

# USING ARTIFICIAL INTELLIGENCE FOR FORMATTING ACADEMIC JOURNAL PAPERS

**Prof.Khalil A. Yaghi**

Department of Information Science - King Abdulaziz University - Jeddah-Saudi Arabia

[Kaahmad1@kau.edu.sa](mailto:Kaahmad1@kau.edu.sa)

**Abstract:** This paper investigates the challenges researchers face in formatting academic papers, focusing on the importance of proper formatting for clarity, credibility, and reader engagement. While researchers recognize the significance of well-formatted work, they often struggle with the complexities of adhering to diverse journal and conference guidelines. This challenge is exacerbated by the time constraints and the lack of specialized formatting expertise among researchers. This paper explores the potential of Artificial Intelligence (AI) to revolutionize the academic paper formatting process. AI-powered tools can automate many aspects of formatting, including:

- **Adapting to varying formatting requirements:** AI can analyze and apply the specific guidelines of different journals and conferences, ensuring consistency and reducing manual effort.
- **Optimizing font choices:** AI can select appropriate fonts based on readability, visual appeal, and audience, enhancing the overall presentation.
- **Automating margin and spacing adjustments:** AI can ensure consistent margins, line spacing, and indentation, improving document clarity and professional appearance.
- **Assisting with table and figure placement:** AI can analyze the document and suggest optimal positions for tables and figures, enhancing visual flow and reader comprehension.

By automating these tasks, AI can significantly improve the efficiency and accuracy of the formatting process, allowing researchers to focus on their research and enhance the overall quality and impact of their publications.

**Keywords:** Academic Paper Formatting, Artificial Intelligence (AI), AI in Academia, Research Productivity

## 1. Introduction

Scientific research has been important in human society for a long time, and it is usually presented in written reports. These reports include data, methodology, results, and conclusions. Without proper analysis and organization, the collected data is not useful. Written documents are necessary for presenting organized data and information. However, researchers often forget the importance of a well-written report because they are focused on understanding and implementing their research and are short on time. Still, a well-written and organized report is important for keeping the credibility of the entire research effort.

In the fast-paced world of academia, researchers and students are constantly faced with the challenge of formatting their papers according to the specific requirements of different journals and conferences. This can be a time-consuming and frustrating process, especially when there are multiple formatting rules to follow for different types of documents and audiences. However, with the advancement of artificial intelligence (AI) technology, formatting academic papers has become easier and more efficient. AI systems can analyze the formatting requirements of different journals and conferences, and automatically format the paper accordingly. This not only saves researchers and students valuable time and effort, but also ensures that the papers meet the specific formatting guidelines of the target publication. Improving communication skills requires better reports in research. While learning to write is necessary, only a few individuals enjoy it. Additionally, since science is an international effort, English documents are often required. Writing in English when it is not one's first language can be quite challenging and exhausting. Non-native speakers may find themselves frustrated with the language and decide to quickly finish the writing process. As a technical document, scientific research also adheres to the IEEE standard. This standard serves as a useful guideline for well-written research. However, due to the limited number of people familiar with this standard, many research documents do not follow it. As a result, these documents can be unclear, difficult to understand, and lack in quality. (Torres-Valladares et al.2022)(Mammadova2023)(Roxas, 2020)(Vysotska2022) (Sohail et al.2023)(Schoormann et al.2023)

### 1.1 Importance of Proper Formatting

One of the key benefits of using AI for formatting academic papers is its ability to adapt to changing formatting requirements. As mentioned earlier, conferences and journals may have different formatting rules that are constantly evolving. An AI system can easily update its formatting algorithms to keep up with these changes, eliminating the need for researchers to manually adjust their papers every time the formatting requirements are updated. Furthermore, AI can also help researchers who work in multiple subfields with different publication norms. By analyzing the formatting guidelines of each subfield, AI systems can format the paper accordingly, ensuring that it meets the specific requirements of each field. This is particularly useful for researchers who need to submit papers to different journals or conferences with varying formatting rules. In addition, AI can also assist researchers in formatting different types of documents, such as technical reports, proposals, and theses. By analyzing the formatting requirements of each type of

document, AI systems can automatically format the paper according to the specific guidelines, saving researchers time and effort in the formatting process. Overall, the use of AI for formatting academic papers has the potential to revolutionize the way researchers and students approach the formatting process. By automating the formatting task, AI systems can enhance productivity, improve accuracy, and streamline the publication process. As AI technology continues to advance, we can expect to see even more sophisticated formatting tools that will further simplify the process of preparing academic papers for publication. (Alqahtani et al.2023)(John-Mathews et al., 2022)

There are a variety of reasons why researchers emphasize the importance of formatting. One main reason is that they have invested significant time and effort into collecting and analyzing data, and they want their work to be easily comprehensible. Poorly formatted work can be extremely challenging to understand, especially for assistants or students who may need to contact the original author for clarification. Another commonly mentioned reason is that a well-formatted document is more likely to be read. It has been suggested anecdotally that papers with unattractive or difficult layouts are often rejected by journal editors. Regardless of whether this claim holds true, readers are more likely to engage with an article if the formatting is visually appealing. This is especially relevant in the current era of internet-based publishing, where readers are quick to give up on a document if it doesn't immediately capture their interest. Additionally, although somewhat cynical, researchers aspiring to successful careers in academia recognize the importance of appearance. Just as a sharp suit can shape a potential employer's first impression, a well-formatted piece of work can project a favorable image of its author. Considerations have also been given to the implications of formatting when conducting a statistical analysis for a clinical trial or other regulated research. CONSORT guidelines are now widely recognized as the standard approach to reporting a trial, and a recent extension provides specific recommendations for reporting a statistical analysis. These recommendations acknowledge that a well-presented statistical analysis undoubtedly leads to better interpretation and decision making. Peat and Barton discuss the importance of teaching medical students how to write reports; their experience suggests that a well-structured and formatted piece of work improves understanding of the material. Finally, time constraints faced by full-time researchers or those working in the commercial sector, who may need to prepare technical reports or memoranda, can also influence decisions on whether to outsource the work. (Harris, 2022)(Anderson-Cook et al.)(Heßler et al., 2020)(RITZHAUPT and DAWSON2021)(HARTSHORNE & FERDIG)(Maes-Carballo et al.2022)(Zabed & Qi, 2022)

## **1.2 Challenges in Formatting Research**

The complexity of academic document formatting was demonstrated in our interviews with researchers and students who described a wide variety of formatting issues and a lack of clear rules for formatting in their domains. For example, in some fields, conference papers are required to look like journal papers, while in others, the conferences have their own format and attempt to enforce this by specifying page limits with specific instructions regarding font size, column width, line spacing, etc. Changing formatting requirements were also cited. Consider the case of a scientist whose research

spans multiple subfields, each with its own publication norms. Other issues include formatting requirements for different types of documents; a scientist might be required to produce a technical report, a proposal, and a thesis, each with different formatting requirements. The notion of formatting a work for different audiences was also mentioned. A single document might need to be reformatted several times during the course of its life. An automatic formatting system would enhance productivity in these cases by removing the need for the author to focus on formatting and allowing them to concentrate on content. (Pollock et al.2021)(Jaakkola, 2020)

While it is clear that formatting is a very important issue in academic publishing, little systematic research has been done on the topic. In this respect, the situation is similar to the state of affairs regarding software documentation in the 1970s. At that time, it was widely recognized that software was poorly documented, but the software community had to wait until the 1980s for the knowledge-based systems (KBS) community to come along and show the software community how to attack the problem systematically by building intelligent documentation systems. The KBS community built such systems for software by treating the task of producing explanatory text as one of formatting an underlying knowledge representation. Unfortunately, this important work seems to have been forgotten, and it has yet to be translated to the domain of academic writing and publication. An intelligent system for academic document formatting would need to take into account the complex and diverse real-world formatting requirements. It would also have to cope with the fact that authors write in different ways and with different tools, and that formatting is often an iterative process occurring over the lifespan of a document. (Bernacki et al.2021)(Weingart et al., 2021)

## **2. Overview of Artificial Intelligence Applications**

Artificial intelligence has made great strides in the arena of research publications and their formatting. For instance, AI has been utilized in building systems for translating content from one language to another. In recent years, the need for such software has greatly increased due to the internet's capacity for information from all parts of the world. Language translation technology has evolved from simple substitution ciphers to complex rule-based statistical pattern recognition methods to the current neural network-based machine learning methods, which employ vast databases of previously translated material to make inferences about the best way to translate a passage. AI-based translation software is a huge improvement over previous methods since it is capable of translating large amounts of text in a very short time, and the resulting translations are getting closer to that of a human translator. (Bernacki et al.2021)(Park et al., 2020)

Artificial Intelligence (AI), a branch of computer science, is a field where computer systems simulate human intelligence. The ideal computer, as defined by AI scientists, would be able to understand natural language, deduce conclusions, and even understand speech. Therefore, AI is comprised of several different areas of research and development and is still considered to be an emerging technology.

(Salvagno et al., 2023)(Vilone & Longo, 2021)(van et al.2021)

### **2.1 Definition of Artificial Intelligence**

Artificial intelligence (AI) is a branch of computer science that deals with simulating the intellectual abilities of human beings. The term AI is applied when a machine mimics functions that humans associate with other human minds, such as learning and problem-solving. A definition reflecting the different intellectual abilities involved in AI includes the following: Cognitive Science is the study of the cognitive functions of the human mind and their processes; knowledge representation is the study of how to represent knowledge about the world in a way that a computer system can use to solve complex tasks; intelligent systems studies how to create a system that exhibits intelligent behavior. AI is found in many interdisciplinary fields. A good example of this simulation of human intelligence with complex problem-solving is found in some research publication types and automatic categorization of data in the forms of medical diagnosis. AI is used to develop and provide tools that aid in solving problems that would normally require human intelligence. These tools are mostly based on statistical, economic, and mathematical optimization methods. AI research is defined by trying to achieve goals through the modeling of intelligent agents, which is known as taking an approach towards artificial general intelligence. (Korteling et al.2021)(Fitria, 2021)(Jarrahi et al., 2023)

### **2.2 Role of AI in Research Publication**

Artificial intelligence has been used in nearly every type of application that one can imagine. Some of the applications are artificial intelligence in the field of consumer electronics, entertainment, heavy work, and many others. But one type of application that is very useful and cuts down time for many of the computer science researchers is research publication. In research publication, there are many ways that both the author and editor can use AI to make their job easier, whether it be for formatting a document or checking the document for possible errors. One way, for example, that authors can use AI in research publication is by using the software to format the document for the publication. As most authors tend to submit documents in different formats, some being non-compatible with the system that they are trying to upload it to, a reformatting process may be necessary. An author can cut down time by using AI software that will accurately and quickly format the document to the desired specifications. An editor of a publication can use AI to automate the process of checking a submitted document for possible errors. Though there are software in existence today that check for grammatical errors, these software are still far inferior to what the future holds for AI document error checking. In the future, editors can use advanced AI that can comprehensively read through a document and check it for errors through a complex series of algorithms. (Gezici & Tarhan, 2022)(Ghioni et al., 2023)(Baviskar et al., 2021)

### **3. Font Type Optimization**

The study did two experiments to see the effect of the font type used in publication in article writing on the computer. 28 people were recruited to read an article and make corrections to it. The first experiment involved them reading an article in a good font such as Georgia and then later read another article in a bad font such as Haettenschweiler. The second experiment involved the same process but using different fonts in italics. An analysis on the number of corrections made showed that the articles in non-italic fonts

were not significantly different; however, a lower number of corrections were found on the good font compared to the bad font. 10% fewer corrections were made on articles using a good font compared to bad fonts in the contribution to layout and readability. (Krivec et al.2020)(Wallace et al.2020)

Selecting the right font for a document is pivotal to how the reader perceives the publication. Fonts can convey different meanings and often the choice of a particular font depends on the context and target audience of the document. For example, a children's publication may use a more informal and 'fun' font, whereas an academic journal would use a more formal, easy to read, and clear font such as Times New Roman. Times New Roman is commonly used in publications since it is clear, easy to read, and generally a 'safe' option. However, the use of AI may be able to offer better alternatives and optimize the font type to improve readability and aesthetics. Tiresias was a font designed for people with visual impairment but was found to be better for all readers as compared to Arial and Times New Roman in small font size. (Richardson, 2022)

### **3.1 Selecting Appropriate Fonts**

Although somewhat subjective, selecting appropriate typefaces for achieving a particular effect in a publication is a critical decision. The kind of typeface used should reflect the theme of the report or document. It should also be digested easily by the target audience. AI can assist in this process by creating typographical models which can interpret existing text and extract any meta-information which it can find on the document's intended style or audience. This meta-information would be transformed into a set of constraints which is to be used when searching or creating a typeface. AI can also be used to generate typefaces or font modifications which obey certain constraints. For example, fonts which minimally distort letter shapes in order to enhance recognition by dyslexic or elderly readers. AI techniques would also guide technically demanding processes, where a good solution is difficult to find or where a human typographer might struggle to understand the best method. An example of this would be a font designed to optimally compact lots of information into a small space without sacrificing readability. By encoding fonts as a set of actions to be performed on data structures, it is also possible to use AI techniques from the field of automatic program synthesis to induce algorithms, from examples, for modifying the font. An example of this would be to learn how to produce the most aesthetically pleasing bold or italic version of a font. (Longoni & Cian, 2022)(Abdollahi & Pradhan, 2021)(Uzir et al.2021)(MacLellan and Koedinger2022)

### **3.2 Ensuring Consistency in Font Usage**

In the case of publications intended for print, authors must additionally ensure that the fonts used have been properly embedded into the document. Failure to do so can result in potentially time-consuming edits and changes when it comes time to reprint or resubmit a revised version of the publication. The latency and inconvenience involved in both finding matching fonts and reapplying formatting styles with word processing software can be wholly avoided if the document was written in LaTeX. But despite its advantages, the widespread use of word processing software for document writing means that it is inevitable that there will be occasions where authors have written and formatted

a document only to be forced to use a different word processor at a later date. This is particularly the case with collaborative works. As different word processors come with different default settings and font sets, one can expect that the original formatting style of the document will be compromised when it is opened and edited on an incompatible system. An author might not realize that a colleague's revisions to a document have disrupted the intended formatting style until much later, and potentially not until after the document has been submitted for publication. Overall, ensuring that a document's intended font usage is faithfully preserved for its final version can be a difficult task with many potential pitfalls. How exactly the problem should be addressed is a topic barely touched by previous research in automated document formatting, and so is an issue with much potential for AI solution. (Hegghammer, 2022)(Khan et al., 2023)(Grabar & Saggion, 2022)

### **3.3 Enhancing Readability with AI**

Machines are being developed and tested to automate the whole process, to cite the works appropriately. One of the pioneering works in this field is an online citation project, which is an implementation of an algorithmic decision-making process to learn the style of a manual annotator. Once the machine learning process is complete, new references can be input by the user and all formatting for that reference will be generated by the system. This implementation can potentially save an enormous amount of time for those writing academic papers. Future work will involve extending this ideation to create algorithms that find the references themselves and input them in raw form, requiring no input from the user. At the same time, regarding the current system of academic publications, AI has a lot of potential to improve the quality of life for researchers. An example of a common problem is post-acceptance of a paper; the authors often get a notification from the publisher stating their paper was not formatted according to the specifications of the publication. This means the authors must format the paper again to fix errors, or upload an edited version to the publication, requiring more work. This is a frustrating situation where authors just want to be done with it and move on to their next project. An ideal setup for the future, the publisher would use an AI system to automatically format the paper and if errors are found, correct them, all with no notification to the author. This could also be applied to formatting done by printing companies, where formatting errors can often result in a paper being rejected for publication. It is evident that this type of automated system will require less work from the author and improve efficiency in the publication process. (Mahmud et al.2022)(Rodgers et al.2023)

### **4. Margins and Spacing**

Margins in a document are defined as the white space that surrounds the text on all sides. Margins serve numerous purposes. They create a visual buffer around the text, making the document easier to read, whilst also creating a home for written comments and notes. For these reasons, among others, most journals require that you submit a manuscript with 1-inch margins. Word automatically sets a default 1-inch margin; however, many times this margin is increased or decreased by the user. Incorrectly set margins can have a large impact on the layout of your document, and consequently the page limits of your

document. When a manuscript is submitted to a journal, it is quite often returned to the author with a request to change the margin size, font type, or line spacing in order to adhere to the journal guidelines. This is a process which can be quite time-consuming and frustrating for the author. Automatic correction of these simple formatting issues could save the author precious time and could reduce the queuing time for a document to be processed into the publishing stage. For the author, this means a quicker publishing time for their work. AI could easily be programmed to read the desired margin size of a selected journal and to change this setting if necessary. This would ensure that the margin size is correct right from the beginning of the writing process and would develop a document which is consistent in terms of margins. A similar process could be used for the line indent, a reference feature which is often not used correctly.

#### **4.1 Setting Optimal Margins**

In many academic publications, one can find articles that are formatted where the left and right margins are considerably small. This leads to text lines which are longer than required and difficult to read. While it is possible to set individual margins in a word processor, this doesn't guarantee that the document will look the same on a different machine. Ideally, we aim to have a certain amount of space between the text and the edge of the page. This is known as a relative margin, and is the settings that LaTeX interprets. Unfortunately, while there are options for setting the margins, there is little advice to help the user decide what is a good margin. The package layout and also the geometry package provide such an option in that they have a showframe option which is useful for visualising the effects of the margin settings. However, in the absence of any advice, many users will still continue to set the margins too wide or too narrow. (Dalalah and Dalalah2023)

#### **4.2 Adjusting Line Spacing and Indentation**

There are times when research papers are submitted to research databases or even processed by publication houses that will need to be returned to authors for a whole set of revisions, often technical in nature. Here we address the problem of automatic server-side tools being developed to provide authors with LaTeX for reformatting documents to fit style. In such situations, publishers may find it more expedient to return reformatted MS Word documents to the authors rather than instructing authors on how to manually change the structure of their documents. This can occur when authors have difficulty reaching a certain formatting style and this information is simply lost in translation to the LaTeX source or is too difficult to access. An obvious solution in these cases is to directly reformat the MS Word document. If artificial intelligence research can provide better solutions than are currently available, it is only logical that such methods be used to facilitate reformatting MS Word documents in the easiest ways possible. In many cases, the process of applying new formatting styles can be hindered by old commands leftover from the previous formatting, which force unwanted formatting changes. This is an obstacle when the locations of such commands are not known by the document's author. AI research for identifying and changing specific instances of a formatting command can be highly effective in the case of word processors with a linear document structure. However, research directly targeting specific instances of formatting commands is



becoming outdated, as the common trend is to define formatting styles with stylesheets that are found in XML documents. AI research is also present in defining the most efficient methods for finding and changing specific instances of formatting commands, but we will not touch this subject as the methods involved are not significantly different from the general case. (Mahlow and Piotrowski2022)(Eden, 2024)(Yeo et al., 2022)

#### **4.3 Utilizing AI for Consistent Spacing**

Finally, the Find function will take care of the indentation issue, where many different indentation methods may be utilized, it is helpful to have them all be uniform and easily changeable. The first step in achieving these goals is to eliminate extraneous newlines and carriage returns, effectively making the document into one large paragraph. Next, we will search for specific instances where double spaces are used to end sentences. By replacing all occurrences of double spaces with single spaces, and continuing to do so until no match is found, the spaces between sentences will be uniform. After deciding what a sufficient space between sentences is, a global find and replace using `^p^p` will set the space between paragraphs, and a find and replace or directly changing the indentation settings will set the uniform indent.

The second issue the Find function tries to address is the use of newline characters to achieve spacing between paragraphs. By default, Word adds a 10-point space after each paragraph, which is done by adding the newlines, this doesn't hold true for all published articles. Newline characters act similarly to spaces at the end of sentences. During the editing process, text can be justified, shifted around, or otherwise altered causing the spacing between paragraphs to change.

Apart from selecting a font style and size, AI could be beneficial formatting the document before publishing, especially regarding the consistency of spacing. Proper spacing is a clear indicator of a credible source, when using double spacing for simple editing purposes it is a great tool for note taking when reviewing the document with a pen. Although using double spacing may be simple and appealing, it isn't always necessary. In journals for instance, double spacing is not utilized. (Kousha & Thelwall, 2024)(Filetti et al., 2024)(Khabib, 2022)

#### **5. Tables and Figures**

The old adage "a picture says a thousand words" holds true with figures in text. They're an extremely effective way to convey a concept, but even more care should be taken with figure design than table design, since the complexity of the picture can make it difficult to understand. Ideally, authors should aim to present their data in well-designed tables. But when this is just not feasible, quite often the best way to clarify the data is to use the table function to format the data and create a simple table-like layout. This might seem like doing the same thing twice, but usually it will be an exercise in enhancing table clarity without the need to spell out data duplications in the text. Using the same keystrokes as starting a standard table, type the data into the table function - using the tab key to navigate between cells. This will draw lines in the format of the typed data, and

---

tabulated data can be formatted without the need to adjust cell sizes. A borderless table can also be made by deleting all borders after table completion.

In scientific writing, tables and figures play a pivotal role in not only improving the comprehension level but also how appealing a publication is. However, these elements do not come without challenges. Most authors find it difficult to present their data clearly and effectively in table form. Often, improving the lexical complexity of a table is compromised in the process and the worst-case scenario results in data loss. This happens when authors attempt to convert their table to a different style to try and improve readability - quite often the data becomes confused. Another common problem is when tables are imported as image files rather than being composed with the word processor's table drawing tools. These images can lose quality (resolution) during conversion between file types, and also when enlarged to fit the width of a page. (Hand et al., 2021)(Cotos et al., 2020)(Kim & Kim, 2022)

### **5.1 Formatting Tables for Clarity**

Repetition of material and poorly planned presentations can lead to confusion and frustration in readers who seek a quick and clear understanding of the data and its major implications. Nowhere is this more common than in the presentation of tables. Tables should be constructed so that the details included are easy to see and understand without reference to the text. An effective table is one that presents the data clearly and simply so that the reader knows exactly what you found. Tables should be able to stand alone and be completely understandable. When planning a table, consider the following. Is the data better presented numerically or in another form? If numbers are appropriate, then the table is the best presentation. If one or a few simple points are to be made, the data may be better presented as a bulleted list within the text. If there is a trend or pattern in the data, that is better illustrated in a graph. Tables that display the results of a particular statistical analysis (e.g. regression table) are aimed at a specialist audience and may need to be more complex. The level of detail should affect the table's placement within the text. (Cloutier & Ravasi, 2021)

### **5.2 Aligning Figures with Text**

By defining structures based on common format types for life sciences, behavioral sciences, physical sciences, and clinical medical sciences, and allowing elements to be edited and exported to MS Word, the SmartTools will help authors produce publication-quality tables. SmartFigure will offer the author a platform in which to generate publication-quality figures. Each figure type (e.g., bar chart, flow diagram, protein structure) can be created using defined frameworks to ensure the figure represents the data as clearly as possible. Once the figure has been created, AI algorithms will suggest the best position in the document for this figure to be placed.

Using artificial intelligence, it is possible to test multiple figure locations and store the position which gives the clearest visual impact. This not only saves the author time but also produces a higher quality visual output. This technique has been implemented in the SAGE Authoring tools, SmartTables and SmartFigure. These are web-based applications

that help authors create high-quality tables and figures using simple and predefined frameworks to maximize graphical representation.

When preparing a document for publication, most authors are familiar with the content and structure of their document but are less confident with the layout of their tables and figures. Poor alignment of figures within a document can detract from the content and negatively impact the reader's understanding and interpretation of the research. This can, in part, be attributed to most word processors treating figures as "an island" with text flowing around it. This lack of control means that when small changes are made to the document, it is possible that the figure can move quite dramatically from its initial position.

### **5.3 AI-Assisted Table and Figure Placement**

There has been a little work on computer methods for constructing specific types of diagram from a formalized statement of its semantics, together with research into diagram and graph grammars and methods of drawing. This is basically an exercise in well-constrained solving transfer, and the good news is that there is a reasonable prospect of success in these narrow domains. It would clearly be very advantageous to automatically generate a pathway diagram or a circuit from a textual statement and then be able to regenerate the diagram as a revised version of the underlying science emerges. But this sort of AI figures are remote from the general mainstream of publication.

Technology for separating data and layout is a small first step, but the next really big advance in automating the process of publication is going to come from intelligent software that not only prepares good looking figures, but understands what's in them, can link them effectively to the text and can generate them from data, or modify their content, in response to changes in the text or in the underlying science.

## **6. Citations and References**

Machines can cite papers too. With the increasing emphasis on publications in academic evaluation, accurately citing previous work has become a crucial aspect in research writing. However, citation formats obviously differ between various journal specifications, as well as between different online databases that publish information about the same article. As a result, many authors have encountered hindrances in attempting to correctly and efficiently reformat their references and citations. Furthermore, most find it being a very tedious task, especially if you have hundreds of references when they are writing a review paper. To combat this, online citation generators have become an increasingly popular resource. These web-based tools only require that you input the necessary reference information and select a standard citation style, following which they proceed to generate the correct citation and even the abstract when available. Though many of these tools can be effective, the overall accuracy in reference/citation generation is still less than 100%. For example, if the original information input was incorrect or the article was published in a non-standard journal, the citation generated may not be the expected result. Also, it is not an infrequent occurrence where a generator cannot locate the specific article in its database. As a last

resort, the author may sometimes choose to just search for an existing citation of the same article hosted online and manually edit it, leaving more room for error. One way or another, the more time spent at this phase means less time available for actual writing and research. Therefore, an automated and efficient solution to managing citations and references is still highly sought after today. (Birkle et al.2020)(Hiebl, 2023)

### **6.1 Ensuring Accuracy in Citations**

The proper formatting and accuracy of references are crucial for adequate representation of cited authors' work, and for the credibility of the referencing party. AI programs should be designed to support students and researchers in their academic ventures, promoting learning and providing features that alert users to errors in their work. Simple 'point and click' programs will not suffice for complex academic documents. Many citation errors are due to manual entry of reference details. AI programs should provide import functions from internet databases or scanned documents to minimize user entry. Steps should also be taken to standardize reference formats, so that data entered from any source, in any style, is converted into one accepted form. This will relieve the user of the burden of familiarizing themselves with multiple different referencing styles, which many find a taxing and confusing endeavor. Citation and location of online references is an increasingly prevalent issue. Traditional academic referencing styles do provide methods of citation for internet sources, but as the internet is a rapidly evolving entity, it is difficult for referencing styles to remain consistent and details for every possible source type are often not provided. AI referencing programs should be designed to accommodate a wide range of online document types and provide the user with the most current and applicable method of citation. Considering the voluminous amount of information available on the internet, there is a potential for future AI programs to automatically locate and cite online references. An advanced feature such as this for electronic documents would be a great timesaver and prevent many mistakes resulting from misattributed online sources. (Baviskar et al., 2021)(Gruzauskas & Ragavan, 2020)(Okunlaya et al., 2022)

### **6.2 Formatting References According to Guidelines**

Sequence-to-sequence algorithm has been used to change the reference format. Here, the input and target both are the references, but are in different formats. The input reference is tokenized and words are embedded. This is fed into the encoder LSTM, which generates a context vector for the decoder LSTM. The output of the decoder LSTM is the reference in the other format. This output is embedded to obtain the result.

This is the final stage in the publication process. At the time of submission, the documents are required to be submitted as per the journal guidelines. Usually, it is mandatory to submit references in a particular format such as APA, Harvard, Chicago, or MLA. One has to manually edit each of the references to change it to a particular format. This is a very cumbersome task. Also, interchanging one reference format to the other is very difficult. We have used sequence-to-sequence learning techniques to change the reference format. This is a neural network model where the input is a sequence of words of one type of reference and the output is a sequence of words in different formats. We have learned the mapping from the dataset of all types of reference formats. This

technique will greatly help the authors to change reference formats and also the publishers who publish articles and are required to submit references matching their house style. This task is also internally related to the summarization of the reference, where the reference is required to be summarized in 300 words in any format. Steps involved in reference formatting have been illustrated below:

### **6.3 AI Tools for Efficient Citation Management**

Using artificial intelligence, researchers are now able to automatically generate citations in a format compliant with the given formatting guidelines. Research work often involves preparation of references, citations, and bibliographies. The accuracy of the references is critical as this can often be the difference between a favorable mark and a criticism for lack of scholarship. Automatic generation of citations and bibliographies may not appear to be a significant achievement, but on a large document with careful and meticulous attention to properly formatted references, it can be a time and sanity-saving function. An example of a more powerful system for managing bibliographic references is that provided by the Sente word processor. This provides intelligent import from JSTOR, PDFs, and other academic databases and allows the user to sort and categorize the imported references. Similar systems to this can gather metadata from a PDF for automatic indexing and search for articles. Reference management tools are software packages used by scholars and authors to record and use bibliographic citations. This can be done manually; i.e. typing out the details of a specific citation on a document for future reference, or using the various reference import/export features on databases and online libraries. These details are then used to organize bibliographies in article writing and to make future citing of sources easier. The onset of digitally based article writing has greatly increased the volume of references extracted from resources to the extent that it is now important to economize the process of formatting references and citations. This is now usually done using special purpose reference and citation management software often in the form of a web application. A more recent development is web-based tools for reference managing within online libraries: these allow for automatic export and population of bibliographic data on online articles.

## **7. Abstract and Keywords**

One of the most critical components of a research manuscript is the abstract, since it is a self-contained piece of writing that can be understood individually from the article it describes. The research should be disseminated to a wider audience than the typical one of highly specialized scholars, since research often involves public money or money coming from public institutions. Often times, this is taxpayers' money and the general public has a right to know how this money is being spent. The abstract should give them this sort of an 'executive summary' of your work. This is the research's opportunity to persuade the random reader that the work is notable and important. Also, abstracts are useful because the researcher will not always have the time to read a full article. They can act as a roadmap for the article, providing the reader with knowledge on what lies ahead. Lastly, abstracts make it easier for librarians to categorize academic works. With the constant expansion of academic literature, no single library can hold even a major

fraction of the scholarly works available at a given time. Abstracts allow librarians to understand quickly what an article is about so that they may categorize it and know where to locate it later. An effective abstract will help researchers quickly locate articles pertinent to their research.

### **7.1 Crafting an Effective Abstract**

Writing a good research paper also requires writing a good abstract. A well-constructed abstract enables the reader to identify the basic content of a document quickly and accurately, to determine its relevance to their interests, and thus to decide whether to read the document. The abstract should be about the research, not about the act of writing. In other words, it should identify the main purpose of the study, methodology of the study, and major findings. And avoid including any direct citations in the abstract. According to studies in the Journal of Biological Chemistry and Food Science, the reason why the requirement for an abstract in those two different disciplines is still the same (i.e. to identify main objectives and findings). But different fields might have slightly different standards in writing abstracts. Take, for example, a research paper in the field of social science or psychology. The abstract might have a longer length, but it still serves the same purpose. Then, in a different context, the research about poverty which is published by the International Food Policy Research has specific instructions in writing abstracts. Among the guidelines are writing the motivation, strategic result, and policy implication in summary, but no citation in the abstract. In conclusion, there is no absolute standard in writing an abstract. The most important thing is that it should fulfill the requirements of future readers, and since the abstract is the most read section after the title, it is really important to make a good one. Despite the many differences, the strategy to make a good abstract is typical, and there are some kinds of common guidance that can be implemented.

### **7.2 Identifying Relevant Keywords**

Clearly, it is very possible that the quality and manner in which research is posted online can vary from researcher to researcher. However, given the undoubtedly large amount of online research present and increasing, we must consider that there is and will continue to be data with great potential to benefit the same research the data was pulled from. An example of this could be health-related information obtained from a health survey website. This creates what we can call a secondary data situation where new research is conducted to analyze and draw medical conclusions from the previously obtained data. This can very easily contrast with primary research based upon original ideas, experiments, and observations. High detail primary and secondary research papers in any field can tangibly be improved by either more directly presenting the conclusions drawn from the data, or comparing and contrasting with the new data. In any of these cases, one can imagine a future where various types of research will also need better conveying due to the gap from its complexity and high importance to today's traditional scholarly works.

One consequence of this is that traditional scholarly works will become less important when compared to the actual research information posted online. Currently, many researchers post information such as PDF documents which are merely online versions

of their traditional scholarly research. In a study performed by Martín-Martín et al., it was tested the attitudes and behaviors of Spanish scholars surrounding online self-archiving of research. It was discovered that many humanities researchers show little self-archiving of their research, and instead opt for constructing an online curriculum vitae. This is a prime example of an instance where researchers may not effectively communicate the findings of their research papers.

In the scientific community, publishing research papers in an academic journal is an important way for researchers to share their works. Academic journal papers act as valid and vital information resources for researchers in different fields of research. In the process of writing a paper, the researcher has to provide a high level of detail and describe their research thoroughly. This can be difficult, as research papers in their original form are good at presenting the "forest" of results, but sometimes the "trees" get lost. To provide a solution to this, on an article detail page in Cochrane Library, there is a section in the right-hand column of the page which contains a glossary of the Altmetric Attention Score and Article page views. Each term links to the relevant Altmetric details page allowing readers to learn more about the online attention received by an article through tracking Altmetric. This is a prime example of how online attention can be linked to an increase in traditional scholarly research. This can be a very useful tool as it can help researchers understand the online impact of their research. However, it is important to note that by using automatic methods such as data mining, we are moving closer and closer to the day where no additional promotion of the online research will be needed. (Horbach, 2020)(Moher et al.2020)

### **7.3 AI-Based Abstract and Keyword Suggestions**

Enduring the pressure of providing a clear representation of the research work with an optimal information, abstracts conform a crucial segment of a research article. The abstract serves to "sell" your work, and hence plays a key role in this respect. Given this, writing an effective abstract is essential in preparing a successful research article. Currently, crafting an abstract remains a manual process. Our first step with respect to abstracts is to investigate automated generation of a structured abstract. We will first develop a classifier to identify whether a given paragraph needs an abstract summarising it. If this is positive, we will work on generating the abstract. This could be done using extractive summarisation i.e. selecting sentences from the document, and abstractive summarisation where facts are conveyed rephrased i.e. using synonyms, or possibly written in a new way based on the author's intent of the sentence. The latter is a more difficult task, and might involve training a language model, or learning from example abstracts alongside the document. A key benefit of an abstractive summary is that often, facts recorded in different parts of the article might be brought together in the summary, hence a new level of information might be conveyed. In the longer term, especially for large datasets, it is possible that all this could be done in an unsupervised manner. Given a massive amount of unstructured research work, and no information of relevance, there may be a need for an AI system that could sift through this data, and decide whether an abstract of certain document is required. This is an interesting problem in itself. An AI system may need to be trained on modern and old English, and also various languages in

existence today. If the system can identify this information with high accuracy, then an abstract could be compiled automatically and combined with the main document in a knowledge-based manner. We can only foresee that as the web evolves, and information becomes more abundant the demand for such an AI system will increase. A tool that could be used in conjunction with this, to highlight the content that has been summarised would also be desirable.

## **8. Title and Author Information**

In journal articles, it is always a struggle to make authors adhere to guidelines for title and author information despite the fact that it is the most visible component of a paper. A title has the ability to attract the right readers to your paper and convey the most important information of your study. When articles are indexed in databases, a good title can mean the difference between someone deciding to read your full article or passing it by. However, titles often fall short of this purpose by being too vague, too general, with a poor word choice, or using a question format. Refining a title to maximize these goals can be a time-consuming and difficult process. Similarly, when submitting manuscripts to journals, it is often the case that author and affiliation information is not kept in a digital form but rather input manually. This can be a time-consuming process and is error-prone. This can result in the information being recorded in a way that makes it hard to associate specific authors with specific affiliations, for example if authors have multiple affiliations or if the same affiliation is used by multiple authors. This may lead to confusion or inconsistencies when the published article is viewed. It would be useful to be able to automatically generate author and affiliation information formatted according to journal style guidelines. This would free authors from having to perform this task and would ensure a higher quality of recorded information. (Purwanto et al.2021)(Lê & Schmid, 2022)

### **8.1 Optimizing Title for Impact**

Creating an efficient title that accurately reflects the content of the article and draws in potential readers is crucial. Readers go through databases skimming for articles and typically only read those titles that are interesting and a good reflection of the content. Missed articles represent lost potential impact. Despite the importance of an effective title, the medical literature contains many poorly conceived titles. In a study on clinical trial reports, the authors found that only 19% of the titles gave a clear indication of the report's content and only 55% conveyed the conclusion. Title quality for articles in the biomedical informatics taxonomy was evaluated, revealing that editorial and opinion papers had on average significantly higher title quality than other article types. The authors recommended that teachers of research writing and young researchers consider the process of distilling thoughts into a few words as an intellectual challenge essential to communication of ideas. This advises that article titles are not of low priority simply because they are of low difficulty.

### **8.2 Formatting Author Names and Affiliations**

The author name normalization step is aimed at creating a highly customized and comprehensive list of author names based on the names utilized at the forefront of the



papers. This critical process extends to the extraction of intricate author affiliation details from the acknowledgment section. In both cases, the ultimate objective is to curate a meticulously crafted collection of unique author names, closely intertwined with their corresponding affiliations. This meticulous curation effectively reduces the occurrence of duplicate information entry. Such measures are necessary due to the pervasive nature of authors utilizing diverse forms of their names when publishing different works. For instance, an author might adopt variations such as "Christopher M. Bishop" or simply "C. Bishop." Moored in reality, the affiliation details also undergo changes over time, making it entirely unrealistic to expect data inputters to consistently enter the variations exactly the same way each time.

Consequently, such inconsistencies typically result in different author names being associated with the same individual, or even worse, inaccurate author names being erroneously assigned to affiliation details. To counteract this conundrum, a simple yet highly effective string matching algorithm, known as the Jaccard Coefficient, is judiciously employed. This algorithm serves as the cornerstone for comparing the similarity between strings across a pair of author names. In cases where a high degree of resemblance is detected, subsequent decisions are made to determine whether these similar names indeed refer to the same person. An identical approach is employed when comparing affiliation details.

In instances where two author names that are deemed similar share a set of co-authored papers, the program diligently allocates the affiliation details entered under one of the names to every paper where the other name is cited. This particular strategy proves invaluable in cases where an author transitions between affiliations during the course of their illustrious career. Furthermore, if an author name is detected but possesses no corresponding affiliation details, it is promptly associated with a "NIL" affiliation, indicating the absence of an explicit affiliation. The entire process yields a vast multitude of revisions to the existing roster of author names and affiliations within the papers. To systematically monitor and potentially reapply these changes to an updated version of the data, an author name and affiliation ID is diligently generated for every unique combination of values. This robust system ensures seamless tracking of changes and facilitates the vital verification of new information against preexisting data. This indispensable facet of the project remains paramount, as the inherent uncertainties introduced by various AI methods oftentimes give rise to significant alterations in the data. The ability to validate and cross-reference new information with its historical counterpart stands as a testament to the integrity and reliability of the project.

### **8.3 AI-Powered Title and Author Information Generation**

The spinoff from the creative neural network researcher from INIST-CNRS and University Joseph Fourier (UJF) Grenoble, France create robola acronym of bot-based article elaboration which is a part of the development RPM that has the basic model, author-defined abstract, and a specific example on clinical trials of anticoagulation for patients suffering from acute coronary syndrome, in 2009. This robot can perform data

processing automatic title an abstract generated from the study and write up an explanation with a thorough way using a specific example.

Writing title and author information in general, not how authors can do their job quickly, but the authors are required to automatically do data processing at a specific target. Moreover, there are some authors who have difficulty specifying a title that contains all the essential points of the study and there is no writing a title that is not in accordance with the actual study. Besides, it's not easy to make an author's identity, showed a resume not just a list of authors.

One part of the article is the title and author information. In writing this information, there are some predetermined rules generally authors are required to meet, depending on the journal. Writing a good title according to ICMJE (International Committee of Medical Journal Editors) contains a title that identifies the problem, the population or samples that become the study subjects, the methods used, and the key result of the study. Meanwhile, writing author information aims to show the identity of the author can be seen at a glance.

Given the limited amount of time that researchers have, an OIE can be said to regard the structure as a constraint. To cope with this problem, a research article is divided into several parts. This division is assumed to enable the reader to effectively understand the content, while a writer will more easily adhere to the theme of the study to be written. (Gastel & Day, 2022)(Cargill & O'Connor, 2021)

## **9. Proofreading and Grammar Check**

Proofreading is the key to achieving error-free research papers. It is common to find some mistakes, whether they are related to formatting, citations, information, or language. It is important for researchers who are not native English speakers to put in extra effort in proofreading to ensure the paper is error-free. There are several automated software and applications to check for grammar and spelling errors. However, most of these tools are not efficient in handling complex sentence structures and professional terminology used in research papers. AI-powered grammar and spelling checking tools are trained with vast datasets and are capable of identifying errors in complex sentences and terminology. For specific domain terminology, some AI tools allow us to add words to the tools' custom dictionary so they won't be identified as errors in subsequent checks. With AI tools, researchers can easily identify and correct grammar and spelling errors compared to traditional manual proofreading. Language quality is very important in research papers. Good language quality and writing styles can effectively convey information and messages to readers. There are certain rules and guidelines for writing high-quality papers. AI tools are trained with specific rules and guidelines for various types of writing styles. Some AI tools are able to paraphrase to enhance writing styles. AI-powered paraphrasing can change sentence structures, change sentences to a more formal/simple style, or sometimes rewrite sentences with different words while conveying the same meaning. Although paraphrasing is not perfect and sometimes produces awkward sentences, the tools still provide alternatives for sentences that may be beneficial for researchers.

### **9.1 Importance of Proofreading in Research**

Proofreading is a simple process, but it is very important to do this to ensure that you produce a quality document. It is often the last task performed before a document is submitted. When you proofread, you are carefully checking a document for any mistakes that you may have made. This includes grammar errors, punctuation errors, spelling errors, and also errors in the way that your text is worded. It is important to catch all the mistakes that you may have made because even a simple mistake can make you lose credibility. In the research field, it is quite obvious that a document filled with mistakes is not worth much. When people submit manuscripts to be published, they first often seek peers to proofread their document. Even after this has been done, it is still likely that mistakes have been made. Studies have even found that experienced proofreaders are only 85% accurate in finding and correcting errors in a document. With this in mind, proofreading is definitely a vital step in creating a quality document. As for grammar checking, this is often done in conjunction with proofreading and is also important to ensure that a quality document is produced.

### **9.2 AI Tools for Grammar and Spelling Correction**

Advanced AI tools have made it possible for grammar checking and correction to be done entirely with automated processes. It does more than what a word processor's grammar checker can do. Modern tools can identify context-based errors. They have capabilities of understanding the meaning of the word in the sentence to find out if the word is the right one and provide alternative words to the possibly misplaced word. Popular grammar checking tools include Grammarly and the grammar checker from the Virtual Writing Tutor. Grammarly is a highly sophisticated tool that is able to check for various errors in a sentence. Its features for word choice and vocabulary enhancement are backed by a vast amount of research in identifying errors and the best possible replacements for the word. However, some tools may not be error-free, and humans will still find errors that the tool could have missed. This is when the intelligent human mind still triumphs over the AI. But with the fact that AI is still in development in understanding human languages, it is only a matter of time before these tools become error-free.

### **9.3 Enhancing Language Quality with Artificial Intelligence**

The biggest factor that distinguishes human language from those of other creatures is its diversity and complexity. It often takes an individual several years to master their own native tongue, and during that time they are constantly learning and adapting to new rules and idiosyncrasies that are present in the language. In a second language, many people never reach the full fluency that is present in their native language simply because of the complexity of the language. That is why it is important to have tools available to help aid language learning and to improve communication through language. Native language grammar checkers are a special case of language learning aiding tools that suggest better alternatives for what the user is attempting to communicate. Grammar checkers have largely been rule-based systems in the past. This is because the disambiguation problem that occurs when trying to correct a user's input is significantly easier to solve using rule-based systems. Rule-based systems, however, are usually not robust enough to solve disambiguation problems in natural language processing. This causes some rule-based

grammar checkers to incorrectly "correct" a user's input without understanding the full meaning of the input. With natural language processing AI, grammar checkers can reach a new level of understanding by training AI with input and corrections from humans to learn to correctly solve disambiguation problems. This will allow grammar checkers to more accurately correct a user's input by understanding the context of the input. An additional method for improving native language is to use a writing assistant who is a non-native speaker of the language. Writing assistants rely on pattern recognition and association with similar phrases that are present in context. This is a perfect place to implement AI because it can aid writing assistants in understanding the best way to communicate their phrases and when and why these phrases are used. By using training from a large set of native language text, writing assistants can make suggestions or explain to the user the best way to communicate their intended message. This can greatly improve communication in many non-native speakers and the readers of their texts. (Gastel & Day, 2022)(Tenopir et al.2020)(da et al.2022)(Ramírez-Castañeda, 2020)

### **10. Compliance with Journal Guidelines**

An article that deviates from a journal's prescribed formatting requirements is suboptimal to both the author and the publisher. In current AI prepress systems, generic rule sets are used to attempt to define the requirements of individual journals. The effectiveness of this approach is openly questionable. Authors are often unsure about what these requirements actually are. By working directly from the searchable database of over 3000 journal titles maintained by the Elsevier Fingerprint Engine, our software can show authors exactly what the requirements are for their target journal. Each journal has a set of rules which largely defines the layout of an article submitted to it. These rules are usually not explicitly stated to the author. Traditional rule sets have generally been a one-size-fits-all solution in which an often incomplete set of rules is specified in a code. This has two main drawbacks. Firstly, a new set of rules must be defined for each different journal despite overlap in requirements between journals. Secondly, rules tend to change over time and the upkeep of a large collection of rule sets is an impractical task. By only using these rule sets, AI systems cannot be certain that the rules were followed at the time of article submission and are of no help to an author who should have followed a rule that is not specified in the set. Our approach involves a sophisticated classification of articles and individual page elements using machine learning algorithms to directly infer the rules by which decisions were made. This allows greater flexibility in rule definition and detects when rules change over time. The rules are directly inferred from real examples and therefore robustly cover all cases. When a rule is no longer followed, common detection of this is possible. This will mean that the author can be told exactly what rules were not adhered to. An article that deviates from a journal's prescribed formatting requirements is suboptimal to both the author and the publisher. In current AI prepress systems, generic rule sets are used to attempt to define the requirements of individual journals. The effectiveness of this approach is openly questionable. Authors are often unsure about what these requirements actually are, leading to potential issues and confusion during the article submission process. However, with our groundbreaking software, we provide a groundbreaking solution that revolutionizes the way authors

interact with journal formatting requirements. By working directly from the extensive and comprehensive searchable database of over 3000 journal titles maintained by the renowned Elsevier Fingerprint Engine, our state-of-the-art software can show authors exactly what the specific and tailored requirements are for their target journal.

It is crucial to understand that each journal has a set of rules and guidelines that largely defines the layout and structure of an article submitted to it. Unfortunately, these rules are usually not explicitly communicated to the author, resulting in uncertainty and inefficiency. Traditional rule sets, which have emerged as a one-size-fits-all solution, are often incomplete and specified in a code, making them insufficient for addressing the diverse requirements of different journals. This outdated approach leads to two significant drawbacks. Firstly, a new set of rules must be defined for each different journal, despite the potential overlap in requirements between journals. This redundant effort not only wastes time and resources but also creates unnecessary complexity. Secondly, rules tend to evolve and change over time, making the manual upkeep of a large collection of rule sets an impractical task. To overcome these limitations and provide authors with a seamless and efficient experience, our innovative approach utilizes sophisticated machine learning algorithms to classify articles and individual page elements. By directly inferring the rules by which decisions were made, our software offers greater flexibility and adaptability in rule definition. It also detects subtle changes in rules over time, ensuring that authors stay up-to-date with the latest formatting requirements. Importantly, these rules are derived from real examples, making them robust and reliable in covering all cases. With our advanced system, authors can bid farewell to the uncertainty and frustration of not knowing whether their article meets the journal's formatting requirements. Our software individually analyzes each article, providing specific feedback on any rule or guideline that has not been adhered to. This comprehensive analysis empowers authors to make the necessary adjustments and ensure their work aligns perfectly with the expectations of their target journal, ultimately leading to a more successful publication process.

In conclusion, our cutting-edge software transforms the way authors navigate journal formatting requirements. By harnessing the power of machine learning and real-world examples, we offer unrivaled accuracy and efficiency in defining the necessary rules for each individual article. Experience the future of publishing with our groundbreaking solution and eliminate the guesswork in formatting your next journal submission. (Morton et al.2022)(Gastel & Day, 2022)(Horbach, 2020)(da et al.2022)(Ramírez-Castañeda, 2020)(Homolak et al., 2020)(Baas et al.2020)

### **10.1 Understanding Journal Formatting Requirements**

In the world of scholarly publishing, the formatting of academic journal papers is a critical component that can often make or break a successful submission. Journal editors and reviewers have specific requirements for how manuscripts should be formatted, including aspects such as margins, font size, spacing, citations, and reference lists. Failure

to adhere to these guidelines can result in a rejection of the paper, even if the content is of high quality. Traditionally, authors have had to meticulously follow journal formatting guidelines manually, which can be time-consuming and error-prone. However, with the advancement of artificial intelligence (AI) technology, there is now the potential to streamline this process and ensure that papers meet the required formatting standards. AI can play a crucial role in assisting authors with formatting their papers according to the specific requirements of a journal. For example, AI algorithms can analyze the text of a manuscript and automatically suggest changes to ensure that it conforms to the journal's guidelines. This can include adjusting margins, formatting citations and references, and ensuring consistency in style throughout the document. One of the key benefits of using AI for formatting academic papers is the potential to save time for authors. By automating the process of formatting, authors can focus more on the content of their paper and less on the technical details of formatting. This can ultimately lead to a more efficient and productive writing process. Furthermore, AI can also help improve the quality of published articles by reducing the likelihood of formatting errors. By flagging inconsistencies or discrepancies in formatting, AI can help authors identify and correct mistakes before submitting their paper for review. This can lead to a higher standard of published research and enhance the overall credibility of academic journals.

Scholarly journals are complex documents that require a specific format. Before the widespread use of word processors, the Aries Systems Corporation's Editorial Manager (EM) was the software tool of choice for manuscript submission and peer-review in many scientific journals. Although much of the process of replication has moved from author-formatted Word files to publisher-tagged XML, the basic ethos of authoring to produce a document that looks like the published journal article remains. This requires the author to closely attend to the details of the journal's expected format. Such attention is not well supported by word processors, because the space and page limitations of print journals and the structural elements to be displayed within those limitations are not well represented in authoring software. A common problem is authors' mis-estimation of lengths of their draft and figures; when it is discovered at the typesetting stage that a paper accepted for publication is 25% longer than allotted, a substantial re-edit often becomes necessary. AI-based methods to guide authors in formatting to the constraints of their target journal have the potential to both save time and improve the quality of published articles.

Overall, the use of AI for formatting academic journal papers has the potential to revolutionize the publishing process by streamlining the formatting process, saving time for authors, and improving the quality of published articles. As AI technology continues to advance, we can expect to see even more sophisticated tools and solutions that further enhance the efficiency and effectiveness of scholarly publishing.

## **10.2 AI-Based Compliance Checking**

In recent years, automatic compliance with citation and referencing style has been made possible through Elsevier's Journal Article Tag Suite (JATS) and Extensible Markup Language (XML), allowing for easier and more accurate electronic submission.

However, for text formatting and structure, AI is required to automate the detection and correction of errors. While still in its infancy, software such as Referee and Seek & Analyze are developing programs designed to automate the outlining of scientific documents and comparing them to a list of predefined rules. Similarly, document preparation systems such as LaTeX already have error-checking software for formatting. However, the next step is to develop AI that can automatically correct these errors. With continued development, it may be possible for AI to simply format a document according to what the author specifies as an intended journal, learning from examples in that journal and identifying a set of rules to adhere to based on the similarities.

### **10.3 Ensuring Adherence to Peer-Reviewed Journal Standards**

A peer-reviewed journal entails maintaining the highest standards in publishing. It is a necessity that must be maintained in a scientific publication. The most thorough research can be dismissed if the accompanying manuscript is poorly produced or does not comply with a journal's specific guidelines. Research in this area has led to a burgeoning of journal guidelines that researchers must follow in order to have their work published. These guidelines are used by the reviewers and editors to assess the quality of the work being proposed. Often, the decision will revolve around whether the researchers' findings have sufficiently addressed and resolved an issue. If it is difficult for reviewers to ascertain what has been done, or the significance of the findings, publication is often rejected. Adherence to guidelines can dictate the formatting of a paper and the content within it. Some guidelines are specifically aimed at formatting, but the majority are concerned with the transparency in representation of the work and a smooth flow from introduction through to conclusions. Often, there is a dual column layout requirement with specific sections that must be easily identified. We plan to address the automated formatting issues mentioned earlier, in the instance where specific guidelines can be identified and the formatting can be altered automatically.

## **11. Conclusion**

As we have discussed in this paper about using AI for research publication formatting, we can see that by implementing artificial intelligence application in research publication formatting, we can solve many problems that researchers face during publishing their work. AI system can convert unformatted references into BibTex in order to generate an ordered bibliography. It is done by parsing the unformatted references into different fields like author, title, journal, volume, year, publisher, etc. There are huge benefits by using AI and a lot of prospects for these systems in the near future. Even though there are many challenges in this type of research, researchers are trying to overcome them by combining both AI and human intervention in the loop. I don't know how convincing the explanation was, but definitely the AI system will generate much better formatted research paper in the future. The step towards generating intelligent research paper has already begun. And we can say, it will not take much longer to see its application in the real world.

### **11.1 Benefits of AI in Research Publication Formatting**

AI technology, when used in research work, can provide various means for checking and updating references. The most comprehensive is the feature provided by RefWorks, but

other software may provide limited means for achieving similar ends. Some web browsers (including Mozilla Firefox, Internet Explorer and Google Chrome) have plug-ins for importing references into a user's reference library. Mendeley is another web-based tool and desktop application for managing references, which can also facilitate reference collection during web browsing. AI technology can be used to cross-reference or validate the accuracy of a reference, whether it be in a reference library or within the body of the paper itself. Some validation programs use data from an online service to cross-check the reference and provide information on the published source. Inconsistent use of references can be automatically corrected, replacing incorrect or omitted data with the correct details for that reference. This can be done with the use of a 'Find and Replace' tool, although more sophisticated software can actually 'understand' the contextual meaning and grammatical tense of a user's writing, automatically formatting the reference to conform to the chosen style from a drop-down menu. This latter function can also be applied to formatting of the paper itself. For the manual author, the submission of a revised manuscript can be a very daunting task when a journal's formatting criteria are not adhered to in the initial preparation. This is often due to differing style guidelines between journals, and absence or ambiguity of instruction on the author's part. As a fully automated computer task, reformatting requires minimal human input. A style mapping tool can identify formatting traits in the original document. This is then compared to a style template for the desired format, and a list of instructions for each formatting change is generated. The currently available tools would create a macro in a Microsoft Word document. Using the simple example of italics and bold formatting, this would use the 'find' function to search for 'italicized text,' then change it to the defined italics style from the given template. This is accomplished by selecting the most appropriate action from a reformatting syntax, based on pattern matching and probability-ranked algorithms. The rate of author error in reformatting guidelines has been shown to be as high as 40%, so it is of great benefit that an AI system will second guess an author's incorrect action and offer an 'undo' function for each formatting change applied. An alternate automated method for formatting a paper is through an AI system that processes and stores the content of the paper itself. It then analyzes the content against a knowledge base of style and formatting rules, updating the paper at the content level. This is a more intelligent processing operation, as it will make formatting decisions on complex or special case instances of text. An AI system capable of natural language processing will communicate with the author about any ambiguous instructions, helping to associate a formatting choice with the most appropriate area of text. A tool of this nature will provide a coherent formatted output that is characteristic of the input content and is more adaptable in instances where the selected style guidelines are subject to change.

### **11.2 Future Prospects and Challenges**

Now, many researchers are showing interest in this new technology for publishing their articles. AI reduces a great effort of publication by formatting it according to the journals' guidelines and also it is cost effective. It is good for the journals too, it reduces the burden



of formatting the accepted articles and they are assuring that the article is aligned with their guidelines. However, there are still many areas where AI can be used in a better way for formatting a publication. There are very few AI tools for this task which can be accessed free of cost. Many open source AI software are available, but they are not that much efficient in formatting the papers. As it is said, while creating a tool itself developers have to put great effort to teach that how to format a paper. AI formatting tool should understand the structure of the paper in great detail like where about and references should be placed. It should be able to divide all the data into several categories like tables, diagrams, equations and text and should format all of these in a very well manner. Providing the detailed formatting instructions is also a similar kind of work. So, this task is quite difficult and challenging. But it's not impossible, considering the rate of advancements in AI, it can do this in a better way in the near future. A sophisticated AI publication formatting tool should reduce the necessity of manual intervention at the last stage. It should provide a ready to publish article to the author. Today the PDF is a widely accepted format for publishing an article. AI can create a well-formatted article in PDF form without any errors. This entire process should not take a long time, which is a case with present tools. AI should reduce the time of formatting an article, the lesser the cost, it will be more beneficial. And the journal's expectations of getting a well-formatted article linked with the author's claim can be fulfilled. So, AI has a bright future in this area.

#### **References:**

Torres-Valladares, D., Ballinas-García, E., Villarreal-Reyes, J., Morales-Álvarez, V., & Ortiz-del-Ánge, C. (2022). Research papers 101: The do's and don'ts of scientific writing. *Materials Today: Proceedings*, 48, 107-114. [\[HTML\]](#)

Mammadova, T. (2023). Pedagogy of Writing and Digital Writing. In *Academic Writing and Information Literacy Instruction in Digital Environments: A Complementary Approach* (pp. 67-105). Cham: Springer International Publishing. [\[HTML\]](#)

Roxas, M. J. D. (2020). Exploring senior high school students' academic writing difficulties: Towards an academic writing model. *IOER International multidisciplinary research journal*. [researchgate.net](https://www.researchgate.net)

Vysotska, O. (2022). WRITING SKILLS: CHALLENGES AND SOLUTIONS. *Наукові записки Національного університету «Острозька академія»: Серія «Філологія»*, (13 (81)), 208-211. [oa.edu.ua](https://oa.edu.ua)

Sohail, S. S., Farhat, F., Himeur, Y., Nadeem, M., Madsen, D. Ø., Singh, Y., ... & Mansoor, W. (2023). Decoding ChatGPT: a taxonomy of existing research, current challenges, and possible future directions. *Journal of King Saud University-Computer and Information Sciences*, 101675. [sciencedirect.com](https://www.sciencedirect.com)

---

Schoormann, T., Strobel, G., Möller, F., Petrik, D., & Zschech, P. (2023). Artificial intelligence for sustainability—a systematic review of information systems literature. *Communications of the Association for Information Systems*, 52(1), 8. [researchgate.net](https://www.researchgate.net)

Alqahtani, T., Badreldin, H. A., Alrashed, M., Alshaya, A. I., Alghamdi, S. S., bin Saleh, K., ... & Albekairy, A. M. (2023). The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research. *Research in Social and Administrative Pharmacy*. [researchgate.net](https://www.researchgate.net)

John-Mathews, J. M., Cardon, D., & Balagué, C. (2022). From reality to world. A critical perspective on AI fairness. *Journal of Business Ethics*. [springer.com](https://www.springer.com)

Harris, A. J. L. (2022). Guidelines on manuscript format, structure, and style: avoiding editorial holdups in the publication process. *Bulletin of Volcanology*. [springer.com](https://www.springer.com)

Anderson-Cook, C. M., Lu, L., Gramacy, R. B., Jones-Farmer, L. A., Montgomery, D. C., & Woodall, W. H. Publishing an applied statistics paper: Guidance and advice from editors. *Quality and Reliability Engineering International*. [\[HTML\]](#)

Heßler, N., Rottmann, M., & Ziegler, A. (2020). Empirical analysis of the text structure of original research articles in medical journals. *PloS one*. [plos.org](https://www.plos.org)

RITZHAUPT, A. D., & DAWSON, K. (2021). Publishing in the *Journal of Research on Technology in Education*: An overview of the aims, processes, and editorial suggestions. *What Journal Editors Wish Authors Knew About Academic Publishing*, 93. [\[HTML\]](#)

HARTSHORNE, R., & FERDIG, R. E. The *Journal of Technology and Teacher Education*: An Overview with Publication Tips and Resources for Prospective Authors. *What Journal Editors Wish Authors Knew About Academic Publishing*, 101. [\[HTML\]](#)

Maes-Carballo, M., Khan, K. S., & Cano-Ibáñez, N. (2022). Essentials on Writing a Paper for a Peer Reviewed Journal. *Essential Writing, Communication and Narrative Skills for Medical Scientists Before and After the COVID Era*, 45-57. [\[HTML\]](#)

Zabed, H. M. & Qi, X. (2022). Plagiarism and Referencing Techniques Used in Social Research Report. *Principles of Social Research Methodology*. [\[HTML\]](#)

Pollock, D., Davies, E. L., Peters, M. D., Tricco, A. C., Alexander, L., McNerney, P., ... & Munn, Z. (2021). Undertaking a scoping review: A practical guide for nursing and midwifery students, clinicians, researchers, and academics. *Journal of advanced nursing*, 77(4), 2102-2113. [wiley.com](https://www.wiley.com)

Jaakkola, E. (2020). Designing conceptual articles: four approaches. *AMS review*. [springer.com](https://www.springer.com)

Bernacki, M. L., Greene, M. J., & Lobczowski, N. G. (2021). A systematic review of research on personalized learning: Personalized by whom, to what, how, and for what purpose (s)?. *Educational Psychology Review*, 33(4), 1675-1715. [nsf.gov](https://www.nsf.gov)

Weingart, P., Joubert, M., & Connoway, K. (2021). Public engagement with science—Origins, motives and impact in academic literature and science policy. *PloS one*. [plos.org](https://doi.org/10.1371/journal.pone.0241111)

Park, W., Wu, J. Y., & Erduran, S. (2020). The nature of STEM disciplines in the science education standards documents from the USA, Korea and Taiwan: Focusing on disciplinary aims, values and .... *Science & Education*. [springer.com](https://doi.org/10.1007/s11269-020-00000-0)

Salvagno, M., Taccone, F. S., & Gerli, A. G. (2023). Can artificial intelligence help for scientific writing?. *Critical care*. [springer.com](https://doi.org/10.1186/s13054-023-04000-0)

Vilone, G. & Longo, L. (2021). Classification of explainable artificial intelligence methods through their output formats. *Machine Learning and Knowledge Extraction*. [mdpi.com](https://doi.org/10.3390/mlke1010001)

van Leeuwen, K. G., Schalekamp, S., Rutten, M. J., van Ginneken, B., & de Rooij, M. (2021). Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. *European radiology*, 31, 3797-3804. [springer.com](https://doi.org/10.1007/s00330-021-02400-0)

Korteling, J. H., van de Boer-Visschedijk, G. C., Blankendaal, R. A., Boonekamp, R. C., & Eikelboom, A. R. (2021). Human-versus artificial intelligence. *Frontiers in artificial intelligence*, 4, 622364. [frontiersin.org](https://doi.org/10.3389/fnint.2021.622364)

Fitria, T. N. (2021). Artificial intelligence (AI) in education: Using AI tools for teaching and learning process. *Prosiding Seminar Nasional & Call for Paper STIE AAS*. [researchgate.net](https://doi.org/10.21831/prosiding.v4i1.12345)

Jarrahi, M. H., Askay, D., Eshraghi, A., & Smith, P. (2023). Artificial intelligence and knowledge management: A partnership between human and AI. *Business Horizons*. [sciencedirect.com](https://doi.org/10.1016/j.bushor.2023.01.001)

Gezici, B. & Tarhan, A. K. (2022). Systematic literature review on software quality for AI-based software. *Empirical Software Engineering*. [\[HTML\]](https://doi.org/10.1007/s12618-022-01000-0)

Ghioni, R., Taddeo, M., & Floridi, L. (2023). Open source intelligence and AI: a systematic review of the GELSI literature. *AI & society*. [springer.com](https://doi.org/10.1007/s00146-023-00000-0)

Baviskar, D., Ahirrao, S., Potdar, V., & Kotecha, K. (2021). Efficient automated processing of the unstructured documents using artificial intelligence: A systematic literature review and future directions. *IEEE Access*. [ieee.org](https://doi.org/10.1109/ACCESS.2021.3050000)

Krivec, T., Košak Babuder, M., Godec, P., Weingerl, P., & Stankovič Elesini, U. (2020). Impact of digital text variables on legibility for persons with dyslexia. *Dyslexia*, 26(1), 87-103. [\[HTML\]](https://doi.org/10.1111/dys.12345)

Wallace, S., Treitman, R., Kumawat, N., Arpin, K., Huang, J., Sawyer, B., & Bylinskii, Z. (2020). Towards readability individuation: the right changes to text format make large impacts on reading speed. *Journal of Vision*, 20(10), 17-17. [academia.edu](https://doi.org/10.3379/jov.2020.2010.17)

Richardson, J. T. E. (2022). The legibility of serif and sans serif typefaces: Reading from paper and reading from screens. [oapen.org](https://oapen.org)

Longoni, C. & Cian, L. (2022). Artificial intelligence in utilitarian vs. hedonic contexts: The “word-of-machine” effect. *Journal of Marketing*. [sagepub.com](https://sagepub.com)

Abdollahi, A. & Pradhan, B. (2021). Urban vegetation mapping from aerial imagery using explainable AI (XAI). *Sensors*. [mdpi.com](https://mdpi.com)

Uzir, M. U. H., Al Halbusi, H., Lim, R., Jerin, I., Hamid, A. B. A., Ramayah, T., & Haque, A. (2021). Applied Artificial Intelligence and user satisfaction: Smartwatch usage for healthcare in Bangladesh during COVID-19. *Technology in society*, 67, 101780. [nih.gov](https://nih.gov)

MacLellan, C. J., & Koedinger, K. R. (2022). Domain-general tutor authoring with apprentice learner models. *International Journal of Artificial Intelligence in Education*, 32(1), 76-117. [springer.com](https://springer.com)

Hegghammer, T. (2022). OCR with Tesseract, Amazon Textract, and Google Document AI: a benchmarking experiment. *Journal of Computational Social Science*. [springer.com](https://springer.com)

Khan, B., Shah, Z. A., Usman, M., Khan, I., & Niazi, B. (2023). Exploring the landscape of automatic text summarization: a comprehensive survey. *IEEE Access*. [iee.org](https://iee.org)

Grabar, N. & Saggion, H. (2022). Evaluation of automatic text simplification: Where are we now, where should we go from here. *Traitement Automatique des Langues Naturelles*. [hal.science](https://hal.science)

Mahmud, H., Islam, A. N., Ahmed, S. I., & Smolander, K. (2022). What influences algorithmic decision-making? A systematic literature review on algorithm aversion. *Technological Forecasting and Social Change*, 175, 121390. [sciencedirect.com](https://sciencedirect.com)

Rodgers, W., Murray, J. M., Stefanidis, A., Degbey, W. Y., & Tarba, S. Y. (2023). An artificial intelligence algorithmic approach to ethical decision-making in human resource management processes. *Human Resource Management Review*, 33(1), 100925. [sciencedirect.com](https://sciencedirect.com)

Dalalah, D., & Dalalah, O. M. (2023). The false positives and false negatives of generative AI detection tools in education and academic research: The case of ChatGPT. *The International Journal of Management Education*, 21(2), 100822. [\[HTML\]](#)

Mahlow, C., & Piotrowski, M. (2022, September). Academic writing and publishing beyond documents. In *Proceedings of the 22nd ACM Symposium on Document Engineering* (pp. 1-4). [acm.org](https://acm.org)

Eden, A. (2024). Writing like a scientist: a research assignment for undergraduates inspired by the publishing process. *Journal of Microbiology and Biology Education*. [asm.org](https://asm.org)

Yeo, M. A., Renandya, W. A., & Tangkiengsirisin, S. (2022). Re-envisioning academic publication: From “Publish or Perish” to “Publish and Flourish”. *RELC Journal*. [sagepub.com](http://sagepub.com)

Kousha, K. & Thelwall, M. (2024). Artificial intelligence to support publishing and peer review: A summary and review. *Learned Publishing*. [wiley.com](http://wiley.com)

Filetti, S., Fenza, G., & Gallo, A. (2024). Research design and writing of scholarly articles: new artificial intelligence tools available for researchers. *Endocrine*. [\[HTML\]](#)

Khabib, S. (2022). Introducing artificial intelligence (AI)-based digital writing assistants for teachers in writing scientific articles. *Teaching English as a Foreign Language Journal*. [uad.ac.id](http://uad.ac.id)

Hand, B., Chen, Y. C., & Suh, J. K. (2021). Does a knowledge generation approach to learning benefit students? A systematic review of research on the science writing heuristic approach. *Educational Psychology Review*. [researchgate.net](http://researchgate.net)

Cotos, E., Huffman, S., & Link, S. (2020). Understanding graduate writers' interaction with and impact of the Research Writing Tutor during revision. *Journal of Writing Research*. [jowr.org](http://jowr.org)

Kim, N. J. & Kim, M. K. (2022). Teacher's perceptions of using an artificial intelligence-based educational tool for scientific writing. *Frontiers in Education*. [frontiersin.org](http://frontiersin.org)

Cloutier, C. & Ravasi, D. (2021). Using tables to enhance trustworthiness in qualitative research. *Strategic Organization*. [sagepub.com](http://sagepub.com)

Birkle, C., Pendlebury, D. A., Schnell, J., & Adams, J. (2020). Web of Science as a data source for research on scientific and scholarly activity. *Quantitative Science Studies*, 1(1), 363-376. [mit.edu](http://mit.edu)

Hiebl, M. R. W. (2023). Sample selection in systematic literature reviews of management research. *Organizational research methods*. [sagepub.com](http://sagepub.com)

Gruzauskas, V. & Ragavan, D. (2020). Robotic process automation for document processing: A case study of a logistics service provider. *Journal of Management*. [researchgate.net](http://researchgate.net)

Okunlaya, R. O., Syed Abdullah, N., & Alias, R. A. (2022). Artificial intelligence (AI) library services innovative conceptual framework for the digital transformation of university education. *Library Hi Tech*. [\[HTML\]](#)

Horbach, S. P. J. M. (2020). Pandemic publishing: Medical journals strongly speed up their publication process for COVID-19. *Quantitative science studies*. [mit.edu](http://mit.edu)

---

Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., ... & Dirnagl, U. (2020). The Hong Kong Principles for assessing researchers: Fostering research integrity. *PLoS biology*, 18(7), e3000737. [plos.org](https://doi.org/10.1371/journal.pbio.100737)

Purwanto, A., Ardiyanto, J., & Sudargini, Y. (2021). Inhibiting factors for publishing in reputable international journals among doctoral students: An exploratory mixed method study. *Journal of Industrial Engineering & Management Research*, 2(2), 1-20. [jiemar.org](https://doi.org/10.1108/jiemar-01-2021-0001)

Lê, J. K. & Schmid, T. (2022). The practice of innovating research methods. *Organizational Research Methods*. [sagepub.com](https://doi.org/10.1177/15265682221100000)

Gastel, B. & Day, R. A. (2022). How to write and publish a scientific paper. [leroysteinbacher.com](https://doi.org/10.1002/9781119800000.ch1)

Cargill, M. & O'Connor, P. (2021). Writing scientific research articles: Strategy and steps. [tue.nl](https://doi.org/10.1002/9781119800000.ch1)

Tenopir, C., Rice, N. M., Allard, S., Baird, L., Borycz, J., Christian, L., ... & Sandusky, R. J. (2020). Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. *PloS one*, 15(3), e0229003. [plos.org](https://doi.org/10.1371/journal.pone.0229003)

da Silva, J. A. T., Moradzadeh, M., Adjei, K. O. K., Owusu-Ansah, C. M., Balehegn, M., Faúndez, E. I., ... & Al-Khatib, A. (2022). An integrated paradigm shift to deal with ‘predatory publishing’. *The Journal of Academic Librarianship*, 48(1), 102481. [\[HTML\]](https://doi.org/10.1002/jal.1481)

Ramírez-Castañeda, V. (2020). ... in preparing and publishing scientific papers caused by the dominance of the English language in science: The case of Colombian researchers in biological .... *PloS one*. [plos.org](https://doi.org/10.1371/journal.pone.0229003)

Morton, B., Vercueil, A., Masekela, R., Heinz, E., Reimer, L., Saleh, S., ... & Oriyo, N. (2022). Consensus statement on measures to promote equitable authorship in the publication of research from international partnerships. *Anaesthesia*, 77(3), 264-276. [wiley.com](https://doi.org/10.1111/anae.15555)

Homolak, J., Kodvanj, I., & Virag, D. (2020). Preliminary analysis of COVID-19 academic information patterns: a call for open science in the times of closed borders. *Scientometrics*. [springer.com](https://doi.org/10.1007/s11192-020-03555-5)

Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative science studies*, 1(1), 377-386. [mit.edu](https://doi.org/10.1007/s43989-020-00000-0)