8

Integrating Generative AI into Analytical Practices in Qualitative Inquiry

PhD María Paz Sandín Esteban
Institut de Recerca en Educació (IREUB), Universitat de Barcelona (UB), Spain
mpsandin@ub.edu
https://orcid.org/0000-0003-1836-6126

PhD Angelina Sánchez Martí
Universitat Autònoma de Barcelona, Spain
angelina.sanchez@uab.cat
https://orcid.org/0000-0002-4719-4688

PhD Ruth Vilà Baños
Institut de Recerca en Educació (IREUB), Universitat de Barcelona (UB), Spain
ruth_vila@ub.edu
https://orcid.org/0000-0003-3768-1105

Abstract
This chapter discusses the use of generative AI for qualitative data analysis, highlighting innovative techniques such as inductive coding, sentiment analysis and opinion mining, applied via ATLAS.ti software. It focuses on presenting how CAQDAS programs have currently expanded their analytical tools by integrating AI and, to exemplify these, the chapter analyses the discourse and policy on Artificial Intelligence and education found in official documents from UNESCO, a recognized international authority in the field. Reports of international forums on AI and education in the last five years were scrutinized, in addition to other relevant UNESCO documents on Artificial Intelligence and ethics, providing an overview of current areas of interest and concern and developments within the international education community. The chapter thus offers guidelines for mastering new qualitative analysis tools using AI by critically integrating these procedures into conventional methods.

Keywords: CAQDAS, education, generative AI tools, qualitative analysis, UNESCO.

8.1. Introduction. Overview of AI-assisted qualitative analysis

The advent of generative AI urges us to address the role it occupies (and can occupy) in educational research (and therefore also among the research community), in order to ensure that the question of epistemology and its relationship with methodology is not, once again, neglected within the educational tradition. Undoubtedly, Artificial Intelligence (AI) has great potential in the field of qualitative research, as it can handle large volumes of data, encompassing both explicit information and the more subtle nuances implicit in discourse. The task of assigning codes to relevant quotations in qualitative analysis, a traditionally laborious analytical exercise even with the support of specialized computer aided qualitative analysis software (CAQDAS), has been considerably simplified with advances in AI, especially in technologies incorporating natural language processing (NLP). Recently, CAQDAS has integrated state-of-the-art AI models, such as OpenAI's GPT, streamlining the coding process and the automatic creation of codes by exploring emerging patterns and producing explanatory insights. An example can be found in ATLAS.ti, a program offering AI tools in beta phase (Lopezosa, Codina, Boté-Vericad, 2023) for automatic open and descriptive coding of textual materials (AI responses in this beta phase may be imprecise or take longer).

This chapter discusses the potentials of two specific AI-based inductive coding tools, namely AI Coding and Intentional AI Coding, and two further applications, i.e., sentiment analysis and opinion mining, whose purpose is to identify and quantify emotions and attitudes in text and to examine word patterns, sentence structures and linguistic contexts. To illustrate this, the chapter shows the practical use of the tools in analyzing official documents and blogs by UNESCO and the United Nations focused on education, ethics and AI. The case study underlines the usefulness of the tools for ensuring rigor and quality in qualitative analyses, aligning analytical processes with research objectives. In this way, the chapter provides a comprehensive understanding of experts' and international organizations' recommendations regarding the integration of AI in forms of education that set out to advance towards a more sustainable world.

8.2. The Integration of AI in Computer-assisted Qualitative Data Analysis Software (CAQDAS): the potentials of ATLAS.ti

In this section, we examine the possibilities offered by the integration of AI in ATLAS.ti (Sabariego, Vilà, & Sandín, 2014) by briefly analyzing its tools.

AI summaries

This is a faster way to extract crucial information, simplifying qualitative analysis and obtaining summarized information quickly with OpenAI (Figure 8.1).

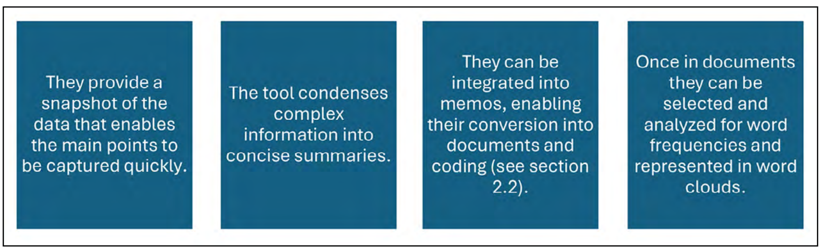


Figure 8.1. Advantages of AI summaries. Source: developed by authors.

AI applications in coding

The applications of AI in encoding documents for analysis are: automated inductive coding (AI coding), intentional coding, and AI-powered code suggestions.

Automated inductive coding: In ATLAS.ti, automated inductive coding, driven by OpenAI GPT models, supports document reading and performs inductive coding automatically. Further review is necessary, of course, in order to refine the automatically generated codes appropriately for the specific study. The suggested process is summarized in Figure 8.2.

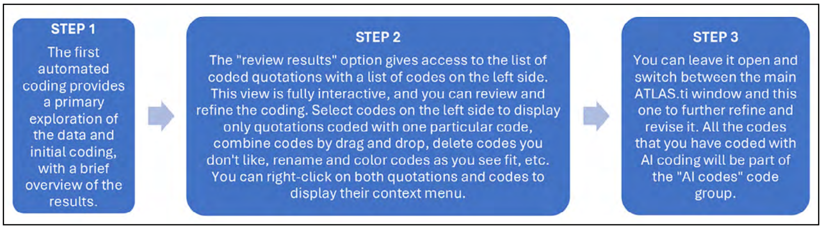


Figure 8.2. *Automated inductive coding process.* Source: developed by authors.

In some cases, GPT models may code for social biases, such as stereotypes or negative feelings towards certain groups, thus careful review of the results is necessary. It is best to submit documents that can combine thematically in the same round of AI coding. Interview questions or participant names in transcripts should be included in their specific response paragraphs. Paragraph structure should be well defined (PDFs can be deficient in this respect). AI Coding skips very short paragraphs, and only parses the plain text of the documents.

Intentional coding: Powered by the technology behind OpenAI's ChatGPT, the AI coding wizard enables us to steer automated coding in the desired direction. It is a tool that allows us to direct AI, obtain codes and explain intentions, concepts of interest and research scope, amongst other features (Figure 8.3).

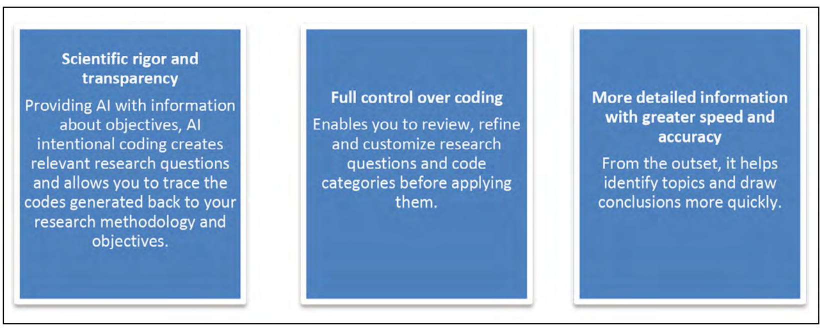


Figure 8.3. *Advantages of intentional coding.* Source: developed by authors.

AI-Powered Code Suggestions: If you prefer to code the data manually and only need a little guidance on how to carry out the coding, the use of suggested codes is recommended. Suggested codes work like AI coding, except that they are applied to a piece of text rather than to the entire document. All suggested codes should be reviewed while creating your own codes (Figure 8.4).

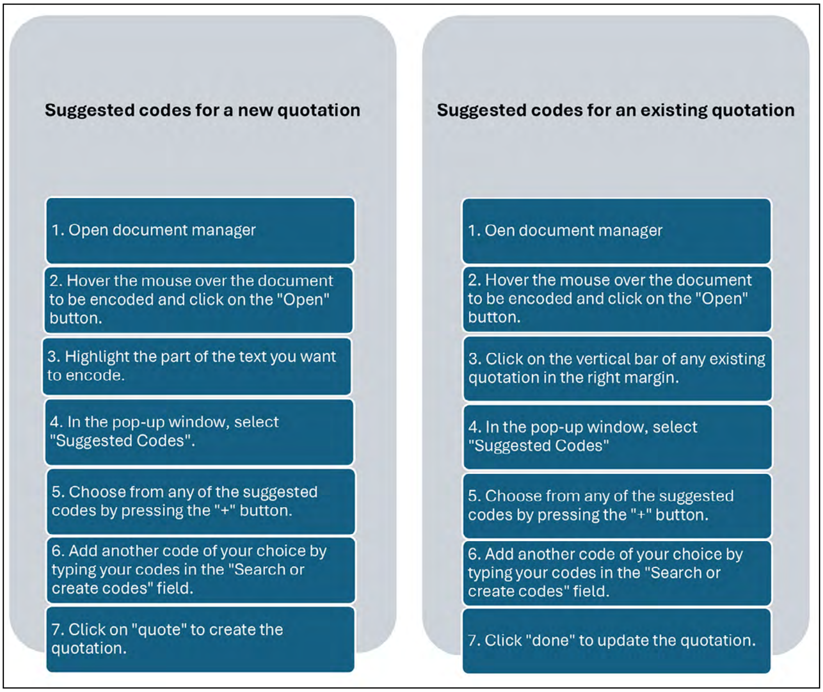


Figure 8.4. AI-powered code suggestions process for new and existing quotations. Source: developed by authors

Conversational AI: chat and interact with documents.

OpenAI's GPT model (used in ChatGPT) utilizes state-of-the-art natural language processing to understand queries contextually, maintain an intelligent dialogue, and provide clarification. Some of the applications of Chatbot AI to qualitative data analysis are shown in Figure 8.5.

AI-assisted sentiment analysis and opinion mining

Sentiment analysis and opinion mining tools are related, but there are fundamental differences in their approaches and objectives. They are often used in tandem to gain a more complete understanding of attitudes and opinions in text analysis, al-

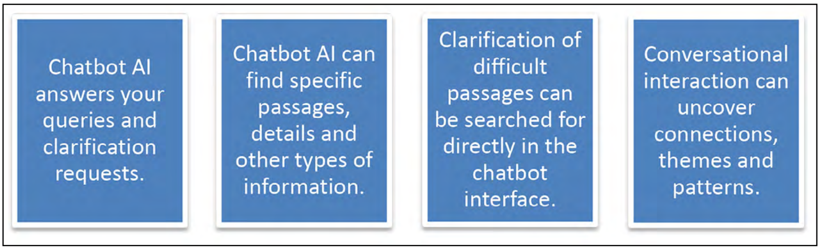


Figure 8.5. Applications of conversational AI to qualitative analysis. Source: developed by authors.

though whether you choose to use them or not depends, as always, on your objectives.

Sentiment analysis focuses specifically on determining the emotional connotations of utterances regarding a topic – i.e., whether the content is positive, neutral or negative – by assessing and quantifying emotional tones, but without delving into more complex aspects of opinion. Opinion mining, on the other hand, enables us to discover and extract opinions and attitudes expressed on a specific topic. Table 8.1 shows some factors that can help researchers decide the best options and the best way to combine them.

Table 8.1. Opinion mining versus sentiment analysis

	OPINION MINING	SENTIMENT ANALYSIS
Definition	Extraction of subjective information from sources such as reviews, comments and opinions expressed in texts.	Focuses specifically on determining the emotional tone of a text, i.e., whether the content is positive, neutral or negative.
Objective	To identify and extract the opinions and attitudes expressed by users on a specific topic.	To evaluate and quantify the emotional tone of the text, without necessarily delving into more complex aspects of the opinion.
Scope	Can cover not only the tone of an opinion (positive, negative, neutral) but also more complex aspects such as the identification of themes, the identification of organizations mentioned and the relationships between different opinions.	Focuses on classifying text into categories of emotional tones and can be part of a broader opinion mining approach.
Applications	For understanding a diverse range of opinions. For deeper analysis.	When tones are sufficient. For simpler applications.

Source: developed by authors.

8.3. Qualitative Analysis Case Study Using Generative AI

The following sections present an example of the use of ATLAS.ti generative AI for inductive coding tools, sentiment analysis and data mining. The case study investigates UNESCO official documents (synthesis reports) on ethics, education and Artificial Intelligence (UNESCO, 2019; 2021; 2022; 2023) and blogs and press releases from both the United Nations (3rd May 2023; 26th July 2023; 7th September 2023; 8th November 2023) and UNESCO (8th June 2023; 16th October 2023), on the same topics.

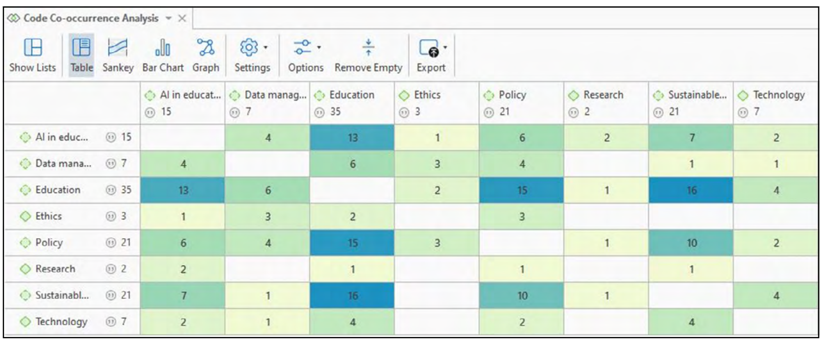
Developing inductive codes and exploratory analysis with assisted AI

The process of applying AI tools to assist in inductive text analysis generally involves three basic steps: induction, organization, andinteraction:

- **Induction**: As noted above, AI-based analysis models read the documents, locate meaningful codes for each data segment and conduct fully inductive coding, presenting an overview of the results. As an example, applying the AI Coding Beta process to the thematic analysis of the report of the Beijing Consensus on AI (UNESCO, 2019), 45 new coded quotations, 145 newly created descriptive codes, the most applied codes (high frequency) and the most co-occurring codes were identified.
- **Organization**: This involves, amongst other things, adjusting the granularity of AI coding results based on your specific needs; eliminating and merging codes; relocating or eliminating quotations; and even discarding codes that, while existing in the quotations, may not be targeted by the study. Thus, in our case, the number of codes was reduced from 145 to 45, following an axial coding process; in the sense, this process of refining the category system sought to discover the most important categories according to: (1) frequency, (2) relevance to the problem statement, and (3) similarities amongst categories. For example, several of the quotations initially in "ed-

ucation" were reassigned to the code "learning"; and quotations from low-frequency codes, such as "learning practices" and "learning outcomes", along with those from codes that were thematically close, such as "personalized learning" or "adaptive learning", were also included in "learning". A further example is that the quotations from "gender equality", "gender gap" and "gender inequality" were added to the code "gender dynamics".

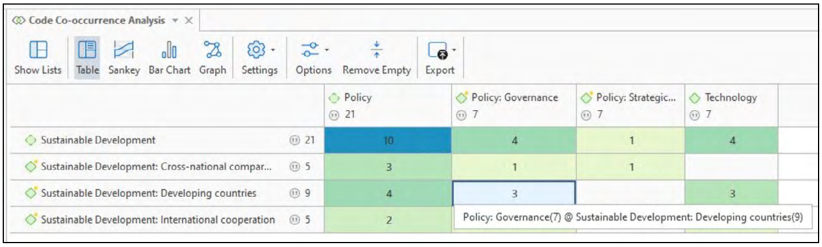
- **Interaction**. Co-occurrence, i.e., when a quotation belongs to more than one code, is the clearest example of relationships and interactions within a data set, and illustrates the richness of qualitative data, which is both complex and multidimensional. AI Coding tells us which topics have the highest co-occurrence of codes. Figure 8.6 shows that, in the case study, these were found at the intersection between "education" and "policy" (15); "education" and "AI in education"; (14) "education" and "sustainable development" (14); "policy" and "AI in education"; (8) and "sustainable development" and "policy" (8).



	AI in educat... 15	Data manag... 7	Education 35	Ethics 3	Policy 21	Research 2	Sustainable... 21	Technology 7
AI in educ... 15		4	13	1	6	2	7	2
Data mana... 7	4		6	3	4		1	1
Education 35	13	6		2	15	1	16	4
Ethics 3	1	3	2		3			
Policy 21	6	4	15	3		1	10	2
Research 2	2		1		1		1	
Sustainabl... 21	7	1	16		10	1		4
Technology 7	2	1	4		2		4	

Figure 8.6. Table of general co-occurrences among codes generated with AI coding in the UNESCO report analysis (2019). Source: developed by authors.

This tool allows us to investigate patterns in the data and suggest hypotheses to deepen the analysis from an exploratory perspective guided by AI automated coding. For example, in Figure 8.7, we extended the analysis window with the details of the subcodes of the association between "sustainable development" and "policy" (8). Thus, we were also able to explore, again by way of example, the relationships between countries' development and governance policies, analyzing the content of the three corresponding quotes from the Beijing forum document, in which experts stated that effective governance policy is crucial to integrating AI in education, facilitating sustainable development by giving rise to a society that is better equipped to face future challenges. Therefore, this procedure helps us make successive inductive, exploratory and thematic analyses of the co-occurrences of interest with the assistance of AI technology.



		Policy 21	Policy: Governance 7	Policy: Strategic... 7	Technology 7
Sustainable Development	21	10	4	1	4
Sustainable Development: Cross-national compar...	5	3	1	1	
Sustainable Development: Developing countries	9	4	3		3
Sustainable Development: International cooperation	5	2			

Figure 8.7. Table of specific co-occurrences for the subcodes of "sustainable development" and "policy" (UNESCO, 2019). Source: developed by authors

Another powerful inductive coding tool offered by ATLAS.ti is intentional AI coding, with which we can indicate to the program our individual research objectives in order to guide the AI application and enhance its efficiency. Figure 8.8, for example, shows the input of our research interests and the context of the study, which framed the analysis of the UNESCO forum report (2019) and yielded auto-generated codes.

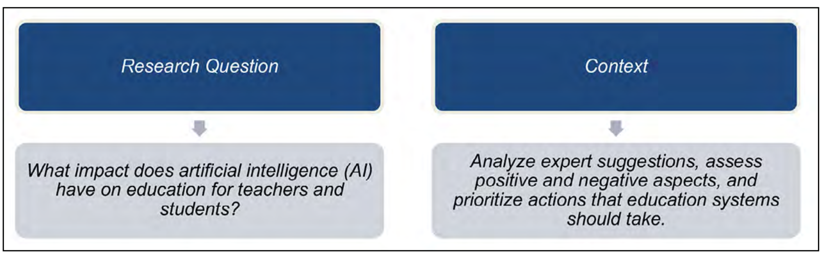


Figure 8.8. Input of intention in ATLAS.ti Intentional AI Coding tool. Source: developed by authors.

The program generated eight questions based on our intention, plus a coding proposal. Before accepting such suggestions, you can add or delete questions and modify codes, thus maintaining control of the coding at all times.

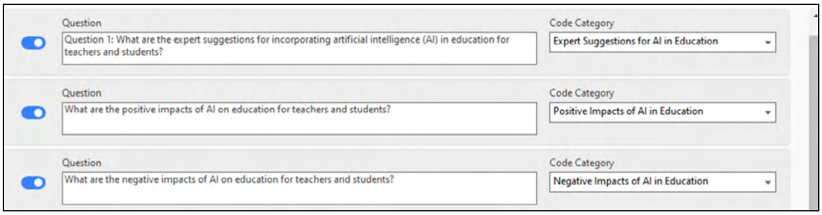


Figure 8.9. Example of questions and code categories proposed for UNESCO (2019), created with intentional IA coding. Source: developed by authors.

With the automatic intentional coding, the result was 121 quotations and 538 codes in eight categories. Again, the analyst's participation in refining the category system offered by AI is important at this point to reduce the number of codes. For example, the codes "ethical concerns", "ethical principles in education" and "ethics" could be subsumed into a single code.

The more context provided for the AI analytical process, the better the results. For example, in our analysis of the four UNESCO documents (2019, 2021, 2022 and 2023), we refined the intention input for intentional AI coding by increasing the number of questions and providing more details about the study, using the instructions below (Figure 8.10). The result was that the location of relevant citations was more accurate and the program provided subcodes with a higher density of quotations (grounded).

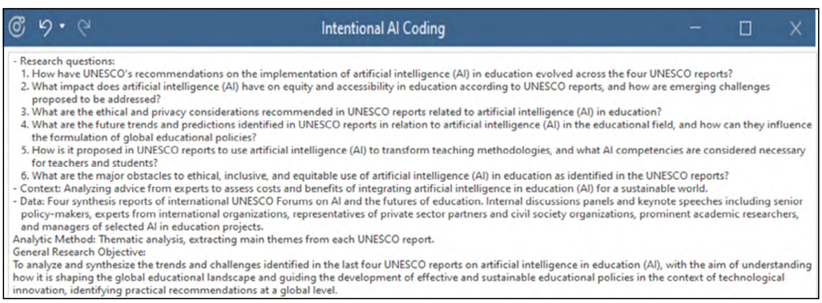


Figure 8.10. Intention in intentional AI coding for analyzing UNESCO forums (2019-2022). Source: developed by authors.

Sentiment analysis and opinion mining

Next, we performed a **sentiment analysis** to obtain an overview of the perceptions and public stance of the United Nations (UN) and UNESCO – as the UN's specialized agency in the field of education, science and culture – regarding the integration and potential of AI in education, identifying the overall tonality of their publications (positive, neutral or negative) and the areas of interest and/or concern in their discourse.

For this analysis, we selected official documents covering a short period of time, since including a temporal approach in this type of analysis can reveal emerging trends, opinion shifts and increases in the importance of key aspects. Once the documents or groups of documents to be analyzed have been chosen, it only remains to choose the base unit for the search and coding (paragraphs or sentences) and to select the type of sentiment or emotional tone to be coded. ATLAS.ti then suggests subcode labels for each sentiment (positive/neutral/negative), but you can rename them or download and install other more complete models.

The analysis tool searches the documents and presents its results through a quotation reader (Figure 8.11), which suggests a sentiment code associated with a sentence or paragraph and allows you to modify the code if necessary. In this case, the analysis yielded 281 quotations, of which 88 corresponded to positive and negative sentiments. ATLAS.ti offers many possibilities for

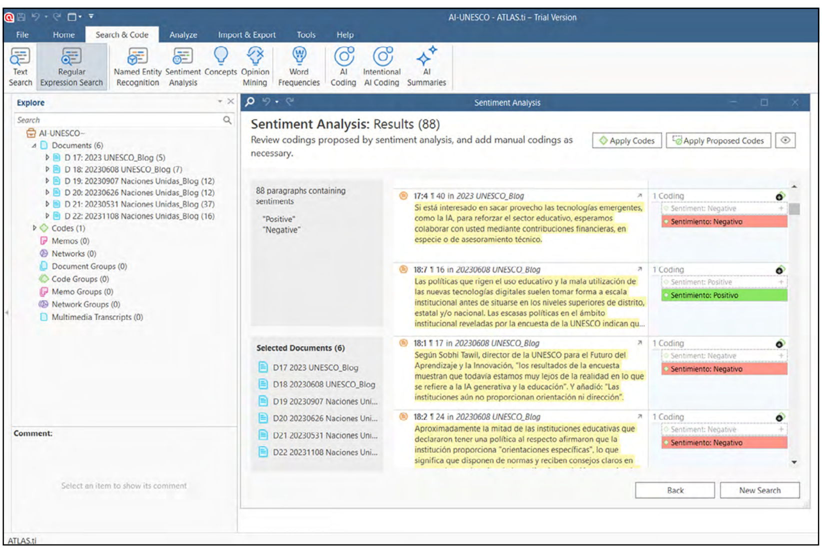


Figure 8.11. Quotations reader after sentiment analysis. Source: developed by authors.

coding. You can, for example, code all the results with one of the suggested codes or with all the suggested codes at the same time. You can also use the regular coding dialog for adding or removing codes.

In this study, we coded all the positive and negative sentiments through a code-by-code review, although automatic coding is also possible. Subsequently, we verified the number of codes identified and assigned colors to each of them in order to make reading clearer and faster, assigning green to positive sentiments and red to negative. In this way, when opening a document, we could see at a glance both the positive and negative sentiments in the compiled news items.

Furthermore, ATLAS.ti enables exploration of the distribution of codes by document, as shown in Figure 8.12. This function was highly useful for exploring the emotional tones in the data – as is usually done when exploring thematic patterns – and interpreting the research results, since it enabled us to visualize the predominance of positive and/or negative feelings in the discourses of the international organizations setting the educational agenda.

To investigate the specific opinions and themes emerging from the positive and negative quotations in more depth, we

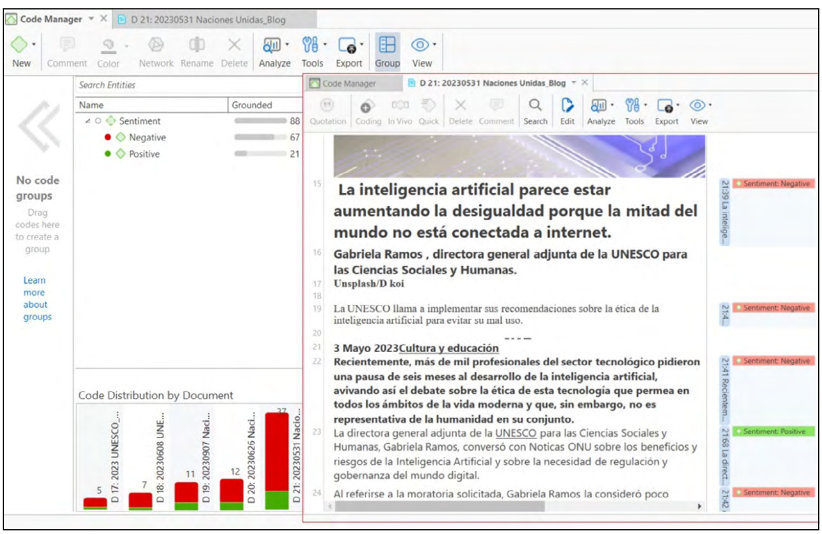


Figure 8.12. Distribution of codes by documents. Source: developed by authors.

used the *opinion mining* tool, which is also extremely useful when combined with a reticular analysis. This qualitative genIA analysis yields a visualization of the most important factors in the positive and negative sentiments, allowing us to review them in context and apply automatic coding.

Opinion mining is a flexible tool that enables analyses on different levels: individual documents, groups of documents, specific codes or sets of codes. Applied to document analysis, it can process one or several documents simultaneously, producing graphs that enable exploration of the results and their text matches.

Once the document or documents have been chosen, ATLAS.ti begins to analyze the data using text analysis algorithms that serve as a starting point for reflecting on the data from different perspectives and theoretical approaches. The aspects detected are displayed in a two-column layout (Figure 8.13): those with positive sentiment in the left column and those with negative in the right. We should also take into account that, since the columns are ordered by the number of occurrences for their respective sentiment, positive and negative results can appear in both columns.

When selecting an aspect, the corresponding text matches are always displayed on the right, and these results can be further

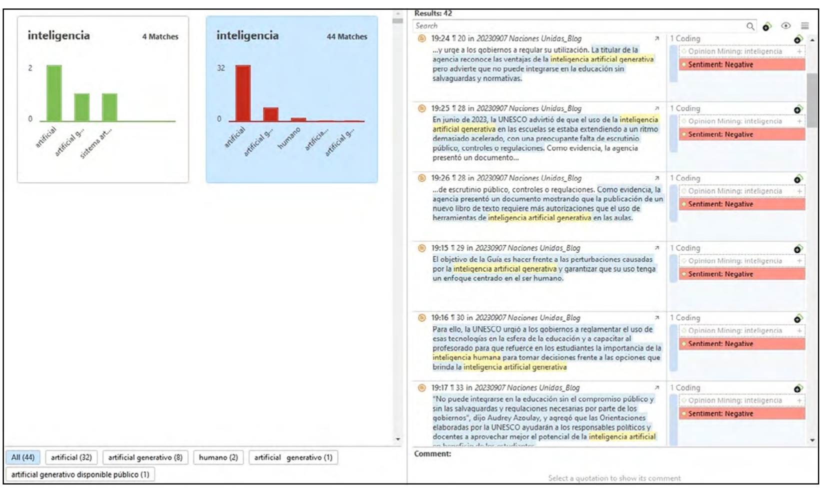


Figure 8.13. AI opinion mining. Source: developed by authors.

filtered by selecting the sentiment modifier themes and coding the text matches progressively. For example, when reading the segments identified as exemplifying a theoretical or descriptive idea of AI, four matches emerged among the positive sentiments. These were references to innovation, diversification and the opening up of possibilities; global collaboration; accelerating the achievement of the SDGs; and efficiency and its associated time savings in research and development. In contrast, among the negative sentiments, 44 coincidences emerged, with references to the challenges and risks associated with AI and its potential unintended consequences, such as a high carbon footprint in model training; ethical issues such as privacy and surveillance; bias and discrimination; security; liability; regulation and supervision; cross-border data flows; unequal data access, and so on. Overall, we found that there was a clear bias towards negative sentiments in the texts, many of which urged UNESCO to focus its discourse on the importance of a consistently human-centered approach in using generative Artificial Intelligence in education, in order to safeguard human rights.

8.4. Critical integration of AI in qualitative analytical approaches

This chapter provides a detailed view of the use of AI in qualitative research, exemplifying its practical application through the analysis of the official discourse of organizations such as UNESCO, which have urged us to pay special attention to the lack of regulation of data use and the ethical issues that we need to address as a society (UNESCO, 2023).

As we have seen, CAQDAS generative AI analytical tools afford advances in line with the most recent and, in many ways, diverse developments in qualitative research. These tools enable us, amongst other aspects, to identify sensitive topics and assess public perceptions (e.g., through sentiment analysis, by detecting topics that generate strong emotional responses; in our case, identifying the international educational community's areas of interest and concern regarding Artificial Intelligence in education); comparison of documents over time (e.g., how perceptions and attitudes towards Artificial Intelligence in education have evolved over time, revealing trends, shifts in opinion and factors of increasing importance); and to transform teaching methodologies and AI competencies (e.g., how it is proposed to transform teaching methodologies and what AI competencies are seen as necessary for teachers and students to prepare them for an increasingly technological and automated world).

However, the integration of AI into qualitative analytical methods is still more strongly related to how we as researchers engage with what we seek to investigate, how we go about the process of investigation, and how we make sense of the knowledge we produce, than anything else. In essence, what we argue here is that the adoption of AI involves strategy and method, and that, while such automation may enable us to work faster, our strategies and methods should still be seen as responsive and always appropriate to the data, not as a set of ready-made AI procedures marking out a predetermined route.

There is no point in using AI simply because it is fashionable or just because it is there. In fact, these are precisely the real risks we run if we get too caught up with the ever-increasing range of AI possibilities, since they can be interpreted as an easy route to qualitative research, without really thinking about what they offer the researcher and the particular study at hand. As with any other method, it is necessary to think about the reasons behind our choosing it. As researchers (trainee or otherwise) we need to be aware of the ever-increasing range of possible methods, learn from them and never stop asking ourselves: Why do I want to use this tool? What kind of data or knowledge can this tool or analytical procedure provide in relation to my research objectives? etc.

We should not lose sight of the fact that the process of validation and development in AI analysis is iterative in the same way as more traditional methods. As researchers move forward in an AI-based analysis, they adjust and improve both its analytical capabilities and the coding and analysis system itself, in the light of feedback and accumulated experience. Thus, we would argue that qualitative analysis occurs in the relationship between the researcher and their research, and although AI can help us facilitate the processes involved, in no case may it replace the researcher. No matter the quantity and quality of the automatically created codes, a researcher should always check their appropriateness.

It is to be expected, for example, that open coding procedures will generate hundreds of codes, single codes will be applied to more than one quotation (hence their large number), and the analyst's refinement of each proposal is therefore essential. The researcher should continue to address the common problems of data quality, such as outliers, information that has no bearing on the research objectives, and the imbalances that often occur in the creation of the category system. It is necessary not to lose sight of how results are generated through AI models and to continue to use constructivist methods – drawn from re-elaboration and abstraction processes – such as constant comparison through open and axial coding (Strauss and Corbin, 2002) or network analyses.

Likewise, it is essential to bear in mind that, despite the advantages offered by these tools, AI models are not infallible and may inherit biases from the data with which they were trained. In addition, human interpretation is necessary to contextualize the results and ensure the validity of the conclusions. Furthermore, before and while disseminating their findings, researchers should ensure that these reflect their specific objectives and conform to scientific rigor. A combination of approaches, using the power of AI together with human research expertise, is still the most effective and ethical way to conduct qualitative analyses.

References

Lopezosa, C., Codina, L., & Boté-Vericad, J.J. (2023). *Testeando ATLAS.ti con OpenAI: hacia un nuevo paradigma para el análisis cualitativo de entrevistas con Inteligencia artificial*. Departamento de Comunicación. Serie DigiDoc. PCUV05/2023

Naciones Unidas (2023, May 3). El debate de la ética de la IA no es tecnológico, sino social. *Noticias Onu*. https://news.un.org/es/interview/2023/05/1520557

Naciones Unidas (2023, July 26). Un informe sobre educación aconseja un uso adecuado de la tecnología en las escuelas. *Noticias Onu*. https://news.un.org/es/story/2023/07/1522972

Naciones Unidas (2023, September 7). Un niño debe tener al menos 13 años para empezar a utilizar la inteligencia artificial en las aulas. *Noticias Onu*. https://news.un.org/es/story/2023/09/1523892

Naciones Unidas (2023, November 8). El poder de la inteligencia artificial y sus desafíos en el marco de las Naciones Unidas. *Centro Regional de Información*. https://unric.org/es/el-debate-de-la-inteligencia-artificial-en-la-onu/

Sabariego, M., Vilà, R., & Sandín, M. P. (2014). El análisis cualitativo de datos con ATLAS.ti. *REIRE. Revista d'Innovació i Recerca en Educació, 7*(2), 119-133, https://raco.cat/index.php/REIRE/article/view/278698.

Strauss, A., & Corbin, J. (2022). *Bases de la investigación cualitativa. Técnicas y procedimientos para desarrollar la teoría fundamentada*. Contus.

UNESCO (2019). Beijing Consensus on Artificial Intelligence and education. Outcome document of the International Conference on Artificial Intelligence and Education. *Planning Education in the AI Era: Lead the Leap*. https://unesdoc.unesco.org/ark:/48223/pf0000368303

UNESCO (2021). International Forum on AI and the Futures of Education. *Developing Competencies for the AI Era*. Synthesis Report. https://unesdoc.unesco.org/ark:/48223/pf0000377251

UNESCO (2022). International Forum on AI and Educatio. *Ensuring AI as a Common Good to Transform Education*. https://unesdoc.unesco.org/ark:/48223/pf0000381226

UNESCO (2023a). International Forum on AI and Education. *Steering AI to Empower Teachers and Transform Teaching*. https://unesdoc.unesco.org/ark:/48223/pf0000386162?locale=en

UNESCO (2023b). *ChatGPT and Artificial Intelligence in Higher Education*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000385146

UNESCO (2023c, June 8). Una encuesta de la UNESCO revela que menos del 10% de las escuelas y universidades disponen de orientaciones formales sobre IA. *Centro de Medios*. https://www.unesco.org/es/articles/una-encuesta-de-la-unesco-revela-que-menos-del-10-de-las-escuelas-y-universidades-disponen-de

UNESCO (2023d, October 16). La inteligencia artificial en la educación. *Centro de Medios*. https://www.unesco.org/es/digital-education/artificial-intelligence