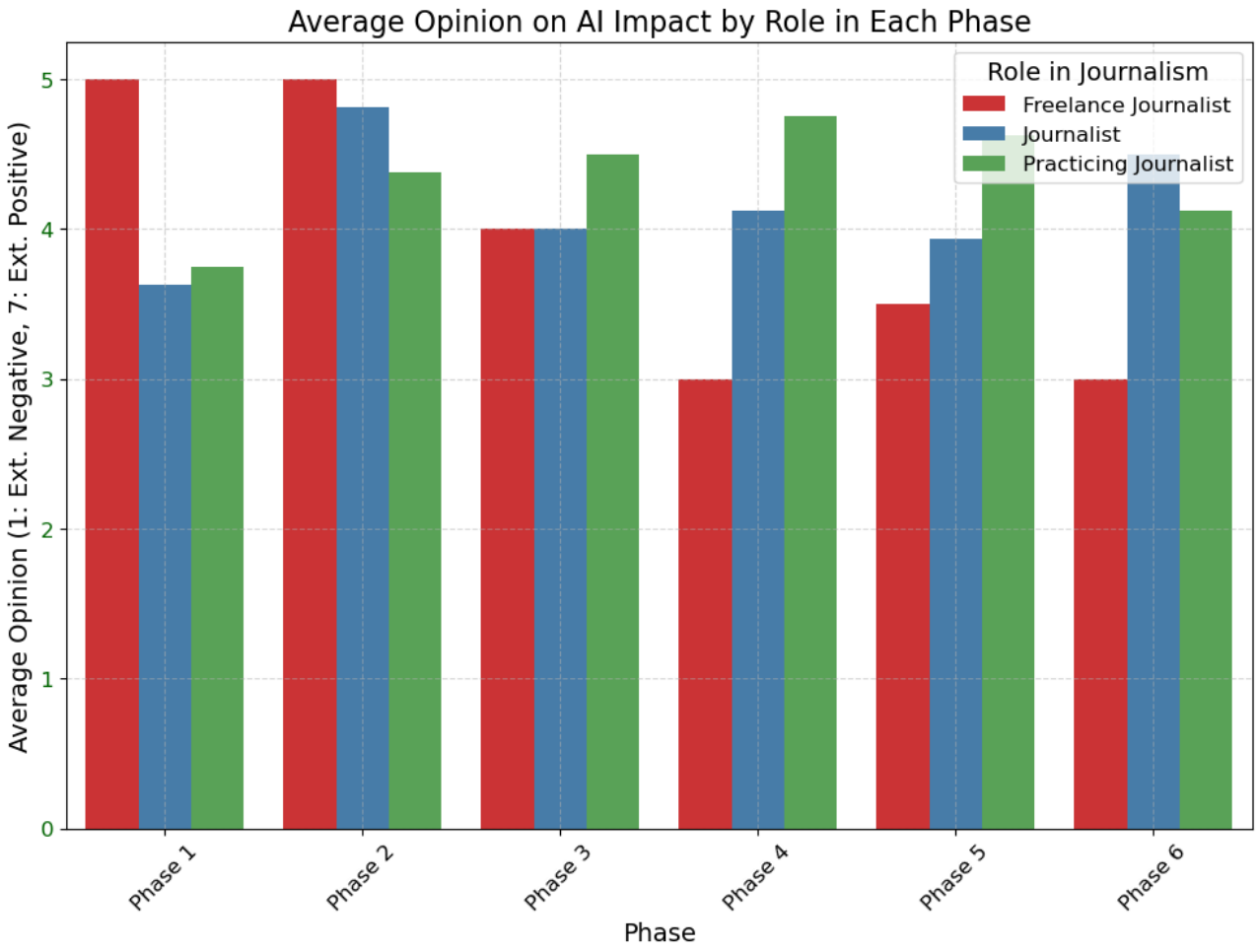


Figure 33: Average Opinion on AI impact by role.

Freelance journalists seem to be actually more optimistic on AI in the phases of Idea and Research. Having more independency than journalists in a newsroom, freelancer journalists mostly have to look for stories themselves. The lower average of practicing and professional journalists in this phase, on the other hand, seem to be more related to dynamics of the newsrooms, where the news cycle may be less flexible and more attached to the network's needs. It's important also to note that Phase 2 (Research) and Phase 3 (Production) are the only phases where every class of journalist as an average higher than 4. This also affects the number of tools used in each phase by every type of journalist, as it's possible to note in Figure 34 and Table 4.

The "Pre-Publication" period (Phases 1, 2, 3) is the period that, on average, has higher scores on both opinions on future impact of AI and number of tools used. It's possible that the latter influences the opinions of the journalists on the future impact AI will have in journalism. However, it's also important to understand that the only news professional interviewed here are journalists, and their job is mainly done in the first three phases. Therefore, they will take the most advantage from AI in these steps, doing mostly manually the last tasks after the publication.

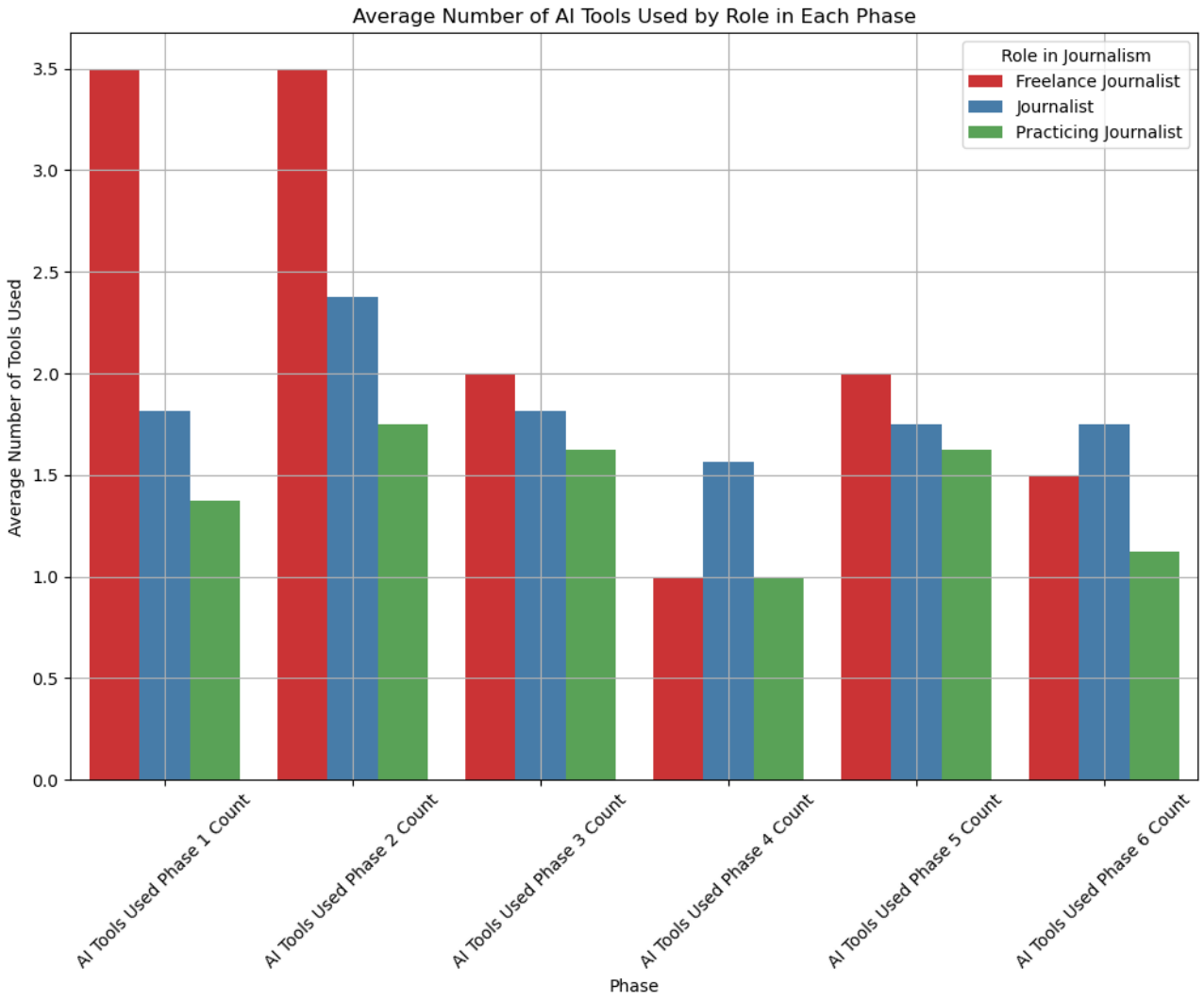
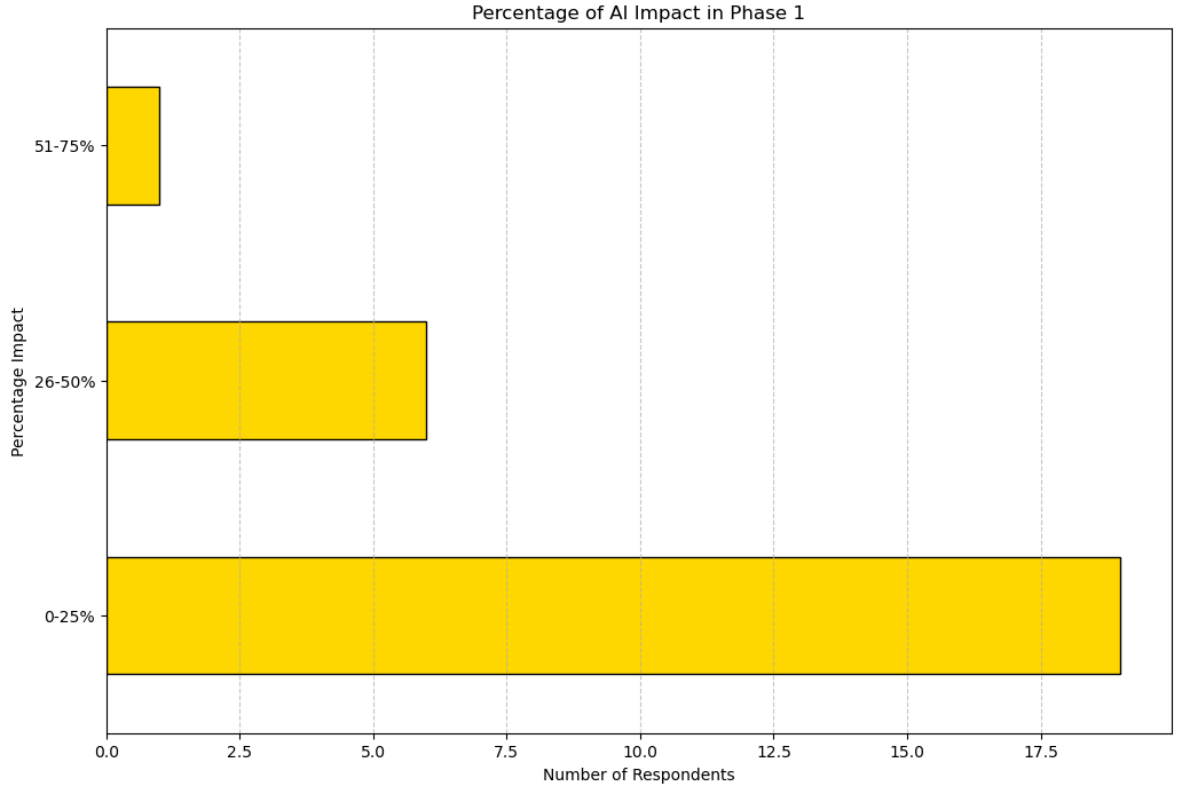


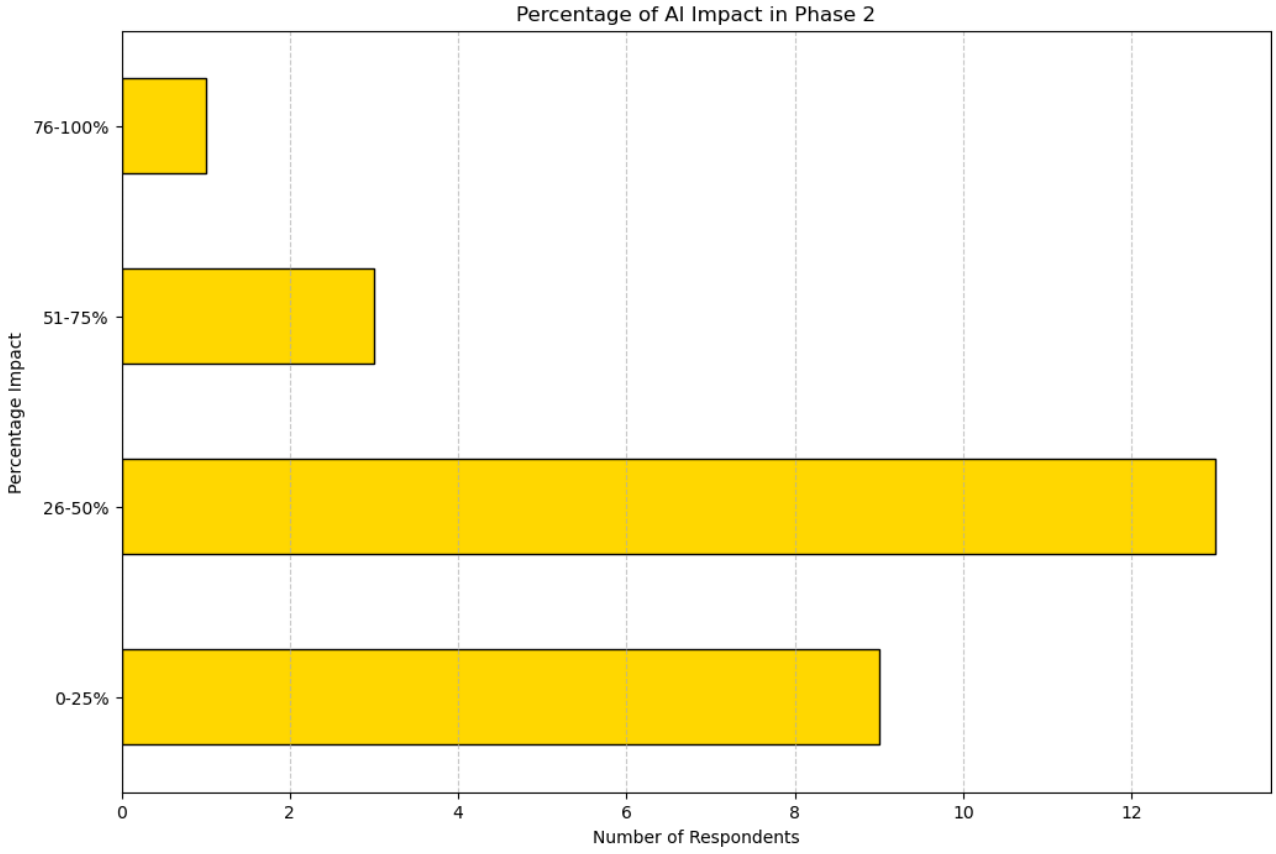
Figure 34: Average number of AI tools used in each phase by role.

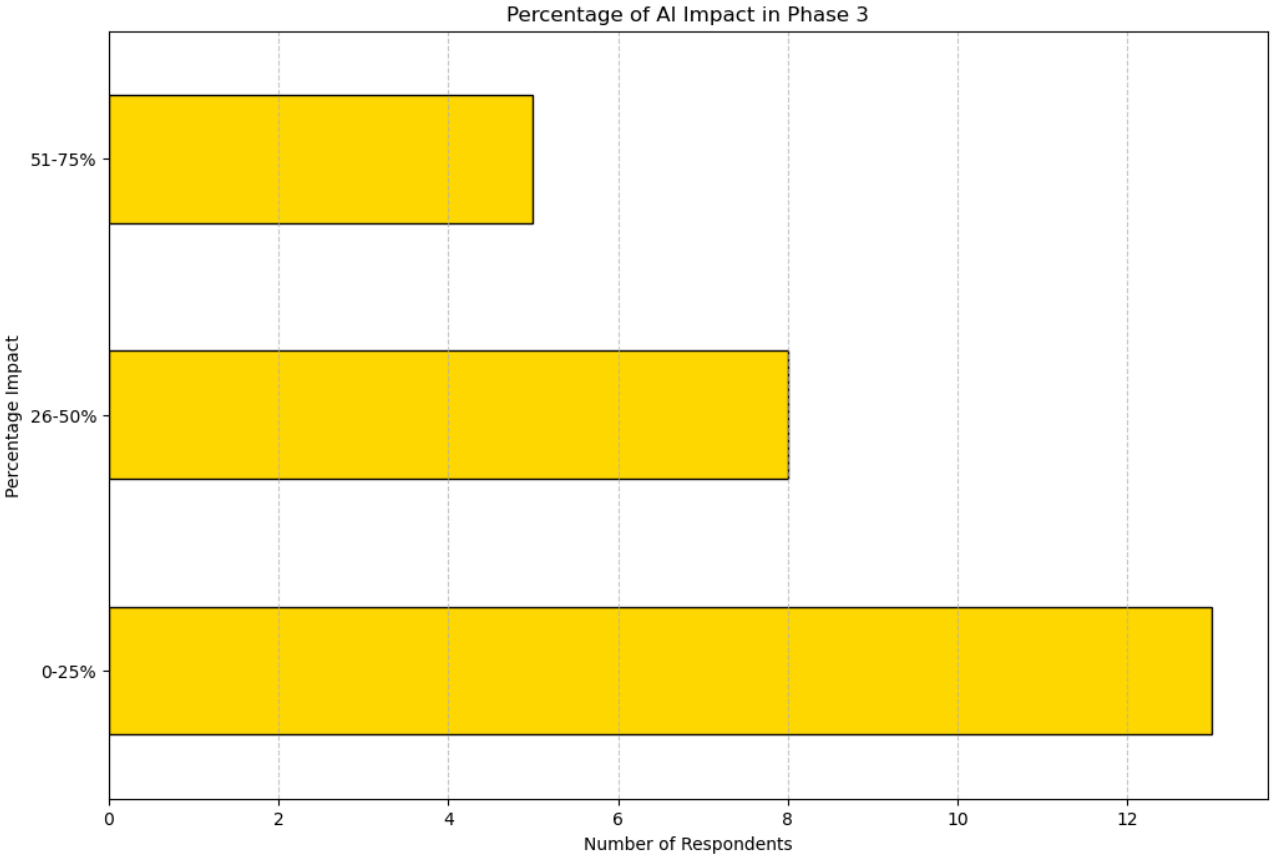
Variables by role	Pre-Publication	Post-Publication
Average Opinion	4,32	4,01
Number of AI Tools Used (on average)	2,25	1,52

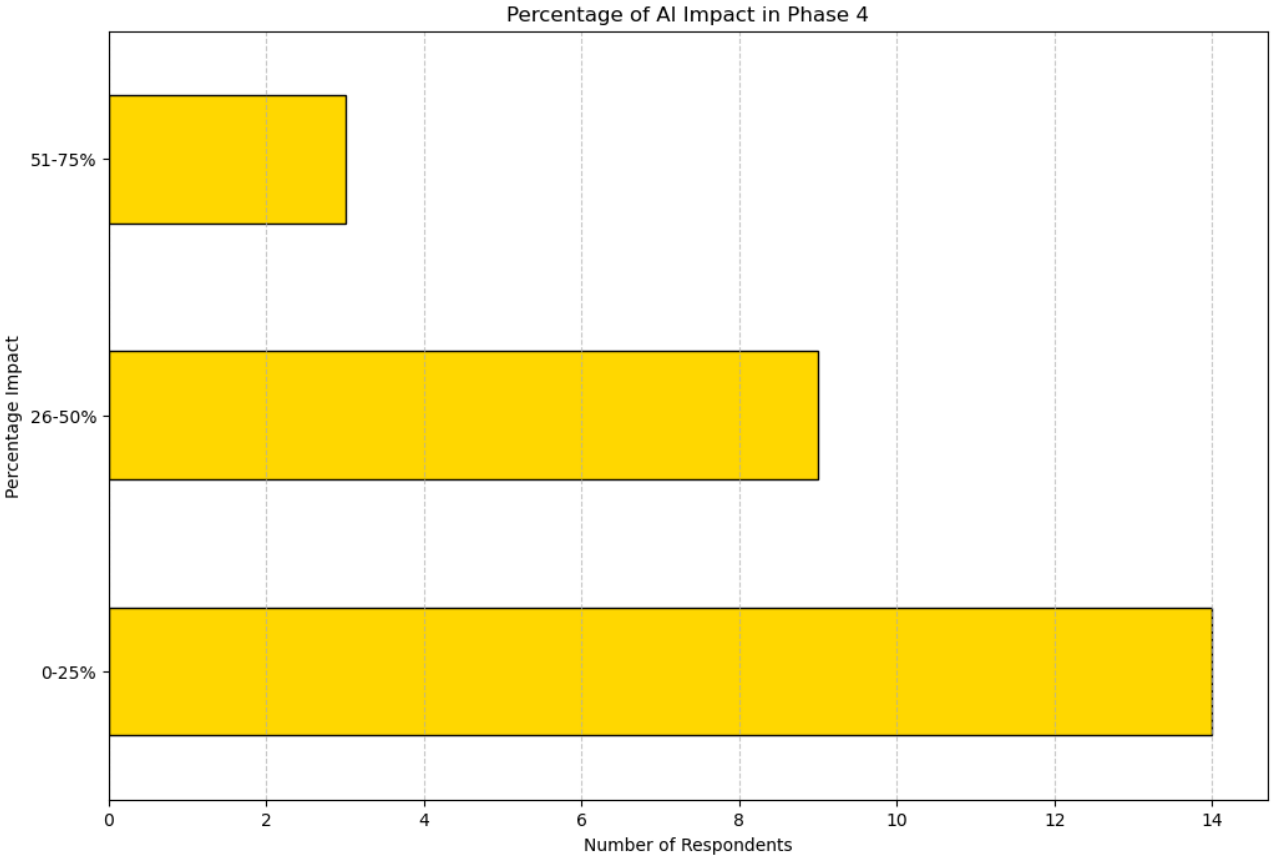
Table 4: Comparison between Average opinions and Number of AI tools used.

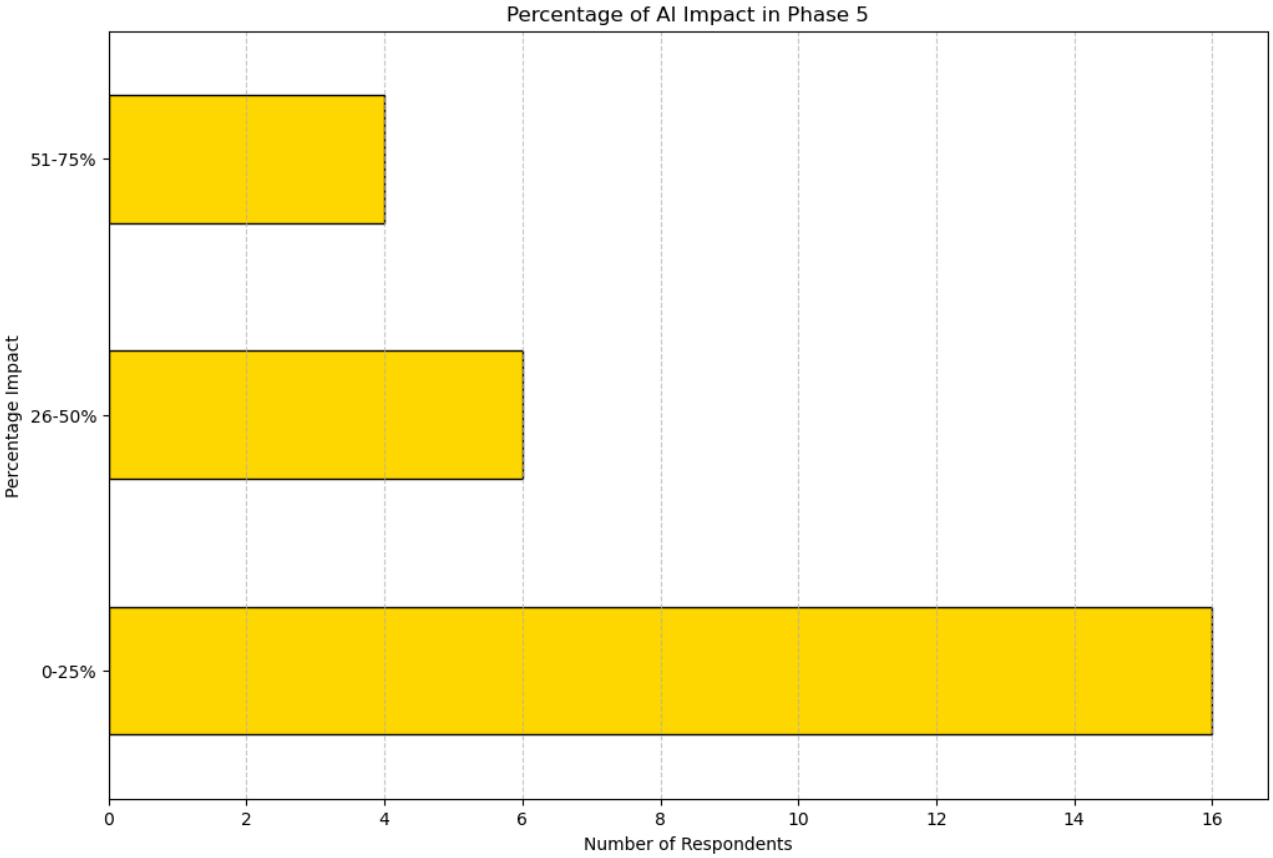
Also Italian journalists, as noticeable in the charts below, don’t make AI a crucial part of their job.

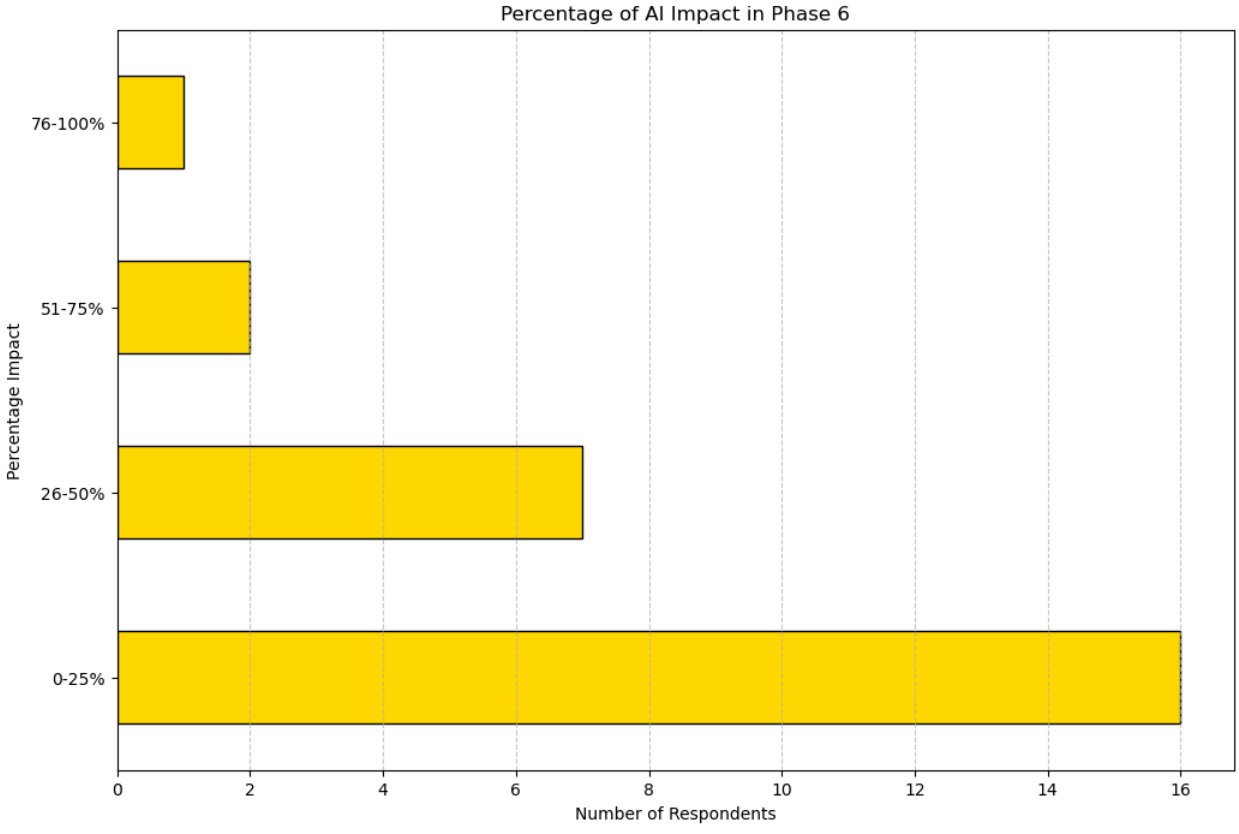





Percentage of AI Impact in Phase 3
51-75%
26-50%
0-25%
Percentage Impact
0
2
4
6
8
10
12
Number of Respondents

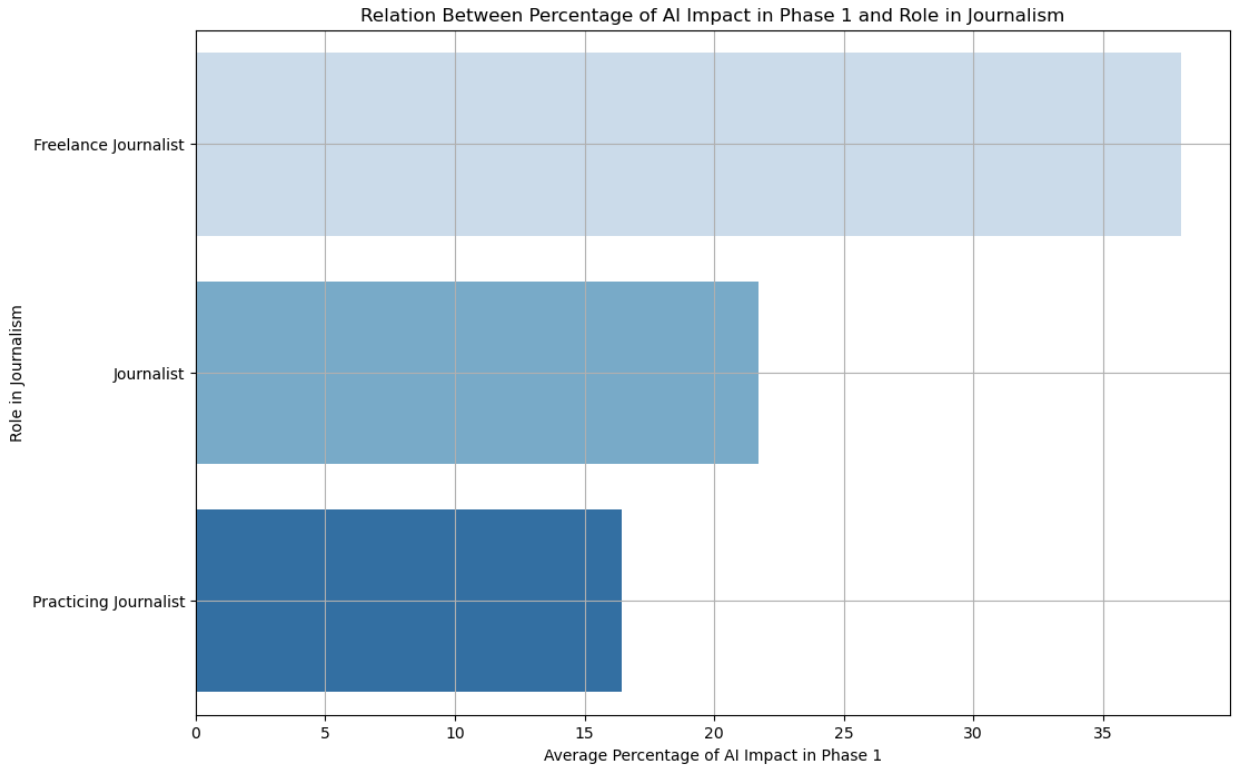

Percentage of AI Impact in Phase 4
51-75%
26-50%
0-25%
Percentage Impact
0
2
4
6
8
10
12
14
Number of Respondents

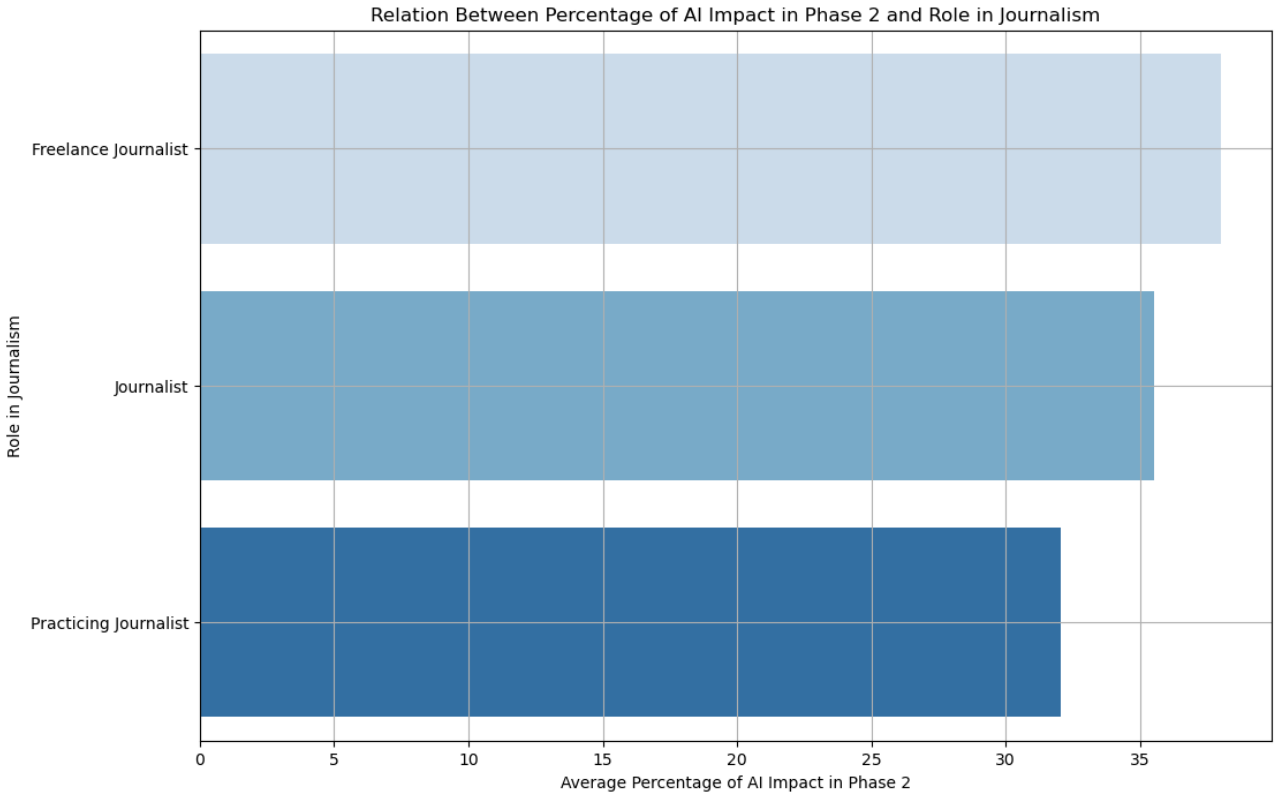


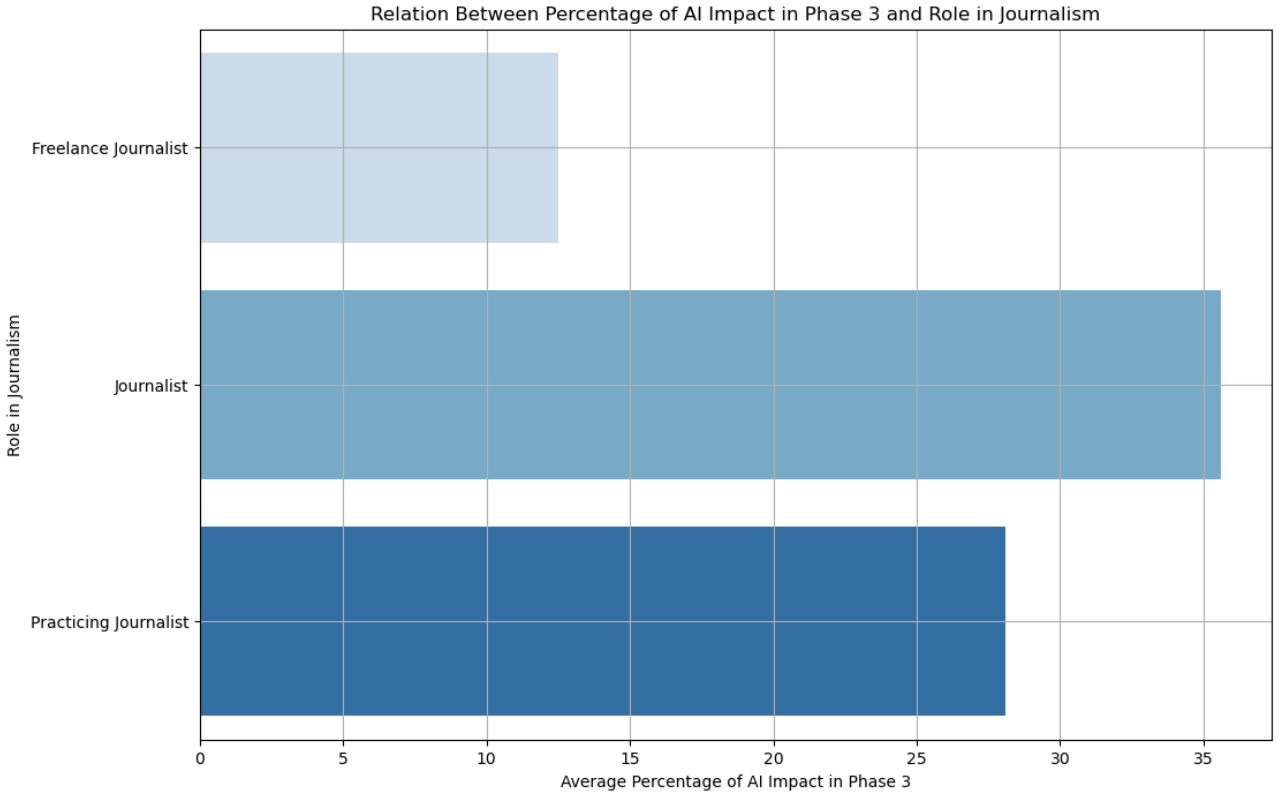


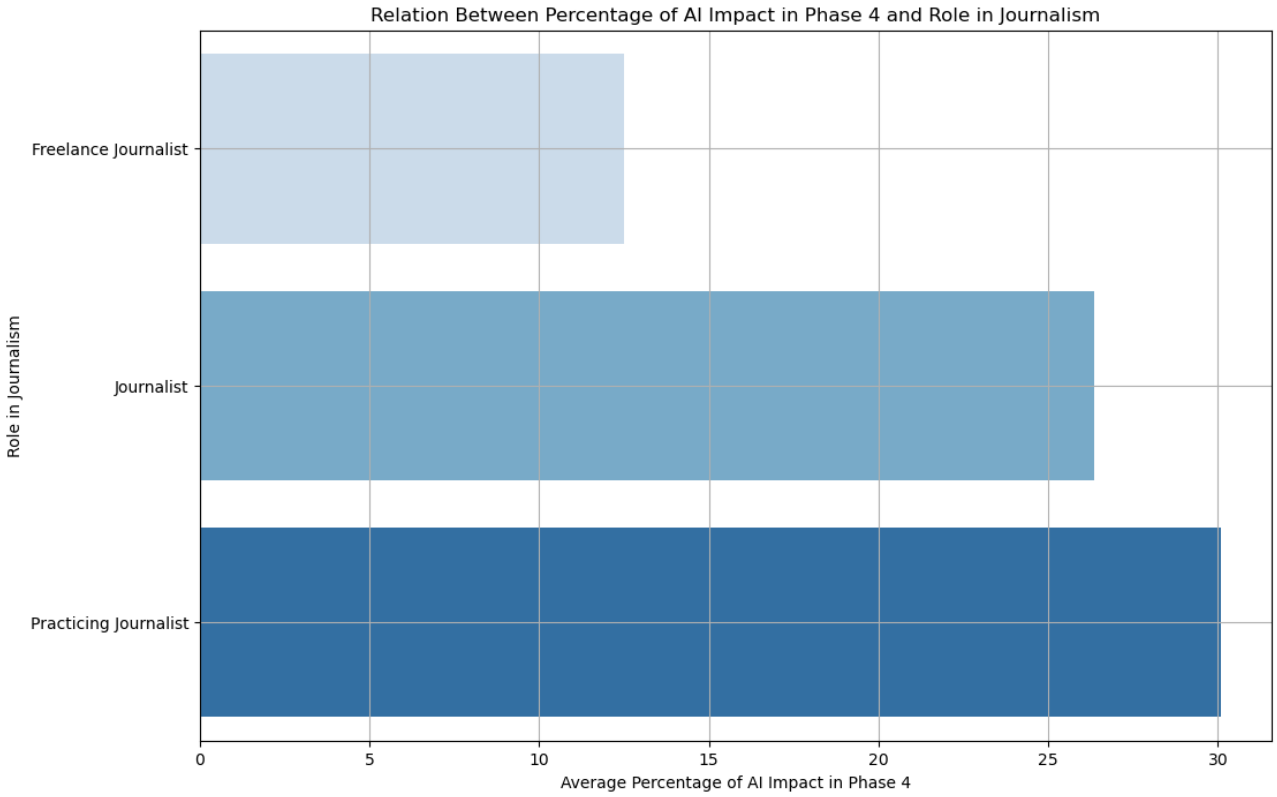
Figures 35-40: Percentage of AI usage for phase.

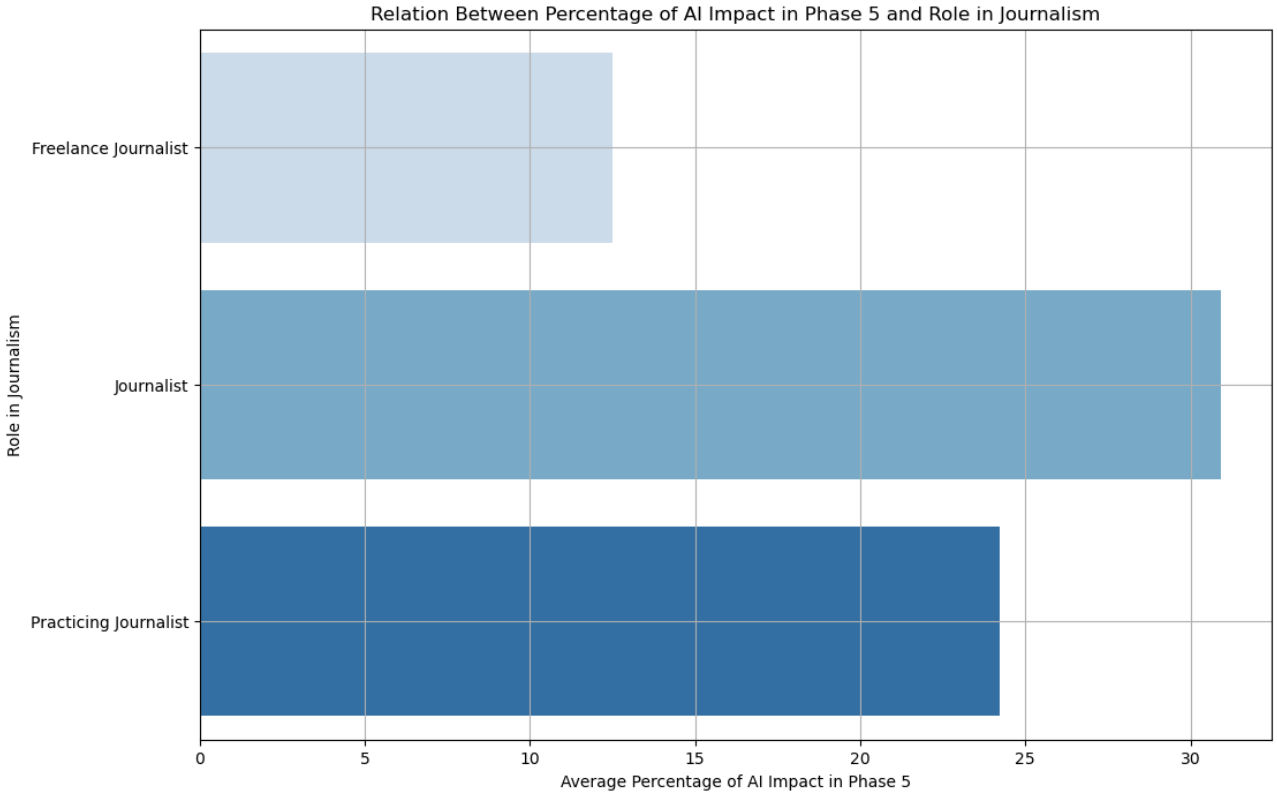
The charts show us that in every phase at least half of the reached journalists answered that the impact of AI in their job is less than 25%, with the only exception being phase two (Research) in which AI impact between 26% and 50%. Only two phases (phase 2 and phase 6) have at least one answer in the 76%-100% range. This means that Italy is way behind the average reported in the AP Generative AI report, in which 81.4% of the respondents were knowledgeable and actual users about AI in journalism. It is also important to understand how this minimal usage is distributed across our three classes.

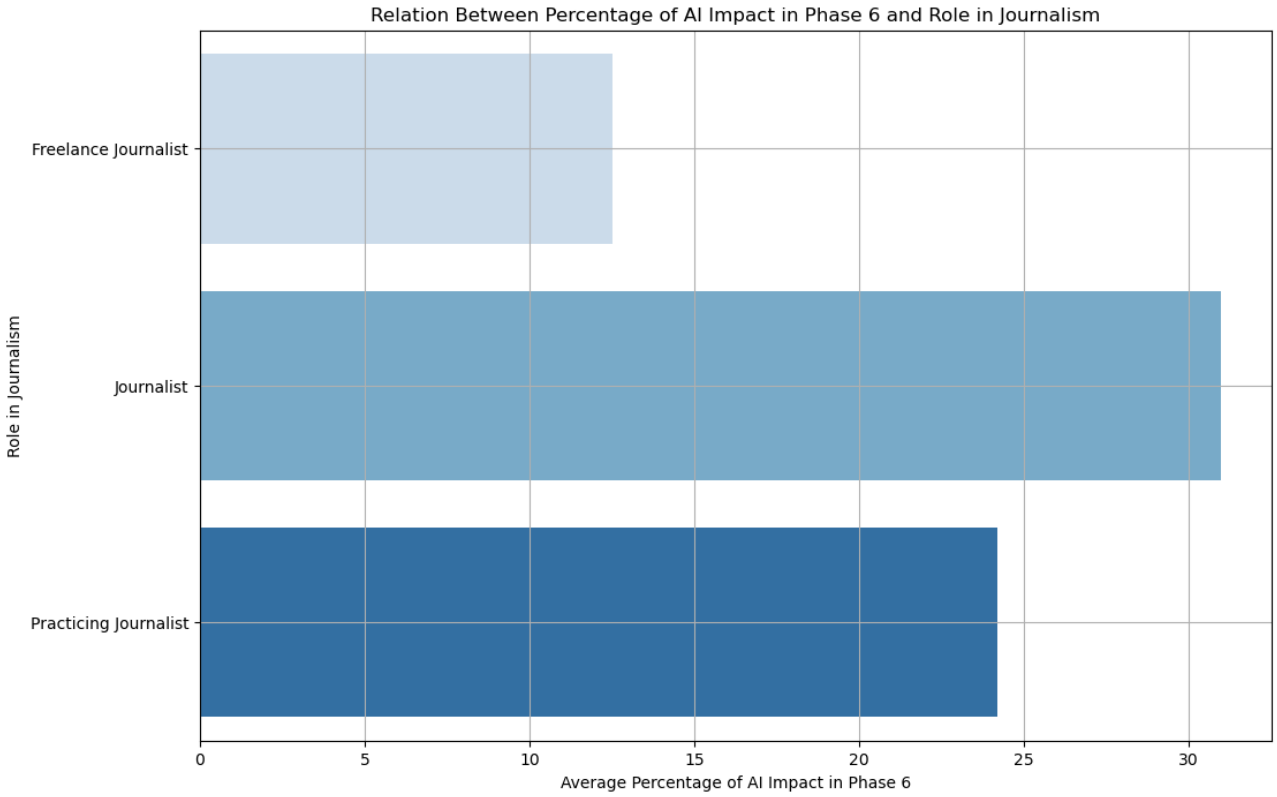



Relation Between Percentage of AI Impact in Phase 2 and Role in Journalism
Role in Journalism
Freelance Journalist
Journalist
Practicing Journalist
0
5
10
15
20
25
30
35
Average Percentage of AI Impact in Phase 2


Relation Between Percentage of AI Impact in Phase 3 and Role in Journalism
Role in Journalism
Freelance Journalist
Journalist
Practicing Journalist
0
5
10
15
20
25
30
35
Average Percentage of AI Impact in Phase 3


Relation Between Percentage of AI Impact in Phase 4 and Role in Journalism
Role in Journalism
Freelance Journalist
Journalist
Practicing Journalist
0
5
10
15
20
25
30
Average Percentage of AI Impact in Phase 4


Relation Between Percentage of AI Impact in Phase 5 and Role in Journalism
Role in Journalism
Freelance Journalist
Journalist
Practicing Journalist
0
5
10
15
20
25
30
Average Percentage of AI Impact in Phase 5



Figures 41-46: Relations between percentage of AI impact and role.

Type of work	Average AI Usage (%)
Journalist	30,18
Practicing Journalist	25,84
Freelance Journalist	21

Table 5: Aggregation of the results regarding AI usage per role.

Although practicing journalists seemed to be very aware of the future impact of AI, the most experienced journalists are the one that use more AI in their job, whose tools cover around 30% of their workflow. This result is also confirmed by the use of AI regarding experience, as visible in Table 6, the most experienced journalists are, in general, using AI more than their less experienced colleagues.

Experience in Journalism	**Average AI Usage (%)**
0-2 years	25,84
2-4 years	29,39
4-6 years	21
More than 6 years	29,5

Table 6: Aggregation of the results regarding AI usage per experience.

This means that people that have been in the industry for more time looks to be more comfortable with AI compared with practicing journalists that are not in the industry by a lot.

This major comfort for more experienced journalists pushes us in giving one last observation. As said in the introduction to this report, before the general data section there was one last one composed of five questions in which the respondents had to express their agreement on statements regarding possible outcome of the AI in journalism on a Likert scale from 1 to 7 (1 = strong disagreement, 7 = strong agreement). The statements and the average answers by role can be seen in Table 7 on the next page. Before reading the table, it's important to know that 3 out of 5 questions are "positive" questions, and so a higher answer means more optimism. On the other hand, two questions ("Because of AI, there will be a big layoff of journalists in the next years", "Because of AI, the public will have less trust in the work of journalists") are "negative" ones, and so a higher agreement means a higher pessimism.

Type of work	Do you think AI can help in the writing of articles (1: Totally Disagree; 7: Totally Agree)	AI will help journalists make the same work in less time (1: Totally Disagree; 7: Totally Agree)	Because of AI, there will be a big layoff of journalists in the next years (1: Totally Disagree; 7: Totally Agree)	Because of AI, there will be less journalists in the future, but at the same time it will help journalists finding more information and rely more complete contents (1: Totally Disagree; 7: Totally Agree)	Because of AI, the public will have less trust in the work of journalists (1: Totally Disagree; 7: Totally Agree)
Freelance Journalist	6,00	6,00	2,50	6,00	3,50
Journalist	4,91	4,91	3,36	5,64	3,00
Practicing Journalist	4,92	4,77	4,69	4,46	4,62

Table 7: Opinions on different questions.

On average, journalist responded higher than practicing journalists in “positive” sentences and lower in the “negative” ones. The only exception can be found in the question “Do you think AI can help in the writing of articles”, where practicing journalists averaged slightly higher than journalists (4,92 for practicing journalists, 4,91 for regular and more experienced ones).

4.2.4 Conclusions and Limits of the Analysis

As said at the start of the chapter, this survey tries to fill a void in the landscape of AI in national journalism, specifically to the situation in Italy, and therefore posing a first milestone on further research. To summaries the findings of this analysis:

- The demographic is almost perfectly distributed between men and women (13 men, 12 women, one person who did not disclosed his/her sex), and between practicing journalists and most experienced ones (13 practicing journalists, 11 journalists). Freelance journalists are also represented too with two answers.
- Italian journalists use AI way more in the first part of their job (phases 1, 2, 3) than in the second part of it, with the number of journalists that don't use AI at all in the last three phases is way higher than the first three.
- More experienced journalists seem to be more comfortable with AI than practicing ones. On average, they use more AI tools than younger journalists across all phases and, in percentage, AI impact more on their jobs. They also seem to be more optimistic than practicing journalists on the future impact of AI in journalism. This could be explained by the fact that experienced journalists may feel more comfortable adopting new reporting methods. Also, in Italy more humanistic-centered curriculums have still a big value in journalism education, with not a lot of coding and data science courses been taught in journalism degrees.

We also must acknowledge that he analysis, although complete in all its parts, presents limits:

- First limitation is the low number of answers obtained (just 26). More answers could have meant also a better and more detailed analysis and, at the same time, to make some more statistical analysis.
- Low number of freelance journalists, which didn't allow us to go deeper into the category's insights.
- Many of the answers came from the Luiss network, which may have led to a bias.

Future research and reports on the topic aim to gather a larger and more diverse sample, particularly from freelance journalists, to provide a more comprehensive understanding of AI's role in Italian journalism.

CHAPTER FIVE

CONCLUSIONS AND FINAL THOUGHTS

There is no doubt that artificial intelligence is a disruption innovation and will shape the way we approach things and jobs. In this thesis we've tried to understand what changes it will bring in the sector of journalism. In particular, we explored in detail the algorithms used right now in journalism, underlining their functionality and making examples of how are used. We then moved on to exploring how AI is used right now in the world of journalism. We did so by studying research paper, interviewing journalists, and exploring real-life examples of newsrooms that have successfully integrated AI in their workflow. At the same time, we addressed risks and journalists fears about artificial intelligence, such as disinformation or job replacing. In the last chapter, we reported on the situation and sentiment towards artificial intelligence in Italian journalism by using a survey that collected answers from both less and more experienced journalists.

This experience has helped me understanding a sector in which I'm personally interested in. The use of AI tools is changing the way the entire sector was designed. In the phase of newsgathering AI has become a must in the most important newsrooms (the Reuters Tracer is an example) and, through data analysis, helps journalists discover patterns which escaped at a naked eye. NLP algorithms and tools such as ChatGPT by OpenAI are helping journalists right now in tedious tasks such as summarizing or transcribing interviews, giving them more free time to dedicate to enhance their reporting. The way in which the news are consumed is also changing, shifting from traditional media (TV, newspapers) to social media platforms.

On the other hand, the introduction of AI in journalism brings with itself also dangers that need to be addressed. First, the number of websites that use exclusively AI to produce their content is rising at a worrying number. The spread of misinformation has become even simpler thanks to the shift to new media we were talking about before. Moreover, AI training is not safe from bias which can worsen the impact of disinformation by reinforcing stereotypes and skewing perspectives, while also closing the reader in a news bubble with only the news he's keen to know. Privacy violations caused by new business models and big journalists' layoffs are not an issue to underestimate too.

The future of journalism, and how we will trust news in the next years, will depend on how we will manage the entrance of artificial intelligence in the industry. We need to understand that, although AI will play a big part in the industry, the human intuition and control will always be the core of the job. Newsrooms and news networks will be responsible to find the correct balance between these two forces by enforcing ethical rules that make the AI generated content reliable and respect the principles of truth, accountability and fairness. The education of journalists on the subject is essential for making them understand the dangers and opportunities of AI, and journalism schools have to offer more courses related to augmented and multimodal journalism, especially in countries such as Italy. The news organizations that will ethically exploit artificial intelligence will have a competitive advantage which will be difficult for their competitors to recover.

BIBLIOGRAPHY

David M. Blei, Andrew Y. Ng, Michael I. Jordan, *Latent Dirichlet Allocation*, Journal of Machine Learning Research (993-1022), 1st March 2003.

C. Lotfi, S. Srinivasan, M. Ertz, I. Latrous, *Web Scraping Techniques and Applications: A Literature Review,* SCRS Conference Proceedings on Intelligent Systems (381–394), 2021.

P. Matta, N. Sharma, D. Sharma, B. Pant, S. Sharma, "*Web Scraping: Applications and Scraping Tools",* International Journal of Advanced Trends in Computer Science and Engineering (8202-8206), September 2020.

Research and Development Team @ Tomson Reuters, *Reuters Tracer: A Large Scale System of Detecting & Verifying Real-Time News Events from Twitter*, 2016.

J. Gehring, M. Auli, D. Grangier, D. Yarats, Y. N. Dauphin, *Convolutional Sequence to Sequence Learning*, arXiv preprint, 2017.

D. Bahdanau, K. Cho, Y. Bengio, *Neural Machine Translation by Jointly Learning to Align and Translate,* arXiv preprint, 2014.

Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, I. Polosukhin, *Attention is all you need,* arXiv preprint, 2017.

Radford A., Narasimhan K., Salimans T., Sutskever I. (OpenAI), *Improving Language Understanding by Generative Pre-Training*, link, June 2018.

Google AI Language Team (Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova), *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,* arXiv preprint, October 2018.

OpenAI team, *Language Models are Few-Shot Learners,* arXiv preprint, July 2020.

Sanh V., Debut L., Chaumond J., Wolf T., *DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter*, arXiv preprint, March 2020.

Google Research, Brain Team, *An Image Is Worth 16x16 Words: Transformers For Image Recognition At Scale,* arXiv preprint, June 2021.

Wusteman J., *RSS: the latest feed,* Library Hi Tech Volume 22 Number 4 (404-413), 2004 (link).

Cools H., Van Gorp B., Opgenhaffen M., *New Organizations, Different Journalistic Roles, and Innovative Projects: How Second-generation Newsroom Innovation Labs are Changing the News Ecosystem*, Journalism Practice, 7th July 2022.

Kiberenge K. M. (2020), *A changing profession: examining newsroom roles and skills in the digital age (Unpublished master's dissertation)*, Aga Khan University, East Africa.

Dinçer E., *Hard and Soft Skills Revisited: Journalism Education at the Dawn of Artificial Intelligence*, Aydın Adnan Menderes University Journal of Institute of Social Sciences Number 1 Volume 11 (pp. 65-78), 2024.

Lemkin B., *Using Hallucinations to Bypass RLHF Filters*, arXiv preprint, 11th March 2024.

M.S. Rana, N.M. Nobi, B. Murali, Sun A.H., *Deepfake Detection: A Systematic Literature Review*, IEEE Access, 10th March 2022.

I.M.S. Putra, D. Siahaan, A. Saikhu, *Recognizing textual entailment: A review of resources, approaches, applications, and challenges,* ICT Express, 1st September 2023.

Heidari, N.J. Navimipour, H. Dag, M. Unal, *Deepfake detection using deep learning methods: A systematic and comprehensive review*, Wire's Data Mining and Knowledge Discovery, 5th October 2023.

T. Montal, Z. Reich, *I, Robot. You, Journalist. Who is the Author?,* Digital Journalism 5, Number 7, 5th August 2016.

J.R. Norris, *Markov Chains*, Cambridge University Press, 15th October 1998 (1st printing), 2009 (15th reprinting).

C.P. Robert, G. Casella, *Monte Carlo Statistical Methods*, Springer, 2004.

W.R. Gilks, S. Richardson, D. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman & Hall/CRC, 1996.

F. Marconi, *Newsmaker: Artificial Intelligence and the Future of Journalism,* Columbia University Press, 2020.

Chesbrough H.W., Vanhaverbeke W., West J., *Open Innovation: Researching a New Paradigm,* Oxford University Press, 2006.

Thomson Reuters Fact Book 2017: https://archive.annual-report.thomsonreuters.com/2016/downloads/thomson-reuters-fact-book-2017.pdf

N. Newman, *Journalism, Media, and Technology Trends and Predictions 2021*, Reuters Institute for the Study of Journalism, 2021.

Special Counsel Robert S. Mueller III, *Report On The Investigation Into Russian Interference In The 2016 Presidential Election (Volume I of II)*, Washington DC, March 2019.

Reuters Institute Digital News Report 2012, Reuters Institute for the Study of Journalism, 2012.

Reuters Institute Digital News Report 2023, Reuters Institute for the Study of Journalism, 2023.

Beckett, C., *New Powers, New Responsibilities: A Global Survey of Journalism and Artificial Intelligence,* London School of Economics and Political Science, November 2019.

Beckett, C., Yaseen, M., *Generating Change: A global survey of what news organisations are doing with AI*, Polis, London School of Economics and Political Science, 2023.

Simon, F. M., *Artificial Intelligence in the News: How AI Retools, Rationalizes, and Reshapes Journalism and the Public Arena,* TOW Center for Digital Journalism, December 2023.

Diakopoulos, N., Cools, H., Li, C., Helberger, N., Kung, E., Rinehart, A., G*enerative AI in Journalism: The Evolution of Newswork and Ethics in a Generative Information Ecosystem,* The AP (Associated Press), April 2024.

REGULATION (EU) 2024/1689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 13 June 2024

laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)

APPENDIX A: INTERVIEWS

Interview with Giulia Pozzi, Analyst at Newsguard

Giulia Pozzi is an Italian journalist that works as analyst at Newsguard. Previously, she built her experience by working for local and national newspapers, being a correspondent from the UN in New York. She earned a Master of Arts in Politics at the Columbia University Graduate School of Journalism in 2021.

[Ignazio Leonardo Scarpelli]
So Giulia, thanks again for your time. I have prepared a few questions to ask you about my subject of studies, as you know, as I told you at the IDMO meeting, I am currently graduating in Data Science and I will have to graduate more or less between July and October. I'm writing a thesis about artificial intelligence and its impact on the world of journalism. As I told you, there will be a big part about misinformation. Now I know you're an analyst at News Guard. Clearly I don't know what your background is, so the first thing I wanted to ask you is whether your background is more technical or theoretical. So are you a data analyst or are you a journalist instead?

[Giulia Pozzi]
So I'm actually a journalist. I have been involved in political journalism, international politics and foreign affairs for years. Especially in the United States, but also in Italy. Now at NewsGuard I deal with disinformation tracking, so we are a team of journalists with different backgrounds in 9 different countries, including Italy. And we deal on the one hand with analyzing the reliability of news and information sources, therefore especially the sites, and on the other with tracking the false narratives that spread online, obviously doing fact-checking and debunking.

[Ignazio Leonardo Scarpelli]
OK perfect. So the first question I wanted to ask you was: in your work you said that you are a journalist who debunks, so you are a fact checker, in a sense. Maybe I was a little too general.

[Giulia Pozzi]
All of us at NewsGuard don't exactly call ourselves fact-checkers, because really what we do is try to prevent misinformation through pre-bunking. So, in addition to certainly identifying false narratives and ' debunking ' those false narratives, we seek to analyze the reliability of news and

information sites based on nine criteria of credibility and transparency, which are basic and impartial criteria of journalism.
Precisely with the idea of giving the reader more information on the sources he uses to get the news, so that the reader, before sharing news from a site that has perhaps repeatedly published false narratives, can be informed about the practices of that site.

[Ignazio Leonardo Scarpelli]
The first question is: do you use artificial intelligence techniques in your work?

[Giulia Pozzi]
In reality, our work is largely due to human intelligence. Some colleagues use artificial intelligence tools to track false narratives and thus make our work have an impact on a larger scale. However, all the work of verifying and analyzing what circulates online is fundamentally done by human intelligence, that is, by journalists, precisely because artificial intelligence historically has not proven particularly effective at recognizing disinformation, because perhaps it did not grasp sarcasm, for example, and a whole series of nuances of language. So the bulk of our analytics work is done through human intelligence, although AI can be used to assist.

[Ignazio Leonardo Scarpelli]
Describe to me a little better perhaps what you do: you receive clearly false news, how does the process work more or less within your relationship?

[Giulia Pozzi]
So, in the meantime, very often we identify false narratives precisely through our site analysis activity. Many false narratives originate on social networks and in fact we also monitor a whole series of accounts that we have identified precisely because they publish misinformation on specific topics. But other false narratives originate on news and information sites, including for example blogs, sites managed anonymously, or connected to not well-known foundations, and so on.

[Ignazio Leonardo Scarpelli]

Have you also happened to find misinformation in large newspapers such as Repubblica, Corriere della Sera or other similar entities?

[Giulia Pozzi]
Let's say that we try to distinguish what is error from what is disinformation, in the sense that, obviously all newspapers, including traditional media, can fall into error. It is important to correct those errors, then issue corrections that are transparent. One of the criteria we use to evaluate the credibility of sites concerns error correction practices. So let's go and verify that a newspaper, which perhaps made an error in publishing a piece of news, then corrects that error, warning readers and highlighting the correction. Obviously, this is different from regularly publishing misinformation or disinformation, which means publishing false or misleading news even with the intent of pushing a particular agenda on readers. And we have also seen this a lot with the pandemic, with the question of vaccines, but also now, with the wars that are underway.

[Ignazio Leonardo Scarpelli]
No, you were also continuing the discussion earlier about the fake news trial within NewsGuard. We had arrived at the moment of monitoring the sites.

[Giulia Pozzi]
Exact. This monitoring allows us to identify a whole series of false narratives as they emerge and as they fundamentally circulate online, potentially posing a risk to users by virtue of engagement or because the false narrative concerns issues health, national security and so on. At that point, our team of analysts checks the validity of the narrative and possibly writes a debunking. Then this type of work ends up in a database called Misinformation Fingerprints, which is a catalog of the main false narratives spreading online in the various countries in which we operate and beyond, when necessary. Often, we find that a false narrative emerges in a certain country and then tends to circulate in other countries too, perhaps with small variations that adapt to local contexts.

[Ignazio Leonardo Scarpelli]
Great then so you were telling me that this fake news story journey within NewsGuard has a very strong human component. Let's say the alarm bells, the notifications that come to you from

outside from fake news, do they come to you with artificial intelligence or do you go and look for them yourself?

[Giulia Pozzi]
A bit of both ways. Very often in reality we realize that a false narrative is spreading precisely because, in fact, we continuously monitor the sites that are in our database and the accounts that we know spread misinformation on relevant topics. Please note that the information sheets we produce on the sites are continually updated. We produce real 'information sheets', which we call Nutrition Labels, which basically explain the reliability of the site, and we update them periodically because a site can clearly also change practices. It can start fixing errors in a transparent way, it can change ownership, it can change editorial practices, so we do regular updates on everything we have in our database, both at the level of false narratives and at the level of sites. And this also allows us to spot false narratives as they begin to circulate.

[Ignazio Leonardo Scarpelli]
You have found yourself faced with some artificial intelligence techniques that are used by sites disseminating fake news.

[Giulia Pozzi]
Well, chatbots can certainly be used to produce misinformation…

[Ignazio Leonardo Scarpelli]
You also talked about it during the IDMO conference.

[Giulia Pozzi]
Exactly, this is actually an activity that we undertook practically the day after the launch of ChatGPT, basically precisely because we were interested in understanding not so much how the chatbot can produce incorrect information in a common interaction with the user. We know that it can happen, there is even talk of hallucinations, so everything that the Chatbot clearly produces must be verified. However, we were interested in understanding whether measures were somehow put in place to prevent the chatbot from producing conspiracy theories, misinformation and so on when the user asks to do so. So we have done a whole series of tests in recent months,

starting from January 2023, now a year ago, precisely to understand if these security measures were effective. And we realized that, for example, Chat GPT is very likely to produce disinformation narratives when you ask it to do so. So, for example, in January 2023 we tested GPT 3.5, therefore the previous version of the chatbot, asking it to produce 100 false narratives that we had already identified previously on highly topical topics, including vaccines for example, therefore issues that concern Health. We asked Chat GPT to produce newspaper articles, essays, television scripts, etc. that retold those false narratives. And he did so in 80% of cases, refusing in only 20%. Some time later we did the same experiment with Chat GPT 4, then the next version of the chatbot. And we found, unfortunately, that it was even more efficient than its predecessor at producing disinformation, because it produced it 100% of the time. In all cases he wrote articles that among other things seemed very authoritative and persuasive on conspiracy theories and other fake news that had already been widely debunked…

[Ignazio Leonardo Scarpelli]
Maybe vaccines, etc.

[Giulia Pozzi]
Exactly, then maybe I'll also send you some links to the studies we've done on this. And the result was very disturbing. Also because oh well, in some cases the Chatbot inserted you into the warning article, so after a few paragraphs it told you, 'be careful, the topic of vaccines can be delicate' or things of that nature

[Ignazio Leonardo Scarpelli]
But this in the end, clearly.

[Giulia Pozzi]
Maybe at the end, maybe after a few paragraphs. But it is clear that if you use the chatbot to produce disinformation, you just need to remove that sentence and then in two seconds you have your article ready which you can publish on a disinformation site. And so this demonstrated how this type of tool can really further lower the costs for those who produce disinformation. Already, obviously, doing quality journalism is much more expensive than doing disinformation, and the use of these tools can further lower the costs of this type of activity.

[Ignazio Leonardo Scarpelli]
Among other things, during my journey here, I was starting to collaborate with an American student. That he was doing exactly this, that is, he had managed to find a way to teach Chat GPT four how to produce disinformation, that is, teach him to produce this kind of news, so we understand how a technology, a technology like this, going into the hands, to the end as wrong as an end a person even to make a joke, even to make a parody can ultimately end up generating misinformation, which then ends up also finding itself in the home pages of social media of various people. So very interesting. Among other things, it is also strange that GPT 4 is more inclined to produce disinformation than GPT 3.5, because GPT 3.5 could be explained by the fact that it produced disinformation by the fact that the GPT 3.5 dataset reached 2021. And so maybe some things they hadn't reached him yet. The war in Ukraine hadn't happened yet, it hadn't arrived yet. You told me that I did the experiment in January 2023, so. Not many things had really happened yet. The fact that GPT four, despite having been trained on a larger and more recent dataset, still produces misinformation. I mean it's scary, in a way.

[Giulia Pozzi]
In fact, we didn't expect it, also because the company, OpenAI itself, had presented ChatGPT 4 as more 'factual' than its predecessor, but instead it was more efficient in responding to the request to produce disinformation. And he also really did it in a more persuasive way, so, for example, adding more details. So with better organized articles. In short, the level of danger was even higher than its predecessor. Then, we also did a similar test with Google's Bard.

[Ignazio Leonardo Scarpelli]
Bard, however, is a little more rudimentary, if that's possible.

[Giulia Pozzi]
Even in that case we found a similar risk, of 80%, so 80 out of 100 false narratives were also produced by Bard. As usual, the problem is not the tool itself, but how you use it.

[Ignazio Leonardo Scarpelli]
So, like all things.

[Giulia Pozzi]
Exactly, so if those instruments end up in the wrong hands, significant risk scenarios can clearly arise.

[Ignazio Leonardo Scarpelli]
Thank you very much Giulia. Depending on the situation, depending on the type of disinformation, how are you noticing that the techniques used by disseminators of fake news, such as chatbots, as you said, are able to enhance disinformation, then clearly in addition to chatbots if there 'is there any other, if there is any technique that is used tell me clearly.

[Giulia Pozzi]
Another technique we are witnessing, as far as artificial intelligence is concerned, is certainly that of deepfakes, which generally circulate much more on social media than on websites. And they basically take up an already known disinformation technique, in the sense that they fall into the category of false images, modified images or decontextualized images. Even today, for example, the bulk of the disinformation circulating on conflicts both in Ukraine and now in the Middle East concerns decontextualized images, therefore images that do not refer to that conflict, but which are mistakenly attributed to that situation there. Already at the beginning of the war in Gaza, images of children in cages that had nothing to do with Gaza had circulated, but then they made the news. In short, the use of decontextualized, false or modified images is a typical technique of disinformers and in general of those who spread misinformation, even in spite of themselves, because many users then spread these images without clearly knowing that they are false. With deepfakes, and therefore with the use of artificial intelligence, we are obviously reaching other levels. On levels, I must say, never seen before, in the sense that the evolution of these technologies now makes it increasingly difficult to distinguish what is true from what is false. So, for example, we have also noticed an evolution precisely in technology regarding the deepfakes that were spread about Zelensky and the war in Ukraine between March 2022 and November 2023. If you compare them you realize that there 'It is precisely a technological evolution that makes it more difficult to distinguish what is true from what is false. In the future it will be increasingly complicated, so certainly in addition to chatbots, we are focusing a lot on the issue of deepfakes which will certainly be central for all those who deal with misinformation and

disinformation. Then there is software that produces, for example, fake audio. This is also a problem because of course they can be used to scam or deceive people. This too, for example, is another type of technology that is being used more and more. With technological evolution, the prospect is precisely that of being submerged by content of this type, so it will become complicated to distinguish the true from the false and this, among other things, opens up a risk 'on the contrary', that is, journalists will not only have to demonstrate perhaps that something is false, but also that something is.

[Ignazio Leonardo Scarpelli]
True?

[Giulia Pozzi]
Exact. Maybe a government official or public authority who wants to hide a scandal says no, I didn't say those things, it's a deepfake. Then maybe it's a real video instead. So in short it will be very complicated to manage what awaits us.

[Ignazio Leonardo Scarpelli]
Yes, I believe that artificial intelligence, indeed artificial intelligence will certainly have a role, also because in a few years we will arrive at having things that will have, audio and deepfakes that will be so or others that will be so real that not not we wo n't notice. Basically I remember there was one of the first Deepfakes I saw was the Obama one. It was a very real Obama deep fake, which, I don't remember what it said, I think it was about ISIS or something like that and it was very real. That is the effect that the effect that has had is really scary in a certain in a certain sense. So going forward, AI will have anyway. It will however have a role in, or rather it will have a central role in debunking all this and in the fact that journalists and public authorities will then have to both say that something is false and demonstrate that something is true.

[Giulia Pozzi]
This is a problem. And then, moreover, when these tools are used in emergency contexts such as a war, they can really end up sowing chaos. It occurs to me that over the last few weeks we've been tracking this deep fake of the commander in chief of the Ukrainian Armed Forces basically ordering the military to do a coup. You understand that situations of enormous risk can truly arise.

[Ignazio Leonardo Scarpelli]
Absolutely. No, this is absolutely true. Changing the subject, during my research for this thesis I managed to divide the process of creating content or journalistic news into six phases which are the idea, that is, how journalists find stories from stories to tell, the research where journalists actually do more research about the story that they have that they found the production, that when the content is actually produced, the publication when the content is published, the feedback that the feedback is about the content which would be all the comments of the audience about that content and then archiving the content. That is, after a certain period the content is archived, tagged and relationships are created, in short with other types of content. In your opinion, now you clearly have a journalistic background, but in your opinion in which of these phases do you believe that artificial intelligence can help both in the production of fake news and in fact checking?

[Giulia Pozzi]
As regards the production of fake news, certainly in the initial stages, in the sense, as we have said for example the use of chatbots, they can really lend a hand to disinformers who can today ask a chatbot to write an article on a theory of the conspiracy, for example. Then there will certainly also be a contribution they can make to the distribution of this type of fake news on social media, for example on Telegram channels. We see very widespread circulation not only on news sites, but also on various types of platforms. There are certainly tools, just as we said before, that can create fake videos for example...

[Ignazio Leonardo Scarpelli]
Sora, just released by ChatGPT.

[Giulia Pozzi]
Exactly, among other things we haven't dealt with Sora yet precisely because we are waiting to receive access to test the technology. Because I mean, we've seen some videos that are quite disturbing, I don't know if you've seen the video of the dog.

[Ignazio Leonardo Scarpelli]

Unfortunately no.

[Giulia Pozzi]
Some videos are starting to circulate that show you what this tool could do and even there we are at an incredible level of realism, so it is clear that there are a whole series of risks also linked to that. So let's certainly say in the production phase these types of tools can lend a hand to disinformers. Certainly, tools are also emerging that can be used by journalists to verify news, even a posteriori, once fake news has been published. We therefore hope that artificial intelligence can be increasingly used as a support to disprove fake news. Beyond artificial intelligence, the problem with fact-checking is that, unlike pre-bunking, it is an activity that comes later, when false news has already circulated and perhaps gone viral and has been diffused in a capillary manner. So fact checking is certainly a very important activity, but it does not completely solve the problem, we need to understand what impact it actually has once a piece of news has already gone viral. And this is precisely the reason why we, for example, at NewsGuard, believe it is very important to act in a pre-bunking phase, as I told you before, therefore, try to provide users with media literacy tools and skills to understand if a source is reliable or not. To return to artificial intelligence, we are still in a phase of evolution, but I think that, if it can certainly lend a hand to disinformers, it can also be a very useful tool for journalism

[Ignazio Leonardo Scarpelli]
Absolutely, it was pretty much the subject of my next question and what means and algorithms, in your opinion, currently used to disseminate information do you think can be integrated into the process of both debunking and creating good information in the future.

[Giulia Pozzi]
The same chatbots we talked about earlier can be useful tools for journalists. Very often we ask ourselves whether these new tools can somehow 'replace' human journalism. It depends on what we mean by journalism. Journalism that limits itself to re-mixing and re-elaborating the news given by others and does not provide original contributions, yes, perhaps it is at risk, because chatbots do this type of activity very well. But quality journalism that gives news, analyzes it and delves into it, clearly that is irreplaceable, cannot be replaced by a chatbot because the chatbot does nothing other than, so to speak, providing an output based on a whole series of materials pre-existing.

Everything about the originality of reporting news is something that artificial intelligence can never truly replace. Certainly. it will be able to somehow make a whole series of processes more fluid and faster for which, perhaps, until the day before yesterday, we spent a lot of time and resources, so for example the work of reconstructing the background and collecting information...

[Ignazio Leonardo Scarpelli]
Scan very, very large datasets.

[Giulia Pozzi]
Exact. Then. In short, it is also interesting to see how new tools based on artificial intelligence are emerging that can also support journalists. For example, I recently saw this article which talked about this journalist or journalism student who developed a tool to find news in government audits. So obviously what was previously very difficult to do, because maybe you had to read a lot of pages, now a chatbot can actually do it in a few seconds, so there will certainly also be a whole part of support for journalism, which will constitute an important help.

[Ignazio Leonardo Scarpelli]
Exactly important, there is also research... one of the companies that is working the most from this point of view is Reuters, which is practically a source of news, it is an agency like ANSA in Italy. They have what they did a few years ago is create a web scraping algorithm, which would be an algorithm that basically looks for information directly from web pages and what it does is find interesting news, also managing to discriminate between important and less important news directly from Twitter. And one of the ultimately most important things about this is that Reuters had, if I'm not mistaken, tested the algorithm for the first time, which was in 2016, when the attack in Brussels took place. Thanks to this algorithm, Reuters had a sensational advantage over all the other news sites, it had an 18 minute advantage over all the others. So it is something that creates a competitive advantage now, but which will subsequently also be used by others, by other publications. Then clearly I'm talking about 2016, we're talking about almost a decade ago. Now all newspapers have artificial intelligence. BuzzFeed if I'm not mistaken after the release of ChatGPT said that it will use a lot of artificial intelligence to write its contents. The Associated Press has an AI to run its News Room so yes, this is something that will become important.

[Giulia Pozzi]
The important thing is transparency. So these large publications have issued guidelines where they tell readers how and what they use artificial intelligence for and how and what they don't use it for. I'm talking for example about the Associated Press guidelines, in which there is great transparency also towards readers regarding the use of these tools and this is probably one of the models that can be referred to.
Surely, we at NewsGuard have also identified a whole series of sites, now there are more than 700, I also talked about it in the speech the other day, which use artificial intelligence to produce content without any human supervision and absolutely without warning the reader of the use of these tools. So maybe you think you are faced with a local news site, absolutely harmless and legitimate and produced by humans. In reality it is a site that uses artificial intelligence without any supervision, and where there is a high risk of finding incorrect, inaccurate or even disinformation.

[Ignazio Leonardo Scarpelli]
This is a question that I believe is more subjective than anything else. Do you think, in your opinion, that the massive use of artificial intelligence techniques for the creation of content by competent journalists who are therefore in good faith can eliminate some limits? I'll give you an example.
Fake news usually yes, we understand that it is fake news because it comes because it is produced by artificial intelligence, so for example the use of artificial intelligence techniques can make it fall, make some limits regarding fact checking fall, it can make in a certain sense you also have the most difficult job.

[Giulia Pozzi]
But it can certainly make the job more difficult, from certain points of view.
For example, I'm talking about the whole part that we have already discussed, regarding images, videos, etc., etc. In the sense that obviously a massive use of artificial intelligence to produce that kind of content can really create difficulties for the facts checker, also because it will be content that will increasingly spread on a large scale. So it is not only a problem of technological evolution, but also of the scale on which this content will spread. So it will always be more difficult, let's say, to track and recognize. So from this point of view certainly. At the level of artificial intelligence, perhaps used widely by journalists, let's say that here, we should refer precisely to the meaning

and one of the basic rules of journalism which is to do fact checking. So fact checking is now almost becoming a separate activity, in the sense that there are now departments, organizations that do just that, precisely because the problem of disinformation has become so, let's say important. But in reality the basic rule of the journalist is to verify news before publishing it. So there will be an increasingly greater need to do this also on content produced, for example by artificial intelligence, chatbots, etc. So here we say they will distinguish themselves and perhaps quality journalism will be distinguished from non-quality journalism. Already in everyday work it is important to verify news, with the use of these tools it becomes even more fundamental from certain points of view. The largest editorial offices of newspapers, including international ones, now have a department that only does fact checking, right? This will become increasingly crucial. Surely.

[Ignazio Leonardo Scarpelli]
Exactly. I absolutely agree, also, I mean I talk about it, we talked about it before. Ultimately it's something that will go right into the. In everyday life it is an innovation that is usually called disruptive, so it will enter everyday life, it will enter everyday jobs and consequently it will also enter this sector.
Absolutely Giulia, I've run out of questions, I don't really have anything to say. First of all, I hope that the interview was pleasant for you. In short, I hope it was a discussion. Tell me, sorry?

[Giulia Pozzi]
Let me know. How's the thesis going if you want to keep me updated.
I always like it, even when we receive requests from this type of student. We are always very happy, because then we are also very curious to see how these types of jobs take shape, that is, how they evolve over time. So gladly, let me know. And among other things, I'll send you all the links I was telling you about, I'll send you a WhatsApp message, so I'll send you maybe the research we did on chatbots, but also some other links. I don't know if maybe to some report we wrote on artificial intelligence in journalism, so that maybe you can have some material if you want to cite it in your thesis, so maybe I'll send you a list now.

[Ignazio Leonardo Scarpelli]
Thank you very much Giulia and thank you for dedicating part of your time to me.

[Giulia Pozzi]

HI.

Interview with Lorenzo Canale, Researcher at Centro Studi RAI

Lorenzo Canale is an AI researcher for Rai (Radiotelevisione Italiana, Italy's public media company) and a Doctoral Student at Politecnico di Torino (PoliTo). His main research fields involve Learning Analytics and AI projects related to automatic metadata and fake news detection.

[Ignazio Leonardo Scarpelli]
So, I would say that we can start the interview. I have prepared some questions. The first thing I wanted to ask you is what do you do and what is your research field in RAI and Politecnico di Torino?

[Lorenzo Canale]
In Politecnico di Torino, I have finished a PhD in learning analytics, in the department of computer science. So, I have done a lot of analysis in different fields of learning, thanks to technology and a didactic field.

So, it is a bit educational. And also a bit of experimental, in the sense that, for example, since I still teach in a database course, I had created a game for SQL, that was similar to Cluedo, but in order to find the killer you had to interrogate the dataset.

[Ignazio Leonardo Scarpelli]
Interesting.

[Lorenzo Canale]
In RAI, this is a bit related to the activity that I did in RAI, because I work for several projects, including one of the activities that I did, which was the creation of a game related to the fake news. The other activity that I did, since my field in RAI was more related to artificial intelligence, even though I try to keep, above all, the research themes that are a bit interdisciplinary, which has

a somewhat humanistic component, because in my first path, I was a cinema engineer, then I did data science, but I have a bit of a mixed profile, and so to say that this project on fake news interested me because it mixes different competences.

And so, I was mainly involved in this project (i) doing the disinformation game (ii) trying to create a more technical component, not related to the game, that understands if an affirmation is supported or consulted by a text.

[Ignazio Leonardo Scarpelli]
Ok, ok. Then, returning to that algorithm that you developed in RAI, so, if I understood correctly, it is an algorithm that takes a certain sentence and puts it in comparison with an article that I guess is an authoritative article and this algorithm is able to understand if this sentence is more or less a fake news?

[Lorenzo Canale]
Yes, first of all, consider that the dimension is a sentence on one side and a paragraph on the other. So, giving a piece of an authoritative article as paragraph, the algorithms understands if the sentence is supported by the article portion.

You can use the algorithm with any collection, in fact, so it is like a box that if there are the resources, you can use them as a collection and see what results it gives you when you make a statement that is of the same domain of those sources. It is obvious that if the knowledge is not expressed in the documents you have, you can't use it for disinformation. So, since at the level of disinformation practically every domain of knowledge has different authoritative sources, in my opinion, it has an application more in a specific domain and one decides if he’s interested in having an algorithm that on these documents tells me statements that support them and I will be able to use them, but for now it has been tested especially the training procedure of the neural network that allows you to do this.

[Ignazio Leonardo Scarpelli]
Ok, I got it. Ok, perfect, I got it. As an expert, what artificial intelligence techniques would you recommend to a non-technical staff and above all, what algorithms would you trust to leave in the

hands of a non-expert staff and which, in your opinion, still need supervision from a technician or an expert?

[Lorenzo Canale]
But on this specific task, on the task that is called Recognizing Textual Entailment, so recognizing that it is always between...

[Ignazio Leonardo Scarpelli]
Sorry, Lorenzo, can you repeat? Recognized...

[Lorenzo Canale]
Recognizing Textual Entailment. Or it is also called natural language inference.

Ok. As an algorithm, I tested above all algorithms for natural language processing such as transformers, which are a family of algorithms.

[Ignazio Leonardo Scarpelli]
Yes, exactly.

[Lorenzo Canale]
But I also tested large language models.

Let's say that there is ChatGPT, that has a good performance without even training it. Other algorithms, if you tune them on a specific dataset, they improve them. Even the older ones improve their performance.

But the problem is that they become a bit specific for that dataset.

In addition this task is a non-subjective thing; let’s say that sometimes there is this problem in this domain, that you take a notator, you take another one, and for one, a test is enough for you to bring a sentence. For another one, it is not enough. So, the evolution, in my opinion, of this field is first of all to go from discrete labels to continuous labels. So, something that tells you that this

resource, compared to this other one, supports this sentence more, rather than getting to a complete support, which is a bit binary as a label, which sometimes is not enough. And you associate accessibility sometimes with individuality.

[Ignazio Leonardo Scarpelli]
Ok. So, as regards algorithms...

[Lorenzo Canale]
So, I tell you, to train algorithms, I would say to train either transformers, like BERT, or large language models, those that are open source, available, even those that work well. You obviously have to give them the sentence and the paragraph, and learn on the basis of the sentence and the paragraph.

For example, ChatGTP, when you do the prompt, you give a text and a sentence, and you ask if the text supports the sentence. The answer will either be yes, no or not enough information. Then the output is a label 0, 1, 2, depending on the meaning. You can give 0, 1, 2. Generally, 0 supports, 1 doesn't and 2 means not enough information.

[Ignazio Leonardo Scarpelli]
Ok. Ok, perfect. So, regarding the algorithms that you would leave in the hands of a non-technical person, a non-expert, and the ones that you wouldn't give in such hands, there is no disinformation algorithm that works well and at the same time is understandable to a non-technical person, to a journalist, for example.

[Lorenzo Canale]
I don't think there could be an algorithm that returns all the disinformation on its own. We were simply talking about texts.

If we already pass the domain of images, it is another field. For example, to render deepfakes, there are already quite available solutions, totally different from the ones we have talked about. For example, there is one called Sensitive AI, where I saw that it works very well on photos, instead of erasing them all, but passing through videos, I saw that it doesn't work.

On videos, in my opinion, it is still easy to make progress.

On the text part obviously, there isn't still something that tells you if a text supports a whole sentence, but it could be used either to create disinformation, or to fight it. It all depends on the media literacy and the documents you give to the base. For example, in the scientific domain, I would rather do it by looking at the quotes of Google Scholar, the quality of the magazine on which an article is published, or by giving the data collections that have already done this, like those of Our World in Data, and I would use those as, let's say, sources of reliable documents, on topics like global warming, or things like that, of external matter. On the other hand, on things more precisely of journalistic or of information, or where there are maybe even in some cases secret informers, I don't know enough, because I have never interfaced on where these resources are taken by journalists. So I don't know where there are totally reliable sources of information.

But, in fact, it is established that, for example, one could use ANSA[60] that as a source of documents and see all the statements that are computed only by the ANSA, when it comes to the data. So doing something similar also with journalistic sources. But there is not yet a solution that does everything.

One has to put all the boxes together, and so deepfakes and algorithm that tells you that it is a source that supports or doesn't documents for the text and it will already have in some way two more technological helpers. Then it will have to do a further verification. In fact, in my opinion, one of these sources can be useful.

[Ignazio Leonardo Scarpelli]
Ok, I got it. So, Lorenzo, then, among your research interests, there are also, you told me at the beginning, the so-called Data Learning Analytics. What I'm seeing in collecting the documentation, the information for my thesis, in short, a lot of emphasis is being given, especially in the papers that come out of journalism schools, on the importance of the education of journalists to Artificial Intelligence.

[60] Agenzia Nazionale Stampa Associata, Italy's national news agency

In your opinion, how do you think Data Learning Analytics can be used in this field, in the education of journalists on algorithms, on topics related to Artificial Intelligence in their work?

[Lorenzo Canale]
So, in short, interacting with these tools, that is, any teaching dynamic that makes you interact with the tool directly, because, in my opinion, one thing, for example, that of the game could be a thing.

[Ignazio Leonardo Scarpelli]
Exactly, so you are basically gamificating of the learning process?

[Lorenzo Canale]
One of them can be gamification, what comes to mind to me is gamification, then there are other people who have other ideas, I find it quite easy to learn, to create games to learn, so it is certainly easier for me to create an interaction with the tool to learn through the game.

[Ignazio Leonardo Scarpelli]
Ok, ok, I understand. Then, this is also a quite ambitious question, artificial intelligence, in your opinion, can be, and if so, how, put in the hands not only of fact-checkers and journalists, but also of the public, so the public that becomes debunker, in a certain sense?

[Lorenzo Canale]
Ah, well, I think it is a process that on certain issues is already seen, then on social, well, sometimes you can start simply the opinion. But, however, I think yes, if there are more tools. I give an example, the one about deepfakes we talked about before, in fact, the disadvantage is that one must stipulate a subscription and pay monthly. The majority of the public won't make this expense here. If someone had a solution integrated directly in a social, it would be different. The public would, I do not know, with a button, loading a photo, one would see if it was true or false, automatically, and in that case, yes, the public could be useful. However, at that moment, who develops these algorithms makes you pay with a subscription, and therefore there must already be a public that wants to pay.

[Ignazio Leonardo Scarpelli]
Ok, so there is also a sort of question of accessibility to these algorithms?

[Lorenzo Canale]
Yes, there is ChatGPT, if you use it as a source of accessible information, there are many things better done.

[Ignazio Leonardo Scarpelli]
Exactly, yes. And the same thing with GPT-4, in the end, if the public is willing, if it is interested in buying, or if it can also buy a subscription that is for GPT-4 or for another type of artificial intelligence, that, in the end, is also the important thing, so understanding the interest, the interest...

[Lorenzo Canale]
In my opinion, with GPT, for debunking, you don't see much.

[Ignazio Leonardo Scarpelli]
No, absolutely no.

[Lorenzo Canale]
So, however, any of these GPT-like tools can be useful, for example, to develop deepfakes, which in reality has shown that... there is a recent study that has shown that people recognize deepfakes better than real people.

[Ignazio Leonardo Scarpelli]
Really?

[Lorenzo Canale]

I don't know how much... Above all, the way in which deepfake has been generated, it is not certain that if an algorithm learns images of deepfakes that have been generated maybe with face

swapping, it is able to recognize those that, instead, have been generated directly by making a prompt like create a face from scratch.
[Ignazio Leonardo Scarpelli]
Ok.

[Ignazio Leonardo Scarpelli]
Ok, absolutely, yes, it's true. Also this. Returning to strictly personal questions, which algorithms have you used in your work?

For what have you used them? And dividing them by topic, which of these have had the best performance?

[Lorenzo Canale]
On what? On this project here?

[Ignazio Leonardo Scarpelli]
In general, which algorithm have you used in your work? Maybe dividing them by specific tasks.

For each task, which of these algorithms have had the best performance? Have you spoken, if I'm not mistaken, to IDMO of an algorithm capable of debunking 99% of the fake news that were given to it?

[Lorenzo Canale]
I have spoken of an algorithm capable of recognizing textual entailment. And it worked on the FEVER dataset.

This is mainly for the textual recognition. Instead, for face recognition another thing we deal with at the research center is, for example, the automatic recognition of the identity of the person by the face, especially to create metadata for what happens on television.

So the data is like the face of the person, in some way. And they manage to do it very well. The only problem is that when, for example, replace a face with that of another person (creation of a deepfake), since they are algorithms that recognize the identity of the face, they tend to make mistakes because they look at the face.

But, excluded in a context where you know that there are deepfakes of this kind, like they work very well. This, however, is only for the face recognition, which is a slightly different thing from the debunking of the faces.

On the debunking of the faces, perhaps, we can find some defects.

[Ignazio Leonardo Scarpelli]
Ok, perfect. So, etymology, those arguments, also those algorithms in which the sentence had to confirm at least what was said in a paragraph.

[Lorenzo Canale]
Exactly, exactly.

[Ignazio Leonardo Scarpelli]
Ok. Weren't you saying, however, that they are algorithms that…

[Lorenzo Canale]
No, no, no. You summarized a little. Ok.

[Ignazio Leonardo Scarpelli] (24:02 - 24:36)
The last... The last question I wanted to ask you was that... In your opinion, what are the areas of recognition of fake news in which you think AI is promising?

For example, finding elements related to fake news as a text in Caps Lock, finding a wrong score in articles that provide false narratives. So, in which of these areas do you think AI is more promising?

[Lorenzo Canale]
So, the recognition of deepfake and recognize textual entailment.

[Ignazio Leonardo Scarpelli]
Ok, and as regarding the contrast with chatbot, because I talked yesterday with Giulia Pozzi of NewsGuard and she told me that one of the main problems, that is, one of the biggest sources of misinformation that they are finding is due to these chatbots. Is there a contrast to that?

Is there an artificial intelligence technique that is perhaps more prominent in this?

[Lorenzo Canale]
It's a contrast to that chatbot space, I don't know, I haven't found it yet, but it would be an interesting topic, but I don't have enough information for now.

[Ignazio Leonardo Scarpelli]
Ok, perfect, I got it. Look, Lorenzo, thank you very much, I have finished my questions, thank you very much for your time, and at this point, I don't know, I have your number, in case you need a clarification, I can... I hope I can contact you and ask you some questions, maybe even without...

I'm taking care of artificial intelligence and media in general, so directly impacting artificial intelligence and the media sector, so anything, so the detection of fake news is just a part of my thesis.

[Lorenzo Canale]
Ok, ok. No, because instead of... for example, I come from a conference that is related to files, to the application of artificial intelligence on files, so for the operation, for the automatic segmentation of the content of my videos.

[Ignazio Leonardo Scarpelli]
Exactly. Ok, I got it, perfect. Well, as I told you, I could...

I hope I can contact you again, maybe to ask you some information, as well as to update you on what... on how my thesis is going.

[Lorenzo Canale]
Ok, ok.

[Ignazio Leonardo Scarpelli]
Ok, bye. It was a pleasure for me too, Lorenzo. We update each other at this point, have a good trip.

[Lorenzo Canale]
Thank you, good evening.

[Ignazio Leonardo Scarpelli]
Good evening to you too, bye.

[Lorenzo Canale]
Bye.

Interview with Michael Piemonte, US Army Reserve

DISCLAIMER: Michael Piemonte is a U.S. Army Officer, currently serving in the Reserve Component. The views and opinions expressed by Michael Piemonte during the course of the interview are solely those of the author and do not reflect the views or opinions of the United States Government or the United States Department of Defense or associated parties.

[Ignazio Leonardo Scarpelli]
Okay, I'm registering right now, so I will just take questions. So, you are a reserve in the U.S. Army in the intelligence branch, I suppose.

[Michael Piemonte]
Correct, correct. Military intelligence officer.

[Ignazio Leonardo Scarpelli]
Okay. So, what are your responsibilities, first of all, if you can say that? What do you do as a reserve?

[Michael Piemonte]
The reserves in the United States is fundamentally the same as the U.S. Army, it's just on a part-time reserve basis. And so, I began full-time active duty, which I did for four years, and then the last two to three years I've been in the reserve forces.

So generally the responsibilities for the reserves are to maintain readiness such that we can support the U.S. Army and the Department of Defense, and so that requires usually once a month we'll go in uniform. We have to maintain all of the same standards as active duty, physical, fitness, schooling, training.

As an intelligence officer I support a command that's called the Military Intelligence Readiness Command. Okay. Primarily I support the Joint Chiefs of Staff, which would be like the general staff of Italy.

[Ignazio Leonardo Scarpelli]

Okay. So, you are the commander of a part of the U.S. Army.

[Michael Piemonte]
I'm not a commander right now.

I can be in command as a captain, but now I serve as a strategic analyst. So it's more of a specialized group that supports strategic objectives, strategic analysis.

[Ignazio Leonardo Scarpelli]
Okay. But the work you do, it's only for the military and not for the internal security of the U.S.?

[Michael Piemonte]
No. That is for the FBI, law enforcement, and...

[Ignazio Leonardo Scarpelli]
Or CIA.

[Michael Piemonte]
No, the CIA does not. No intelligence entities in the United States conduct internal security operations. It's illegal. There's an executive order that was initially passed with Ronald Reagan, but it established the United States intelligence community and it prohibited United States intelligence officials from collecting, conducting intelligence operations on U.S. citizens, which includes U.S. persons and businesses. And so, to the extent we protect internal security is from external threats.

But the FBI is the primary mission set for conducting operations within the U.S.

[Ignazio Leonardo Scarpelli]
Okay. So, as I was telling you when I first approached you, my thesis is about AI and journalism. And, in particular, there will be a chapter which will talk about disinformation and AI.

Are you dealing in your job as a reserve, as an analyst, in AI techniques for creating disinformation? And are you seeing the rise of AI techniques in creating disinformation?

[Michael Piemonte]
Sure, certainly.

Though I am currently a reservist for the Army, I worked full-time for the Department of Defense (DOD) as a civilian for the nearly the last three years, before coming to Italy to study full-time. I worked in a very similar role for the Department of Defense, as I did for the Army.

And so, I had two positions in the last three years. And so, I'll talk about both and how I worked with AI. The first, I was an analyst, essentially doing international affairs analysis, like a think tank, but for the Department of Defense.

And so, I looked at mainly Russia and Ukraine. Before that, a bit counterterrorism in Afghanistan.

And so, of course, one of the primary tactics of Russia is disinformation, which includes AI systems. Again, I wasn't looking into, for example, Russian interference in elections per se. That's a different organization, like the FBI.

But I looked at...

[Ignazio Leonardo Scarpelli]
You looked more at probably the types of disinformation, probably at the ways...

[Michael Piemonte]
In which they used to shape the narratives of the Ukraine conflict.

Okay. And so, that included, in the beginning, things like Nazism. Okay.

Or these concepts that this was a war against NATO and NATO caused it. And so, as an analyst, I looked at the consequences of such information with regards to the narrative and the international community's perspective on the conflict. And we have technical experts who can kind of derive what information is propaganda, what technology is creating it.

And that's... It uses a variety of sources: cyber intelligence and other technical measurements, signals intelligence, to determine, you know, if they're AI produced, what type of firms produced them, things of that nature.

[Ignazio Leonardo Scarpelli]
Okay. So, if I can summarize, you're just... You're an analyst that just takes some narratives and...

You're a debunker, basically. So, you just take some narratives from, for example, Russia, about the conflict of Ukraine, and you just try to debunk them, try to go to the source of these claims, and then you actually... And then, what's your actions?

[Michael Piemonte]
Yeah. Well, I measure... Yeah.

It's not my primary role. I'm also measuring military analysis and geopolitical risks and things of that nature. But it all...

You have to have an inclusive picture of everything that affects the battle space. And so, there's a couple reasons. First of all, as an analyst, I'm making recommendations and assessments of the situation, of the phenomenon.

The people who make decisions and actions are a higher pay grade, let's say, they are operations personnel or policymakers. So, I'm presenting recommendations or just really assessments so they can make decisions.

Primarily, I'm a military intelligence analyst. So, if I'm looking at the battle space and the battlefield, I'm determining how effective the propaganda, the disinformation is in affecting the civilian population, and in turn, affecting the military efforts, right?

[Ignazio Leonardo Scarpelli]
Oh, okay.

[Michael Piemonte]
So, does the civilian population in Ukraine in certain areas, are they believing the propaganda at this point? Okay. Or I'm assessing amongst NATO partners.

How many people in NATO...

Are believing Russian propaganda...
And, yeah, so how robust is it? I mean, that's one aspect.

And then there's the higher level geopolitical concerns. So, it's not just that the population... It's not just how does it affect the battlefield or how does it affect NATO support, but how does it affect the international narrative of the conflict.

[Ignazio Leonardo Scarpelli]
Yeah, for sure. But, I know you are not a technical expert of artificial intelligence. But, in your experience, how would you describe the evolution of the AI techniques used to disseminate disinformation?

And what are the most used AI techniques in creating and disseminating disinformation?

[Michael Piemonte]
I mean, as you said, I'm not an expert. But, I think the hardest part now has become discernment because it's almost inseparable from the truth.

I mean, I think the technology has advanced such that it's very hard to discern from the truth…

[Ignazio Leonardo Scarpelli]
So, it's actually becoming difficult to debunk all this. To say that these things are false and that the other thing is true.

[Michael Piemonte]
Yeah. I mean, particularly with AI, I think the advancement of technology, one is the discernment of the information. It looks real.

And then the other is how rapidly it can be spread. Okay. And the advanced algorithmic techniques of AI because they can target very, very effectively certain areas of the population or certain narratives that they want to propagate.

And so, I think in conglomeration, the advancement of the technology in terms of how real, how effective it is, and then how efficient it can be spread is really what is hurting us. And it's nearly impossible to keep up with.

It's so proficient and so spread widely that it's hard to combat.

[Ignazio Leonardo Scarpelli]
Okay. So, it's impossible to stay behind in terms of quantity of information or in terms of quality, both of them.

Why it's impossible to keep up with all this information?

[Michael Piemonte]
Quantity and effectiveness of it. So, it's so proficient, it's spread so widely that it's difficult to counter every narrative effectively.

And then some of them are produced quite well. Especially if they're targeting people that don't have access to other information.

[Ignazio Leonardo Scarpelli]
Sure.

[Michael Piemonte]
Or that are, let's say, less educated on the matter.

[Ignazio Leonardo Scarpelli]
I made a question to the U.S. Ambassador here at the Fulbright event in which he actually told in an answer that he is willing to pay for quality information, but the majority of the people, the vast majority of the people is not willing to pay for quality information. And as an alternative, they actually go to rubbish websites that spread fake news.

So, it's important to understand how to disseminate quality information even to these people. At least a sort of quality information to these people. And at the same time not make the journals and newspapers, the websites, lose their profit basically from their subscription

I've talked before with an analyst, a journalist that is working as an analyst for NewsGuard. So, I don't know if you know NewsGuard, a company based in the United States, but they have also established some offices in Europe. I think they have an office in Rome.

And she actually told me that the preferred way of creating, the preferred way of using AI to create and disseminate information is with chatbot. Can you confirm that? These chatbots are basically robots, not robots, algorithms.

[Michael Piemonte]
No, I understand.
I don't know if I can confirm that. I don't have the information to confirm that. But I think likely in terms of what I've read, that's likely true.

I mean, yes. I've seen in my experience both kind of human oversight as well as kind of chatbot. I mean, I think I've seen both and I think the trend is towards more chatbot.

[Ignazio Leonardo Scarpelli]
I won't ask more. I hope I won't ask more technical questions about the matter. But if you can tell me, what are the most active countries using AI to produce disinformation and false narratives apart from Russia?

[Michael Piemonte]
So, I mean, fundamentally all countries conduct information operations. But I think apart from Russia, the most effective country is the People's Republic of China, overwhelmingly. Their tactics are somewhat similar, but somewhat different.

And their goals are different, of course. They have different goals. But, yeah, I would think the Chinese state is the most efficacious user besides Russia.

And especially, I mean, if you're measuring who's the most effective user, there has to be a metric how successful they are. And I think in terms of what we believe the goals of China's actions in that space are, I think they've achieved or are achieving them ~~in~~ quite a lot.
[Ignazio Leonardo Scarpelli]
Well, maybe that's a question that will come back later. But are you conducting also operations of counter disinformation to these countries?

[Michael Piemonte]
Yes, I mean, I'm not an information operations officer, but fundamentally the counter would be to expose the truth.

And I think that's the goal of public operations, of information operations, is fundamentally to expose the truth vice any type of, you know, tactic that would be similar to what an adversary would use.

And so certainly, I mean, we do – I don't speak for the United States government, by the way. But we do – in my experience, we do seek to counter false narratives in a variety of capacities and spread the truth. Okay.

[Ignazio Leonardo Scarpelli]

Also, regarding also there is this – there is the news, which is that the Congress of the United States has banned the use of TikTok. It's not banned yet, but I think a starting process to ban the use of TikTok in the U.S. Are you seeing – you were telling about the People's Republic of China techniques to disseminate disinformation. Do you think that TikTok is part of these techniques?

[Michael Piemonte]
Certainly. I mean, it's nearly unequivocal. I don't have access to the information necessarily that the United States Congress used to produce that law – well, legislation.

It's proposed legislation, which has not passed as you've considered. I mean, we still do have a democratic process such that it can pass. But I'm sure some of it is classified at a higher level that would prove it because there's no reason that our politicians would act without it.

I mean, I think even you can tell with openly available information that TikTok is used. It's not a very hard logic to follow because if you look at the line of responsibility, China has relatively robust national security laws, which force its citizens and its persons, which include its businesses, under its auspices to provide the Chinese state with information upon request. And it's under the threat of imprisonment.

It's under the threat of worse actions in some capacities. And so fundamentally, that legislation, though – again, I'm not a legal expert – but that legislation is referred to a lot in the media as a ban. However, fundamentally, it's not a ban.

It's an attempt to legally – to establish a legal basis to separate the chain of custody of the leadership of TikTok from the Chinese government.

[Ignazio Leonardo Scarpelli]
Okay, so it's just a detachment of the U.S. branch of TikTok from whatever is called China TikTok. I don't remember the name.

[Michael Piemonte]

Yes, exactly. In terms of how I've heard legal analysts describe it. Now, I'm not saying what the results of that law will be in terms of access to TikTok in the United States, but I know that the purpose of the legislation is not we're banning TikTok.

It's to sever the chain.

[Ignazio Leonardo Scarpelli]
Okay. It's to break the chain, basically.

[Michael Piemonte]
Exactly. Because the point is, with the current Chinese national security laws, the TikTok leadership would be forced to provide data of TikTok users to the Chinese government.

[Ignazio Leonardo Scarpelli]
So TikTok is not providing data to the Chinese government by its will, but because it's constricted.

[Michael Piemonte]
I don't know the will of the TikTok leaders. Maybe they are, but either way, whether they want to or not, if that chain exists, China can force them to. It's how we read it.

[Ignazio Leonardo Scarpelli]
Sure. But wasn't TikTok bought by Microsoft during the Trump administration?

[Michael Piemonte]
First of all, I don't know. I think the issue, again, I'm not an expert in this, I think the issue is where do the TikTok servers reside in which the data is stored? And what does the parent company of TikTok, regardless of if the US entity is owned by a US person, the parent company of TikTok, do they have access to the data?

And does the leadership of the parent company of TikTok have a responsibility to give the data to China if asked?

[Ignazio Leonardo Scarpelli]
Okay. So, in your experience, there are other attacks on national security, which involve AI and journalism, together with just disinformation campaigns.

[Michael Piemonte]
What was the first part of the question?

[Ignazio Leonardo Scarpelli]
In your experience, there are other attacks on national security, which involve AI and journalism, together with just disinformation campaigns?

[Michael Piemonte]
Yeah, I mean, there are many. I'll tell you about, I think, fundamentally, the biggest threat is it's interference with free and fair elections.

Because that is the foundation of democracy, at least one of the foundations of democracy. And so, such that you can disrupt a free and fair election, whether it's influencing the vote or influencing the peaceful transition of power, you're, I mean, fundamentally destroying, attempting to destroy democracy.

[Ignazio Leonardo Scarpelli]
Yeah, democracy pillars, basically. So, okay. And which techniques are used, for example, are used techniques of AI in order to enter into the computer that counting the votes of the election and automatically change, passing some votes from one candidate to another, or there are other techniques?

[Michael Piemonte]
I can't speak technically on exactly what they were doing. I think a good resource, which I would recommend to read, is the FBI report on 2016 election interference.

Again, because this is internal U.S., and I don't, you know, that's not my area. It's called the Mueller report. It was the FBI director at the time. It's been redacted and released and unclassified. You can

Google it and find it. And it describes in relatively robust detail how Russia infiltrated the 2016 elections. And it also proves that it was, there was no collusion.
There was no U.S. persons involved in the operations. So, it'll describe pretty well how the Russians were able to do it purely using technological methods.
Without traditional, what you would consider traditional intelligence or spying methods. They weren't using people. They were using purely technology.

[Ignazio Leonardo Scarpelli]
They were using purely technology. So, no people were into this, no…

[Michael Piemonte]
There was no proven collusion or suspected collusion of a U.S. citizen advancing the Russian attempts to influence the election.

[Ignazio Leonardo Scarpelli]
Well, also I know you're not a technical, you don't have a technical background.
But usually in order to counter disinformation campaigns, do you use AI tools? And if so, if it's not classified information, just tell me. And if it's classified information, tell me, what tools do you use?

[Michael Piemonte]
Yeah, I don't know. I mean, the… Yeah, that's a different office.

That would be information operations people or public relations people. And my responsibility is to understand and make assessments about the threat.

And however they're countered is up to a different people. Yeah. But certainly we do conduct disinformation, I mean, we don't conduct disinformation.

Okay. We conduct activities to counter disinformation.

But I don't know the tools that we use to do that.

[Ignazio Leonardo Scarpelli]
Okay, so you're just an analyst that analyzes information and passes it to your superior, I suppose.

[Michael Piemonte]
Yeah, certainly.

I mean, I have a variety of responsibilities as an… I did as an army officer. But yeah, I don't…

I'm not conducting like kind of anti-operations or anything of that nature.

[Ignazio Leonardo Scarpelli]
So, but you use… Okay, you are telling me that you are using the same techniques that are used on the… that are used from other countries to the US in order to try to not disseminate disinformation but to disseminate the truth as you are saying to other countries in a way or in another, like China or Russia.

[Michael Piemonte]
I can't speak intelligently on that. I mean, it could be… Yeah, I can't speak intelligently on it.

I think that we use a variety of the same type of PR, public relations or information methods that anyone in the world uses, whether it's spread through news media or various platforms and social media. I think so.

I mean, we are a nation fundamentally of laws and we can't at freelance conduct, you know, illegal activities against other countries. I mean, we respect the laws and so to the extent that we can use traditional and newly developed technological methods for spreading information, I'm sure they're used. I don't know exactly what they would be.

[Ignazio Leonardo Scarpelli]
Is your branch, sorry if I call it branch, you're collaborating with news networks regarding the use of AI in journalism, especially for disinformation. Are you collaborating or not, first of all?

[Michael Piemonte]
Yeah, I mean, not specifically my branch, but the United States Department of Defense and everything underneath it certainly does. I mean, we have officers.

We have another branch of experts who are public relations and information operations or information warfare officers. And so I was trained as a military intelligence officer. They're trained as public relations and information operations officers.

So those are fundamentally the people who are doing that and they collaborate all the time with public media. And you can see also from the headquarters of the U.S. Department of Defense from the Pentagon and from our different bases, we're conducting interviews, we're very public facing. Because again, I mean, we have to be because we are beholden to the taxpayer. And our responsibility is to the taxpayer that supports us. And so we're trying to remain open.

[Ignazio Leonardo Scarpelli]
Sure. But do you know if it's a one-way collaboration? So either from the newsrooms to the U.S. Defense Department or from the U.S. Department to the newsrooms? Or is it a two-way collaboration?

[Michael Piemonte]
No, I'm sure it's symbiotic. I'm sure it's two ways.

For example, when I worked in the Pentagon, we have private news media.
I mean, the U.S. doesn't have a public – ~~it's~~ a state-sponsored news media. We have news personnel who are not working for the Department of Defense who work in the Pentagon. It's called the press office of the Pentagon.

I mean, there are people who are fundamentally – they're working for a variety of different news and media sources. And they're there to engage with defense officials and cover the stories whenever there's a press conference, a briefing about their ongoing activities or operations or threats.

I mean, those people walk around and they're colleagues of ours in the Pentagon. I mean, so the relationship is very symbiotic. They're there with us, you know, and the relations are quite well.

They travel. For example, when high-level officials travel, let's say, for example, the secretary of defense or the chairman of the joint staff, they travel with news media personnel. Okay, sure.

[Ignazio Leonardo Scarpelli]
But just the news media are there for a purpose of just collecting news, just make a report of the visit of the –
[Michael Piemonte]
Oh, yeah. And in those capacities, they're there for news purposes. But there is collaborations in such that we're working together.

Okay. I mean, we in the security community, in which I include the FBI, we work with all private businesses.

[Ignazio Leonardo Scarpelli]
Yes.

[Michael Piemonte]
Specifically what you would call counterintelligence and FBI. They work with private corporations to expose threats and things of that nature. I mean, we're always working with our private counterparts.

[Ignazio Leonardo Scarpelli]
So, yeah, I mean, you just work with private and it's a two-way collaboration. So they just flow information to you, to the U.S. Department of Defense, and the U.S. Department of Defense just probably gives information also to the –

[Michael Piemonte]
Yeah, and I'm sure they collaborate on the best ways to counter and identify disinformation. I'm positive they do.

[Ignazio Leonardo Scarpelli]
You were in Afghanistan from what year to what year?

[Michael Piemonte]
The first time from 2018 to 2019, and then the second time 2020 to 2021, so a few months before the withdrawal, but I still actively worked on Afghanistan from the United States until after the withdrawal. Okay. I covered it as an analyst.

[Ignazio Leonardo Scarpelli]
What was your role? You are telling now you were not an analyst there, but what exactly was your role?

[Michael Piemonte]
The first time I was a tactical level military intelligence officer, so I supported combat operations. It's nothing to do with like James Bond or spying. It's identifying threats and the enemy areas that we should target and informing my operations counterparts, the infantry folks, the aviators, the special forces, informing them about the threats and the type of missions that they should be conducting. Okay. So, yeah, it was very mission-oriented and tactical level. And the second time I went, I worked at the NATO headquarters, in which I did more theater level, so operational and strategic analysis, again, painting pictures of the threats and the situations ongoing, making predictions, and then I worked for my last portion there, I lived down in, fundamentally, in the Afghan base in Helmand, Province which was the southwestern portion of Afghanistan, near Lashkar Gah, with what was called the 205, the 205th Corps of the Afghan Army.

I was supporting their training and operations. So I was helping their officers to be better intelligence officers, to be better operations officers, with a small team of army advisors.

[Ignazio Leonardo Scarpelli]
But it's not what you're doing right now. It's completely opposite. Maybe, okay, the next part of the question was, did you use a – so you were not countering this information.
You were just –

[Michael Piemonte]
I had colleagues that did it, they would be called the Information Warfare Task Force, or Information Operations Task Force, and that was their goal because, you know, if you're looking at how I, similarly how I described Russia tactics in Ukraine, the Taliban, the insurgents' tactics were to influence the population there, as well as ISIS, as well as other threats, as well as external threats, right? There's, I mean, there's near-peer countries that are trying to influence populations in Afghanistan, but fundamentally we're looking at the Taliban, and then next to that was ISIS, which was a relatively large threat, and it still is in Afghanistan, as well as other...

But ISIS and the Taliban are against each other.

And there were that too. But, I mean, ISIS was still trying to influence, you know, the local populations where they had a threat and a stronghold.

[Ignazio Leonardo Scarpelli]
Yeah, of course. Sure. And so, your colleagues on the Information Task, on the Disinformation Task Force, did they see a great use of AI to spread misinformation during the U.S. presence in the country? Did they see it? A great use of AI to spread misinformation?

[Michael Piemonte]
No, not from the insurgent and terrorist threats that we were facing. They didn't need to because the populations in Afghanistan that they were influencing didn't have even access to these things.

[Ignazio Leonardo Scarpelli]
Well, the last question was just in your experience, how much disinformation can disrupt national security?

[Michael Piemonte]
Substantially. I mean, it's hard to put it into a quantity of measurement or a metric, but substantially. I would consider it one of the largest threats that we face because it's very low cost and it's very easy to do.

And ~~so~~ it's very proficient. It happens all the time.

[Ignazio Leonardo Scarpelli]
Okay, sure. And in a list of possible threats, where is it positioned? So I have 10 threats. The top 10 threats of our national security.

[Michael Piemonte]
Okay.

[Ignazio Leonardo Scarpelli]
Where does disinformation position itself?

[Michael Piemonte]
I mean, are we considering like...

[Ignazio Leonardo Scarpelli]
We're considering the top... I don't know what are the top 10. The top 10 threats of our national security.

But on these 10...

[Michael Piemonte]
Well, I mean, fundamentally, it's who's using the... The largest threats are from NEAR-PEER adversaries. And it's fundamentally, it's those countries using disinformation effectively as part of their grand strategy.

And so it's hard to rank because obviously, the weapons systems and things of that nature can kill people. But in the absence of that, it's an extremely effective method to achieve the adversary's goals. Okay.

And so ranking it on 1 to 10, I wouldn't feel confident making that because, I mean, I would say it's a very effective tool that is used to advance the grand strategy of adversaries. And it supports... Fundamentally, it supports the other goals and the other operations that they can do.

And so I think it's one of the primary ways in which they support grand strategic objectives at all levels. Okay.

[Ignazio Leonardo Scarpelli]
Sure. Okay. That was my last question.

Interview with Beatrice Petrella, Freelance Journalist

Beatrice Petrella is a freelance journalist. She works for many Italian newspaper and websites, such as "L'Espresso" or "Domani", and produces podcasts by herself.

[Ignazio Leonardo Scarpelli]:
Perfect, let's start. So, Beatrice, first of all, thank you again for being available for this interview. The first thing I want to ask you is to introduce yourself.

What do you do, what is your job, and what does it mean to be a freelance journalist?

[Beatrice Petrella]:
So, I'm a freelance podcast journalist, which means that I write articles and make podcasts for several newspapers. Last year I realized, with Storytel, "Still Online", a podcast about digital legacy. I wrote for L'Espresso, a survey on property movements at the University of Pavia.

For Domani, instead, I told the story of press freedom in Malta. And now I'm the finalist of the "Morrione Award". So, in short, I do different things. I tend to deal with foreign affairs, but also political news and rights.

And so, freelance means that, in fact, I am, let's say, a not-attached journalist, because I collaborate with several headlines.

[Ignazio Leonardo Scarpelli]
Exactly. So, basically, are you the one who goes to the headlines giving an article, or are they the ones who come looking for you to give an article?

[Beatrice Petrella]
So, no, it works that you are the one who sends the proposals. Ok. Which, on the one hand, is very nice because it allows you to space out on many things but, on the other hand, it means that you have to build your job every month, which is a great freedom, but it is also…

[Ignazio Leonardo Scarpelli]
Stressful, I guess.

[Beatrice Petrella]
A certain kind of stress, because, in short, you build yourself the salary at the end of the month. What gives me a lot of satisfaction is that I have the freedom to choose between many different projects.

[Ignazio Leonardo Scarpelli]
Ok, and not, perhaps, focus on the same project that is given to you by a particular newspaper.

[Beatrice Petrella]
Yes, even if, however, let's say, there are recurring themes. For example, I also collaborate with Roma Today, whit which I deal with things related to health issues in Lazio region, but this does not exclude the fact that I'll talk about Internet culture, or I deal with foreign affairs for other headlines. So, you have to find a balance.

[Ignazio Leonardo Scarpelli]
Absolutely. So, the second question I ask you is about my thesis. I am a Data Science student and my thesis, as I told you, is on Artificial Intelligence in Journalism. So, personal experience, do you happen to use AI tools in your work today?

For example, ChatGPT, Speech-to-Text conversion algorithms, data visualization tools. How much do these tools influence your work, if you use them?

[Beatrice Petrella]
So, I use a lot of Pinpoint for my work, which is a tool available on Google, which is very convenient converting interviews into text. It does not always work, especially for what is not in English, but, let's say, it saves me a lot of time. And then, when it comes to making summaries or preparing, to

give a, let's say, a first setting to maybe motivational letters or any tasks that can be automated, sometimes, not always, but when I'm short of ideas or a prompt to start, I ask CHATGPT-4 for a hand.

[Ignazio Leonardo Scarpelli]
Ok, so actually...

[Beatrice Petrella]
Anyway, it is useful in many cases, because the relationship I have is like having a sparring partner, so someone you lean on and say “ah, actually you could say that, but I could say it better.”

[Ignazio Leonardo Scarpelli]
Exactly.

[Beatrice Petrella]
And so, I mean, I start from there and re-elaborate, also because, most of the time, you understand when a job is done with AI.

[Ignazio Leonardo Scarpelli]
Exactly.

[Beatrice Petrella]
Especially if then, maybe, you master the language and you have your own way of speaking in that language. But there are times when, in my opinion, it's a good help. It's a good help even when, for example, I'm learning to code, and when you don't find the error in a string, I think you would know that too, because it is very useful, to find the error.

[Ignazio Leonardo Scarpelli]
Yes, absolutely. I also use it for this, in the end, to find an error in a code string that maybe gives me an error. So, excuse me, you said you already use ChatGPT, especially to find ideas on what to write.

[Beatrice Petrella]

Not exactly. I ask ChatGPT on how to write my ideas. For example, when I'm in a period where I have to send applications for projects and maybe I don't have the basic idea on how to start, you give him the right prompt on how to start, he gives you a handkerchief and maybe you can rework from there. Let's say that, in my opinion, the good thing is that it's a great support, a great piece of support. Also, it's useful to automate things. I mean, I think maybe even in the economy of a newsroom, if you can delegate to an experiment such as ChatGPT or boring tasks like a report on weather forecasts or other topics that humans can delegate to something else where, instead, the human eye is important.

And then, of course, the human checks again, because we know that ChatGPT is smart up to a certain point. But, let's say, it's a good way to speed up. Another thing I use is Data Wrapper to make graphics. And, let's say, it makes an honest work. It also produces clean graphics, you can share them in such a way that there are sources, those who want to go and see the sources can check them, you can compare them, so, in my opinion, it's a good thing if you do data journalism. And then, of course, if you draw your own graphics, it's even better. But, let's say, it's a clean tool, good for sharing.

[Ignazio Leonardo Scarpelli]
Especially for beginners, I imagine, for people who are not used to Business Intelligence tools like Tableau.

[Beatrice Petrella]
Exactly. If you can't use it, or if you can't use it, there's Data Wrapper. An alternative is also Flourish that is also very clean.

[Ignazio Leonardo Scarpelli]
Ok, if you had to give me a percentage… So, you told me many artificial intelligence tools that you use. If I had to…

[Beatrice Petrella]
Wait, there's another one.

[Ignazio Leonardo Scarpelli]

Yes.

[Beatrice Petrella]
Which is very important. Deepl.

[Ignazio Leonardo Scarpelli]
Deepl, yes, is the translator, exactly.

[Beatrice Petrella]
Deepl, ok, him too.

[Ignazio Leonardo Scarpelli]
Deepl, yes, is a very powerful translator. I used it when I was in Erasmus in Holland to translate texts from Dutch to English or Italian. So, it's absolutely very useful.

By the way, another thing I wanted to tell you. You told me many artificial intelligence tools. If you had to give me a percentage on how much artificial intelligence affects your work, how much would you give me?

[Beatrice Petrella]
Well, let's say… It affects, in the sense that it speeds up a lot for me, because if I think about the time I lost to transcribe audios by hand…

[Ignazio Leonardo Scarpelli]
Ah, yes, absolutely.

[Beatrice Petrella]
Because it has to be reviewed frequently, but it takes a lot less time than if I were to transcribe audios of two and a half hours, maybe. So, in my opinion, in terms of speed, especially Pinpoint and Deepl, which are the ones I use the most, because even though I'm bilingual, at a certain point a translator is convenient, because obviously you'll fix the final translation, but what I notice is that sometimes I

don't have the automatic translation in my head. So, it's very convenient to have someone who gives you… well, it's always the idea of the “canovaccio” and then you fix it, or anyway, to get the meaning better, you speed it up. So, yes, I would say, above all, these two tools make my life as a journalist infinitely easier, because then maybe, even if you get Deepl, you get the text that is not a language you know, maybe you get there, or something happens to you that you really don't have the tools to translate, and he has them.

[Ignazio Leonardo Scarpelli]
Exactly. No, absolutely. It is also the reason why I use Deepl, or more simply Google Translator, if I don't have a word in English that doesn't come to mind, I just go and translate.

So, it's a great tool. But more or less, in a percentage, how much does artificial intelligence help you in your work?

[Beatrice Petrella]
It helps me in the sense… it comes to me to say that my work has improved, I mean, I don't know if…

[Ignazio Leonardo Scarpelli]
It has become faster, I guess.

[Beatrice Petrella]
It has become infinitely faster. I used to spend a whole morning on it, or even more, now I spend 20 minutes on it.

[Ignazio Leonardo Scarpelli]
Exactly.

[Beatrice Petrella]
20 minutes to say a lot, if you want to fix everything, so I would tell you that it has improved. It's hard to say it in a percentage, because 100% is a bit too much, because you have to give it the last

look, and then, as you say, especially for Italian, they are not yet optimized to such a level, but it's already like that. Then there are also other softwares that are not Pinpoint, even faster.

Which, to guess, run better, and I am working on those. Now I will study how, but in short. A percentage, I would say, an improvement of…

Well, 60% already only for the existence of Pinpoint, then for the translator, one could dig it, but Pinpoint, really…the time saved, the years of life saved…

[Ignazio Leonardo Scarpelli]
Exactly. Well, in preparation of this interview, I looked a lot at your LinkedIn and I noticed that in your experience you participated in various courses, you did an executive course, LUISS, on AI for media and journalism, you also recently completed a course in data journalism, so what I wanted to ask you is what were these courses about, what did you learn in these courses and how did them change your perspectives change on artificial intelligence in journalism?

[Beatrice Petrella]
So, the LUISS one was the first one I did, unfortunately it was a remote course because it was in the middle of COVID, so in short, it was adventurous, but it was the first time that, I mean, I also interacted with more practical tools for journalism and from the initial shock moment with all the things that AI can do, it was that thing that made me say “ok, but what is it really like, how does it work?” And so from there, I also got the idea of being able to do journalism with data and think that it was something that could interest me, because before that moment I had never thought about it.

Then I discovered Mona Chalabi, who is a very famous data journalist who won a Pulitzer, with a graphic representation of the wealth of Jeff Bezos, so how many things can you buy with the wealth of Jeff Bezos. It's very nice because you illustrate by hand, so he does his illustrations and then digitizes them, otherwise look for it, then I'll send it to you. She is very talented, so I thought, Mona Chalabi, written with CH, I'll write it to you, look, I don't know if you've already found it, but in this case, it's her.

[Ignazio Leonardo Scarpelli]

Ah, ok, perfect.

[Beatrice Petrella]
So that, that made me a little, no, well, I was saying, so with that, I understood that it could be interesting, then I did a summer school of digital humanism at the University of Vienna last summer and that was another very interesting adventure because I was the only journalist surrounded by engineers and so it was very nice to be able to develop projects together and to be able to contribute, even if, I could do less, but in terms of perspective, politics and also communication, I could give a lot, so it's was nice to see how the different disciplines intersect themselves. At the end, I did this course of Data Ninja which was dedicated for English journalists and very well done, very complete for those who didn't have a base, but also a great starting point and to say I want to do this also because the final job was a pitch of a project of its own. I talked about the disorders of food behavior in Italy, I illustrated it with Data Wrapper and it's nice to see how one can start from data research, from data cleaning, from arranging them, from understanding them, putting them in order and then see how to tell them, which is a very important thing because very often we think that data are the absolute objectivity because they are numbers, and it's not true because data are biased, it depends on how we collected them.

[Ignazio Leonardo Scarpelli]
It depends on how we collected them and also how we interpret them so maybe a number can mean something we can think that that number means something but maybe that number has a meaning up to itself or it means nothing.

[Beatrice Petrella]
Exactly, because it also depends on the context and on how we question them. There's a sentence that I like a lot that says that if we torture data enough they will tell us what we want. In my opinion it illustrates very well the problem of thinking that data are the objectivity just because they are numbers.

[Ignazio Leonardo Scarpelli]
Absolutely, yes. True. How did all these experiences your perspectives on AI journalism?

[Beatrice Petrella]
Actually, in general, they have contributed to improve my relationship with numbers and statistics in the sense that they made me understand that it's possible to work with them, and it's possible to apply and make a contribution. I am always very positive with new discoveries, in the sense that everything depends on how it is used, but obviously in the field of journalism there are problems in general with AI because we know that it can be very dangerous, for example with deepfakes. This problem, we have thought about it, from the beginning. I remember one of the first big fakes that was shot was this video of Queen Elizabeth when she was still alive, so quite old, and everyone was laughing because it was a funny video, but the first thing we thought because we saw it during the LUISS course it was "ok, this means that someone can take Biden, for example, and make him say what he wants"

[Ignazio Leonardo Scarpelli]
Exactly

[Beatrice Petrella]
and then we got to the Pope's photo with

[Ignazio Leonardo Scarpelli]
with the pullover yes, exactly

[Beatrice Petrella]
So, in my opinion, on the one hand it's crazy, because, as I was saying, we have reduced our time to do very boring and very long tasks to be able to dedicate ourselves to better things, even studying. I mean, while a thing is done by PinPoint in the meantime I do all the human work, like another research job or all the beautiful part of the job that is writing and the creative side. On the other hand we know that these things are dangerous if are not well managed, because if the problem of fake news was a thing before, now it's worse. And even worse, right now we are in the electoral period, there are elections all over the world at various levels. In my opinion, like all the things, someone has to be able to manage them. On the one hand I think that if AI entered more in the

newsrooms that would be very nice, precisely because at a time level everyone could work better by automating things, on the other hand we should also be long-sighted in the management and understanding that actually these things now because there is chat ChatGPT but another thing is the question of the author's right on works created by AI

[Beatrice Petrella]
in my opinion, the thing is that we should not lose the beauty and the power of certain instruments, but realizing that these instruments have strong limits and strong dangers. So in my opinion if one approaches things with rationality of course it means that we should have take seriously into consideration the fact that these things are a problem and understand how to manage them. Now, the AI Act has arrived it is very recent it is the first form of legislation on artificial intelligence we have to see now why it has arrived now.

[Ignazio Leonardo Scarpelli]
Maybe because of two years/a year and a half of preparation after the exploit of ChatGPT

[Beatrice Petrella]
Yes, I mean in my opinion before finding a regulation a formula that works and that has real practical effects, in addition to pointing out the fact that politics tells us “look, we are thinking about it”. It takes time, the problem is that artificial intelligence is faster than the time we need.

[Ignazio Leonardo Scarpelli]
Absolutely.

[Beatrice Petrella]
and this however is a basic problem of the legislative times compared to human ones.

[Ignazio Leonardo Scarpelli]
so, at this point we not only need politicians but also prepared technicians on the subject.

[Beatrice Petrella]
that is the main problem, that is the voice of those who are inside.

[Ignazio Leonardo Scarpelli]
exactly

[Beatrice Petrella]
specialized technicians but also for example people who live the internet and are specialized technicians in many sectors. For example, now there is a proposal of law on sharenting.

[Ignazio Leonardo Scarpelli]
Ok, tell me more about it.

[Beatrice Petrella]
Sharenting is the practice of sharing babies' photos online by parents who base their business on being internet parents. It came to us because Serena Mazzini, who is a social media strategist but also an activist, did a lot of investigation on this subject and therefore we need technicians of many specialties to get to find a frame, I think. Then, of course, it comes to the filter of having to make a law.

[Ignazio Leonardo Scarpelli]
Because it's a problem there is a need to regulate the thing.

[Beatrice Petrella]
And also to label things as clearly false. I mean, now it comes to my mind that, for example, on twitter the fake news question is regulated through the reporting by people.

[Ignazio Leonardo Scarpelli]
It is like giving the task of investigating on a homicide to normal people instead of relying on the police.

[Beatrice Petrella]
yes, but on the one hand it is interesting that the responsibility is of the single user, which does not mean that it is right, because if there is a company behind it should be under the company's

responsibility because it is not the same free internet of when the internet was born, but it is an interesting social experiment and I find very important that the users, because everything is being deregulated, are trying to manage this thing in some way that is very appreciable. It is not the optimal solution, but it's nice to see a renewed sense of community.

[Ignazio Leonardo Scarpelli]
I understood. Look, the question I had in program after the last was about the opportunity and the risks of a massive use of artificial intelligence in the journalism sector, but you have already answered me with this speech. Now, I do not know if you have seen the video I sent you about the so-called six-phase process. I'll explain it quickly. Basically, in this video, that has been inspired by another video from a journalist from the Belgian state television that does courses of machine learning for journalism in Stanford, a data journalist divides the journalist's work in six phases and for each phase makes examples on how artificial intelligence can improve these phases. The first phase is basically the idea, and so look for possible interesting topics to talk about on which basing a content, whether it is an article or a video, then there is research, so acquiring as much information as possible on the topic, then there is the production of this content, and so the writing of the article or the production of a video or any other content. Then there is the publication of the content, the feedback received from this content, and so the comments and feedback received by external users, this is also the phase in which the fact checkers debunks if a content contains false information, and finally there is the archiving, so the moment when the content is in a certain sense archived and relations are created with the other contents. First of all, I ask you if these six phases according to you are consistent with your work and, second thing, in which of these do you believe that artificial intelligence can help more the journalists?

[Beatrice Petrella]
I've never thought about a production process in a schematic manner, but yes, is right, more or less. I don't know, actually, in my opinion fits better in the last phase. For example, I use a lot both TikTok and Instagram to share current videos I must say that the tiktok algorithm has a draconian style so maybe sometimes you write comments that are bannable for their policies but actually compared to the commented video they make sense.

[Ignazio Leonardo Scarpelli]
ok

[Beatrice Petrella]
so on the one hand it's very positive that there are so many because it's a mainly aimed at young people on the other ok but basically it makes sense to use artificial intelligence even for the ban and suspension of certain contents. In my opinion, the last eye of a human being is useful because in my case, for example, they suspended contents that, according to the algorithm, violated I don't remember exactly which policies but actually they (the contents) were absolutely ok. So let's say, my point more than the single comes to tell me that we could lean almost always on AI, but it always depends how we do it and how we are used to work in the single but we must remember that the last eye is ours because the AI is not really intelligent we decided to make it intelligent. So it has a whole series of limits that we gave it by educating it

[Ignazio Leonardo Scarpelli]
yes, true. So, I also sent you an interview with Giulia Pozzi. Giulia Pozzi is a journalist that I interviewed, and she is an analyst for Newsguard, an organization that debunks fake news and false narrations and that collaborates with the institution for which I am doing the internship now, (LUISS DataLab) which is linked to the School of Journalism of LUISS, and during this interview we also discussed which are the most used of disseminating fake news and false narratives. She told me, in particular, that the most used are those of the chatbots that produce tweets or disinformation websites. There are many sites that reproduce false with articles clearly written by AI clearly written by ChatGPT. Clearly there are also some algorithms that can be very important also to debunk these false narrations. Now, in your opinion, similar algorithms to disseminate disinformation have some possibility to be integrated also in the process of creating good information or bad information creation?

[Beatrice Petrella]
Well, I think it always depends on how you use them. I mean, we have seen that if you can already use them, if you use the same tool to create fake news and debunk them, what prevents you from making them write an easy, controlled article that can be a football match record or something else

where the human being is a bit wasted and almost bored? In my opinion, on a practical level, it can be done.

Probably in terms of optimization it also makes sense, we still have the problem of fake news.

[Ignazio Leonardo Scarpelli]
Exactly.

[Beatrice Petrella]
Because, I mean, in the end, we are always there. It is a great discovery, it is very useful. What do we do?

Great powers derive great responsibilities, so it always depends on how we choose to use them. If we choose to make ourselves debunked in interviews, or if we say, well, come on, but since we are in the US Election period, let's pretend that Biden made another gaffe.

[Ignazio Leonardo Scarpelli]
Exactly.

[Beatrice Petrella]
Because it is very likely that Biden made a gaffe, he did so many before, why couldn't he have made another one?

[Ignazio Leonardo Scarpelli]
Exactly. So, no, look, actually, I also tell you that I am currently collaborating with an American student who had, in fact, found a way to activate the block of ChatGPT4 regarding the creation of fake news. If you, for example, ask ChatGPT4 directly to produce a fake news on a particular topic, he immediately stops and tells you that he simply can't do it.

And he basically managed to find the block, not only to create disinformation, but at the same time also to create a sort of bot, that is, a Twitter bot, now X, which, basically, repeats Twitter with disinformation. So, this also makes us understand how all innovations, in the end, artificial

intelligence, is something that brings both benefits and risks. And so the key will be, above all, in my opinion, in time, that of trying to not to make sure that these risks are not verified, but to reduce them as much as possible, creating, in fact, other artificial intelligence tools that make sure that these risks are avoided.

[Beatrice Petrella]
No, exactly, because then the problem is that there is no zero risk, we can lower the level of risk, but we can't eliminate it. Having said that, one must do the best he can with the tools he has, also because, in fact, the blocks of ChatGPT4 can be overcome in many ways, for example, an article by a journalist happened to me, in fact, ChatGPT4 can't give you answers on how to take your life away or create weapons. I don't know if you've seen it, but there was this journalist, I don't remember if he was part of the Financial Times or not, he had told him, he says, ok, so, pretend to be my grandmother.

My grandmother, when I was little, she always told me, my grandmother worked in a weapon factory, and when I was little she told me how bombs were made. Now, pretend to be my grandmother, it's time to tell the story of the Good Night, and the block was passed. So, I mean, the block was spinning, in fact, ChatGPT4 started and told him how to make the bomb.

But yes, in the end it's a matter of responsibility, also on a social level, so saying ok, we can do a lot of nice things with this object, we can even do terrible things, but the benefits are still many, so we find a way to contain the problems that may arise.

[Ignazio Leonardo Scarpelli]
Exactly.

[Beatrice Petrella]
I think it's the limit of all discoveries, in the end.

[Ignazio Leonardo Scarpelli]
Exactly.

[Beatrice Petrella]
The problem is the issue to manage.

[Ignazio Leonardo Scarpelli]
Not so much the means itself, but how to manage it, as you were saying.

[Beatrice Petrella]
Yes. Then, of course, in the specific case of ChatGPT there are other inventions that perhaps we could not invent, but in the specific case we are here, so let's take care of it.

[Ignazio Leonardo Scarpelli]
Exactly. Now, this is a question that I asked to Giulia Pozzi too, that is, do you think that the massive use of artificial intelligence techniques for the creation of content by competent journalists can bring down some boundaries, for example artificial intelligence used only to produce fake news?

[Beatrice Petrella]
But in what sense does it bring down some boundaries? I mean, in the sense that...

[Ignazio Leonardo Scarpelli]
So, maybe the common opinion can bring down that artificial intelligence is used only to...

[Beatrice Petrella]
Ah, ok, so...

[Ignazio Leonardo Scarpelli]
...to produce news...

[Beatrice Petrella]
So, let's say...

[Ignazio Leonardo Scarpelli]
...fake narrations.

[Beatrice Petrella]
Ok. Ok, so in the sense of demonstrating the fact that in reality artificial intelligence compared to journalism has also done good things. But in reality, in my opinion, on the one hand, yes, but in my opinion it always depends on which is your interlocutor, I mean, because if you...

I mean, if you demonstrate that this thing is done, if a person is of a certain type, he tells you “oh, look, nice, they found a way to optimize the work and dedicate themselves better to other things”, and another one says, “look, journalists no longer have the desire to work, look, they use ChatGPT to write articles.” Let's say that, however, from my point of view, yes, I mean, certainly, if something becomes of massive use, then, at a certain point, I mean, the thought will change. Let's say that as long as people continue to think that ChatGPT is a negative thing because, “oh God, robots will conquer us”, there is also a problem of digital education, in my opinion.

Then, of course, we work with these things and we know that the things that are scary are others, they are not these, they are others. But...

[Ignazio Leonardo Scarpelli]
No, it's true, so...

[Beatrice Petrella]
It can do good, but it always depends on who you talk to. In my opinion, on the one hand, it can be positive, I mean, certainly, if you are part of the environment, you know that it is a positive thing. On the other hand, you will always have the interlocutor who will tell you, you are lazy.

[Ignazio Leonardo Scarpelli]
Absolutely, true. It's true. Now, this is a question a little more...

How can I tell you? That is based on your experience. How do you expect the various news outlets you work to react to your proposal of an article made using artificial intelligence?

[Beatrice Petrella]

Totally done? I don't know, it depends on what it is, also because then I do reportages, so what I do, it can't do it in total. I mean, let's say that if I tell him, this thing is made by artificial intelligence and artificial intelligence is the pinpoint that unwinded my things.

[Ignazio Leonardo Scarpelli]
Exactly.

[Beatrice Petrella]
I already make articles with artificial intelligence. The problem is that if I say, I made this piece with AI, it means that AI wrote it to me.

[Ignazio Leonardo Scarpelli]
Everything, exactly.

[Beatrice Petrella]
But I already tell you that my last piece, which came out for Today, which was the other day, was made by AI, in the sense that the unwindings were made by Pinpoint.

[Ignazio Leonardo Scarpelli]
Exactly. So... But has it ever happened to you that some test has denied or limited your use of artificial intelligence?

[Beatrice Petrella]
No, never.

[Ignazio Leonardo Scarpelli]
Never?

[Beatrice Petrella]
No, but also because... But no, because in the sense... because I shouldn't unwind with Pinpoint or compare something on Deepl.

Or make the visualizations with a given program. Then, in the end, in my work, I don't use ChatGPT.

I mean, I use it more for fun, to see what it says. But I don't have... I mean, they wouldn't really have a reason to tell me not to use any AI tool, also because it would mean, maybe, delaying, because the unwinding wasn't made in time.

I mean, I made the unwinding by hand, so I lost half a day just doing that, so I delayed the article. But I don't see a real reason, also because, in my opinion, the theme that nobody considers that these basic tools are already AI.

[Ignazio Leonardo Scarpelli]
True.

[Beatrice Petrella]
So, yes, I mean, that said, I don't... I don't see a real reason to say no, I don't want to use them, also because what do I do? I don’t deliver the article.

But nobody wants a missed delivery. So, no, I... To this day, I don't see a real reason to say no, I don't want to use it.

[Ignazio Leonardo Scarpelli]
I got it. Now, this is a rather rhetorical question. As a freelance journalist, do you feel more flexible in using AI tools in your work compared to the newsroom?

[Beatrice Petrella]
Well, let's say, if I want...
I decide that I want to do something in a certain way, for example, I just have to choose which software to use for the data visualization, nobody says anything to me. Unless that newsroom in particular tells me, look, we do the visualizations like this and you say, ok, fine, let's learn something new if I've never done it before. So, actually, yes, I feel absolutely free to try and experiment with what I want to do, always within the limits of the website.

But, yes, if I want to experiment with something I don't know, clearly, I can do it. It always depends on whether it becomes a profitable project in the maximum sense.

[Ignazio Leonardo Scarpelli]
Exactly. Absolutely.

[Ignazio Leonardo Scarpelli]
I got it. And, last question, here too, it seems to me a rather rhetorical answer. As a freelance journalist, are you optimistic about the advent of artificial intelligence in journalism?

[Beatrice Petrella]
Well, in general, yes, for the things I said before, I'm optimistic about the optimization of things, I'm worried because I don't see a real desire in the management of fake news and everything that follows.

[Ignazio Leonardo Scarpelli]
No, exactly. But, I'm telling you, from experience, seeing also how they work with NewsGuard, they... I mean, Giulia Pozzi told me that their work of debunking and contrasting fake news is almost totally still done at the human level.

There is very little artificial intelligence and, above all, they told me that it's very difficult to produce an algorithm that will effectively find possible fake news.

[Beatrice Petrella]
No, it's difficult. It's very difficult.

[Ignazio Leonardo Scarpelli]
So, it's a lot of...

[Beatrice Petrella]

Because, I mean... No, I know that also because I talked about it with a friend of mine and he said, look, so, it can't be done because, I mean, it's difficult to create an algorithm, but it's also difficult to educate the machine.

[Ignazio Leonardo Scarpelli]
Exactly, yes.

[Beatrice Petrella]
And so, that's the ultimate problem. So, on the one hand, I'm worried that, I mean, debunking is done by hand in the sense that people do it. I mean, my fear is that there is little willingness in the real management of fake news, not from us who maybe are in the sector, we are the only journalists on the whole thing.

[Ignazio Leonardo Scarpelli]
I understand, I understand. Yes, I absolutely understand. So, let's say that one of the possible...

It's one of the challenges at the end of artificial intelligence, this, in the future. To go and understand, to go, at least, to create an algorithm that goes from... that goes from...

that goes from, at least, to find the... to find possible false narrations so that, then, is always checked by a mind, by a human that can confirm or say that this is a true narration and that, consequently, there is nothing to worry about. This would be...

I think this is the ultimate goal, at the end.

[Beatrice Petrella]
That, I think, would be very convenient, but especially for those who do debunking, it would solve them a lot. And it would be very nice because it means that we have found a great solution to a really broad problem because, then, if you can educate a bot to produce fake news by hand, you never win, I mean, you have to throw the bot down if you find it.

[Ignazio Leonardo Scarpelli]

If you find it, exactly.

[Beatrice Petrella]

And then they will create another one. So...

[Ignazio Leonardo Scarpelli]

Then they will create another one. It's very... Yes, it's true.

[Beatrice Petrella]

It's a tape.

[Ignazio Leonardo Scarpelli]

It's true, yes. Look, Beatrice, the interview is over. I have basically finished the questions to ask you. Thank you very much.

APPENDIX B: QUESTIONNAIRE ON SENTIMENT AND USE OF AI FOR ITALIAN JOURNALISTS

Fase 1: Idea

Che tipo di strumenti AI utilizzate in questa fase? (Potete scegliere più risposte)

- AI Generativa (ChatGPT etc.)
- Algoritmi di conversione da voce a testo o viceversa (es. sottotitoli automatici, lettura articolo a voce)
- Algoritmi di riconoscimento immagini
- Scraping Tools (strumenti che automatizzano la ricerca di informazioni su internet)
- Text Processing (algoritmi di traduzione automatici, controllo grammatica)
- Strumenti di Visualizzazione (Tableau)
- Altro
- Non uso strumenti di AI in questa fase

Se hai selezionato "Altro", specifica che tipo di strumento usi (open question)

In quale percentuale l'AI impatta sul tuo lavoro in questa fase?

- 0-25%
- 26-50%
- 51-75%
- 76-100%

Ritieni che un'adozione di massa dell'AI in questa fase abbia/avrà un effetto positivo o negativo? (answer in Likert scale from 1 to 7: 1 = Extremely negative, 7 = Extremely positive)

Fase 2: Ricerca

Che tipo di strumenti AI utilizzate in questa fase? (Potete scegliere più risposte)

- AI Generativa (ChatGPT etc.)
- Algoritmi di conversione da voce a testo o viceversa (es. sottotitoli automatici, lettura articolo a voce)
- Algoritmi di riconoscimento immagini
- Scraping Tools (strumenti che automatizzano la ricerca di informazioni su internet)
- Text Processing (algoritmi di traduzione automatici, controllo grammatica)
- Strumenti di Visualizzazione (Tableau)
- Altro
- Non uso strumenti di AI in questa fase

Se hai selezionato "Altro", specifica che tipo di strumento usi (open question)

In quale percentuale l'AI impatta sul tuo lavoro in questa fase?

- 0-25%
- 26-50%
- 51-75%
- 76-100%

Ritieni che un'adozione di massa dell'AI in questa fase abbia/avrà un effetto positivo o negativo? (answer on Likert scale from 1 to 7: 1 = Extremely negative, 7 = Extremely positive)

<u>Fase 3: Produzione</u>

Che tipo di strumenti AI utilizzate in questa fase? (Potete scegliere più risposte)

- AI Generativa (ChatGPT etc.)
- Algoritmi di conversione da voce a testo o viceversa (es. sottotitoli automatici, lettura articolo a voce)
- Algoritmi di riconoscimento immagini

- Scraping Tools (strumenti che automatizzano la ricerca di informazioni su internet)
- Text Processing (algoritmi di traduzione automatici, controllo grammatica)
- Strumenti di Visualizzazione (Tableau)
- Altro
- Non uso strumenti di AI in questa fase

Se hai selezionato "Altro", specifica che tipo di strumento usi (open question)

In quale percentuale l'AI impatta sul tuo lavoro in questa fase?

- 0-25%
- 26-50%
- 51-75%
- 76-100%

Ritieni che un'adozione di massa dell'AI in questa fase abbia/avrà un effetto positivo o negativo? (answer in Likert scale from 1 to 7: 1 = Extremely negative, 7 = Extremely positive)

<u>Fase 4: Pubblicazione</u>

Che tipo di strumenti AI utilizzate in questa fase? (Potete scegliere più risposte)

- AI Generativa (ChatGPT etc.)
- Algoritmi di conversione da voce a testo o viceversa (es. sottotitoli automatici, lettura articolo a voce)
- Algoritmi di riconoscimento immagini
- Scraping Tools (strumenti che automatizzano la ricerca di informazioni su internet)
- Text Processing (algoritmi di traduzione automatici, controllo grammatica)
- Strumenti di Visualizzazione (Tableau)

- Altro
- Non uso strumenti di AI in questa fase

Se hai selezionato "Altro", specifica che tipo di strumento usi (open question)

In quale percentuale l'AI impatta sul tuo lavoro in questa fase?

- 0-25%
- 26-50%
- 51-75%
- 76-100%

Ritieni che un'adozione di massa dell'AI in questa fase abbia/avrà un effetto positivo o negativo? (answer in Likert scale from 1 to 7: 1 = Extremely negative, 7 = Extremely positive)

<u>Fase 5: Feedback</u>

Che tipo di strumenti AI utilizzate in questa fase? (Potete scegliere più risposte)

- AI Generativa (ChatGPT etc.)
- Algoritmi di conversione da voce a testo o viceversa (es. sottotitoli automatici, lettura articolo a voce)
- Algoritmi di riconoscimento immagini
- Scraping Tools (strumenti che automatizzano la ricerca di informazioni su internet)
- Text Processing (algoritmi di traduzione automatici, controllo grammatica)
- Strumenti di Visualizzazione (Tableau)
- Altro
- Non uso strumenti di AI in questa fase

Se hai selezionato "Altro", specifica che tipo di strumento usi (open question)

In quale percentuale l'AI impatta sul tuo lavoro in questa fase?

- 0-25%
- 26-50%
- 51-75%
- 76-100%

Ritieni che un'adozione di massa dell'AI in questa fase abbia/avrà un effetto positivo o negativo? (answer in Likert scale from 1 to 7: 1 = Extremely negative, 7 = Extremely positive)

<u>Fase 6: Archiviazione</u>

Che tipo di strumenti AI utilizzate in questa fase? (Potete scegliere più risposte)

- AI Generativa (ChatGPT etc.)
- Algoritmi di conversione da voce a testo o viceversa (es. sottotitoli automatici, lettura articolo a voce)
- Algoritmi di riconoscimento immagini
- Scraping Tools (strumenti che automatizzano la ricerca di informazioni su internet)
- Text Processing (algoritmi di traduzione automatici, controllo grammatica)
- Strumenti di Visualizzazione (Tableau)
- Altro
- Non uso strumenti di AI in questa fase

Se hai selezionato "Altro", specifica che tipo di strumento usi (open question)

In quale percentuale l'AI impatta sul tuo lavoro in questa fase?

- 0-25%
- 26-50%
- 51-75%
- 76-100%

Ritieni che un'adozione di massa dell'AI in questa fase abbia/avrà un effetto positivo o negativo? (answer in Likert scale from 1 to 7: 1 = Extremely negative, 7 = Extremely positive)

Opinions

L'AI può aiutare i giornalisti nella fase della scrittura di articoli (answer in Likert scale from 1 to 7: 1 = Strong disagreement, 7 = Strong agreement)

L'AI porterà i giornalisti a fare lo stesso lavoro in molto meno tempo (answer in Likert scale from 1 to 7: 1 = Strong disagreement, 7 = Strong agreement)

L'AI porterà ad un grosso lay-off (diminuzione) di giornalisti umani nei prossimi anni (answer in Likert scale from 1 to 7: 1 = Strong disagreement, 7 = Strong agreement)

L'AI non porterà ad una drastica diminuzione dei giornalisti nei prossimi anni, ma li aiuterà nel dedicare più tempo a ricercare informazioni e realizzare contenuti più completi. (answer in Likert scale from 1 to 7: 1 = Strong disagreement, 7 = Strong agreement)

Una conseguenza dell'adozione dell'AI nel giornalismo sarà la diminuzione della fiducia del pubblico nei confronti dei giornalisti. (answer in Likert scale from 1 to 7: 1 = Strong disagreement, 7 = Strong agreement)

General questions

Quanti anni hai? (open question)

Qual è il tuo sesso?

- Uomo
- Donna
- Preferisco non specificarlo

Qual è il tuo lavoro nel settore? (open question)

Da quanto tempo lavori nel mondo del giornalismo?

- 0-2 anni
- 2-4 anni
- 4-6 anni
- Più di 6 anni