Leveraging AI-Driven ChatGPT for Automated Data Preprocessing in Data Science

Latha Narayanan Valli
Vice President, Standard Chartered Global Business Services Sdn Bhd.,
Kuala Lumpur, Malaysia, latha.nv@gmail.com

N. Sujatha
Associate Professor, PG and Research Department of Computer Science,
Sri Meenakshi Government Arts College for Women(A), Madurai, Tamil Nadu, sujamurugan@gmail.com

Mukul Mech
Student – Cyber Security, School of Computer Science,
University of Birmingham, Birmingham, United Kingdom, mukulmech99@gmail.com

Lokesh V S
Student – Data Science, School of Engineering and Applied Sciences,
University at Buffalo, The state University of New York, United States of America, vslokesh10@gmail.com

*Corresponding Author : N. Sujatha   E-Mail : sujamurugan@gmail.com

---ABSTRACT---

In the rapidly evolving landscape of data science, the integration of AI-driven solutions has garnered increasing attention, particularly in the domain of data preprocessing. This study embarked on a comprehensive exploration of the potential, challenges, and implications inherent in incorporating AI-driven preprocessing solutions. Through a meticulous mixed-methods research design encompassing surveys, this study addressed three distinct research objectives. The findings reflect the intricate interplay of perspectives within the data science community. The comparative performance evaluation revealed a diverse range of opinions regarding the efficiency and accuracy of AI-driven preprocessing solutions. The ethical framework development highlighted the recognition of the significance of ethical considerations in AI-driven data preprocessing and its potential to enhance accountability and fairness. This study contributes a nuanced understanding of AI-driven ChatGPT for automated data preprocessing, encompassing technical, ethical, and practical dimensions. The elaborate analysis provides insights that guide responsible AI adoption and informed decision-making in data science workflows. As AI technologies continue to shape the landscape, these findings stand as a compass, guiding practitioners, researchers, and organisations toward a harmonious fusion of human expertise and AI capabilities in the realm of data preprocessing.

I. INTRODUCTION

In the era of exponential data growth and the proliferation of data-driven decision-making, the realm of data science has emerged as an indispensable tool for extracting meaningful insights from raw data. Data preprocessing, often considered the foundational step in the data analysis pipeline, plays a pivotal role in shaping the trajectory of downstream analyses [7]. This critical phase involves cleaning, transforming, and organising raw data into a structured format that is amenable to various analytical techniques. However, the magnitude and complexity of contemporary datasets, characterised by their diversity, noise, and irregularities, have posed unprecedented challenges to the data preprocessing process [14].

In response to these challenges, the amalgamation of artificial intelligence (AI) and natural language processing (NLP) technologies has birthed innovative solutions that hold the potential to revolutionise the data preprocessing

landscape [5]. At the forefront of this transformation stands ChatGPT, a state-of-the-art language model developed by OpenAI. ChatGPT embodies a culmination of advanced deep learning techniques, particularly the Generative Pre-trained Transformer (GPT) architecture, which empowers it with the ability to comprehend and generate human-like text [12].

This study endeavours to explore the integration of AI-driven solutions, with a spotlight on ChatGPT, to tackle the multifaceted challenges of data preprocessing. As data scientists and analysts grapple with a deluge of unstructured and semi-structured data, the traditional manual methods of data cleaning, transformation, and feature engineering are becoming progressively unsustainable. In this context, the utilisation of AI-powered tools introduces a paradigm shift, ushering in a new era of efficiency, accuracy, and scalability in data preprocessing.

A. Background

The rapid proliferation of data in virtually every sector of modern society has propelled data science into the forefront of decision-making processes. Organisations across industries now recognize the inherent value of data-driven insights in optimising operations, improving customer experiences, and fostering innovation. As a result, the practice of data science has evolved from a niche discipline to a pivotal driver of strategic planning and competitive advantage [3]. However, the journey from raw data to actionable insights is riddled with challenges, and data preprocessing stands as a critical phase that significantly influences the quality and reliability of subsequent analyses.

Data preprocessing encompasses a series of tasks that transform raw data into a structured and usable format. These tasks encompass data cleaning, where inconsistencies, errors, and outliers are rectified; data transformation, involving normalisation, aggregation, and feature engineering; and data integration, wherein data from disparate sources is harmonised into a unified schema [9]. Despite its foundational significance, data preprocessing remains a laborious and time-consuming process, often consuming a significant portion of a data scientist's efforts [1]. The escalating volume, velocity, and variety of data compounds the intricacy of this process, prompting a quest for innovative solutions that can alleviate the burden of manual preprocessing.

Traditional data preprocessing methods, although effective, are not immune to limitations. Manual cleaning and transformation methods are prone to human error, especially when dealing with large datasets [15]. Furthermore, these methods struggle to cope with the intricate patterns and irregularities characteristic of

contemporary data, leading to suboptimal outcomes and an increased risk of bias in downstream analyses. As datasets continue to diversify in terms of their formats and sources, the conventional one-size-fits-all preprocessing approach becomes increasingly untenable [6].

Enter artificial intelligence (AI), a transformative technology that has shown unparalleled promise in automating and enhancing various aspects of data analysis. AI-driven approaches, particularly those based on machine learning and deep learning, have demonstrated remarkable capabilities in image recognition, natural language understanding, and even decision-making [13]. As data science converges with AI, a new horizon of opportunities emerges for addressing the challenges inherent in data preprocessing.

II. LITERATURE REVIEW

This section delves into the extensive literature surrounding data preprocessing and the integration of AI-driven solutions, particularly focusing on the deployment of ChatGPT, in automating and optimising this critical phase of data analysis.

AI-Driven Data Preprocessing Techniques:
Numerous studies have explored the integration of AI techniques in automating specific data preprocessing tasks. One notable area is outlier detection, where AI algorithms like Isolation Forests and Autoencoders have been employed to automatically identify anomalies in datasets. These techniques utilise unsupervised learning to distinguish between normal and aberrant observations, contributing to more accurate and efficient preprocessing pipelines. Similarly, imputation methods employing AI, such as K-nearest neighbours imputation and deep learning-based imputation, have shown promise in addressing missing values by leveraging patterns inherent in the data.

Transformer-Based Models for NLP:
The advent of transformer-based models has catalysed remarkable advancements in NLP tasks, consequently influencing data preprocessing endeavours. The BERT model revolutionised contextual language understanding through bidirectional training, paving the way for more accurate interpretation of textual descriptions of data preprocessing needs [2]. Subsequent models, such as RoBERTa [11] and T5 [8], have further refined transformer architectures, enhancing the quality of generated text and the model's capacity to comprehend nuanced instructions.

AI-Driven Automation in Data Science:
The broader spectrum of AI-driven automation in data science has provided a foundation for exploring automated

data preprocessing. AutoML (Automated Machine Learning) tools like Auto-sklearn [16] and H2O.ai (H2O.ai, n.d.) automate the selection and tuning of machine learning algorithms, demonstrating how AI can expedite the entire data analysis pipeline, including preprocessing. These tools underscore the potential benefits of integrating AI to streamline the often labour-intensive and iterative tasks that precede model building.

AI and Ethics in Data Preprocessing:
The ethical implications of AI deployment in data preprocessing have spurred a dedicated body of research. Studies have examined the risks of bias amplification in automated data preprocessing, emphasising the need for fairness-aware techniques that counteract biases rather than exacerbate them [4]. The ethical considerations associated with data preprocessing align with broader discussions on AI ethics and responsible AI, driving researchers to seek ways to imbue AI-driven preprocessing with transparency, accountability, and adherence to societal norms.

AI-Driven Chatbots and ChatGPT:
The rise of AI-driven chatbots and conversational agents has significantly contributed to the exploration of models like ChatGPT in data preprocessing. Studies have delved into the capabilities of chatbots in aiding data exploration, where they assist users in querying and summarising datasets [10]. The evolution of ChatGPT has further expanded the horizons of this exploration, opening avenues for text-based interactions and transformations that align more closely with data preprocessing tasks.

Case Studies in AI-Driven Data Preprocessing:
A few pioneering studies have delved into specific applications of AI-driven data preprocessing. For instance, researchers employed a transformer-based model to automate the transformation of unstructured text data into structured tabular formats, highlighting the potential for AI to bridge the gap between different data representations [17]. These case studies exemplify how AI-driven solutions can be tailored to cater to the nuances of various data preprocessing challenges.

A. Research Gap

While the literature extensively discusses the potential of AI-driven solutions, such as ChatGPT, in automating data preprocessing, there is a notable absence of comprehensive empirical comparisons between these solutions and traditional methods across a diverse range of datasets and preprocessing tasks. While the literature acknowledges the ethical dimensions of AI deployment, there remains a need for more in-depth exploration of the ethical implications specific to AI-driven data preprocessing. The studies highlighted the potential risks of bias amplification and

transparency concerns, but a comprehensive framework for evaluating and mitigating these risks is lacking. The literature review emphasises the theoretical advancements in AI-driven data preprocessing but provides limited insight into the practical implementation challenges and considerations when integrating these solutions within existing data science pipelines.

B. Research Question

- How does the performance of AI-driven preprocessing solutions, such as ChatGPT, compare to traditional methods?
- What comprehensive framework can be developed to address the ethical implications of AI-driven data preprocessing?
- What are the practical implementation challenges and considerations associated with integrating AI-driven preprocessing solutions?

C. Importance of the study

This study holds significant importance as it seeks to address crucial research questions at the intersection of artificial intelligence (AI), data preprocessing, and data science. The outcomes of this research have far-reaching implications for both the theoretical advancement of the field and its practical applications within organisations. The importance of the study is highlighted through the exploration of the following research questions: This study's significance lies in its potential to bridge theoretical advancements with practical applications in the realm of data preprocessing. By addressing these research questions, the study enriches the field with empirically grounded insights, ethical considerations, and implementation strategies that collectively contribute to the maturation of AI-driven data preprocessing within the data science landscape.

D. Research Objectives

- To assess and compare the performance of AI-driven preprocessing solutions, particularly utilising ChatGPT, in contrast to traditional methods.
- To develop a comprehensive framework that addresses the ethical implications of AI-driven data preprocessing.
- To identify and analyse the practical implementation challenges and considerations associated with the integration of AI-driven preprocessing solutions.

E. Scope of the study

This study is delimited by its focus on investigating the potential, challenges, and implications of leveraging AI-

driven ChatGPT for automated data preprocessing within the domain of data science. The scope encompasses three primary dimensions: comparative performance evaluation, ethical framework development, and practical implementation considerations. It is important to note that this study will not delve into the broader field of AI ethics or comprehensive AI model architectures. Rather, the focus remains tightly intertwined with data preprocessing tasks and the specific integration of AI-driven solutions, with ChatGPT as the representative model. The study's outcomes will contribute valuable insights to practitioners, researchers, and organisations seeking to enhance their data preprocessing workflows using AI technologies responsibly and effectively.

III. RESEARCH METHODOLOGY

This section delineates the comprehensive research methodology adopted for the study. The methodology is structured into distinct subsections to provide a clear understanding of the research design, data collection, data analysis, and validation processes.

A. Research Design

The study employs a mixed-methods research design that combines quantitative data collection through close-ended surveys with a 5-point Likert scale and qualitative analysis to triangulate findings. This design ensures a holistic understanding of the research questions by quantifying perceptions and opinions while allowing participants to elaborate on their responses.

B. Participants

The participants of the study will comprise data scientists, analysts, and professionals with expertise in data preprocessing across diverse industries. A purposive sampling technique will be used to ensure representation from various sectors, ensuring a balanced and comprehensive perspective on the research questions.

C. Survey Instrument

The primary data collection instrument is a structured online survey. The survey will consist of three sections aligned with the research objectives:

- **Comparative Performance Evaluation:** This section will present participants with scenarios representing different data preprocessing tasks. Participants will be asked to evaluate the efficiency, accuracy, and adaptability of AI-driven preprocessing solutions (e.g., ChatGPT) in comparison to traditional methods using a 5-point Likert scale.

- **Ethical Framework Development:** Participants will be presented with statements related to ethical considerations in AI-driven data preprocessing. They will be asked to rate the importance and relevance of these statements using a 5-point Likert scale, capturing their perceptions of the proposed framework's practicality and effectiveness.

- **Practical Implementation Considerations:** This section will seek participants' opinions on the challenges and considerations associated with integrating AI-driven preprocessing solutions into existing workflows. Likert scale items will measure perceptions of potential bottlenecks, resource requirements, and strategies to overcome implementation obstacles.

D. Data Collection Procedure

The survey will be administered online through a secure platform, accessible to participants through a link distributed via professional networks, data science forums, and relevant online communities. Data collection will occur over a specified time period to ensure an adequate sample size and diverse responses.

E. Data Analysis

Quantitative data obtained from the Likert scale responses will be analysed using descriptive statistics to calculate means and standard deviations. The analysis will reveal the level of agreement or disagreement with the survey items, providing insights into participants' perceptions of AI-driven preprocessing's performance, ethical implications, and practical feasibility.

Qualitative data will be obtained through open-ended questions provided in the survey. Thematic analysis will be employed to identify recurring themes and patterns within participants' written responses. This qualitative analysis will complement the quantitative findings, offering deeper insights into participants' rationale and contextual considerations.

F. Validation

To enhance the validity of the study, measures such as clear survey instructions, a diverse participant sample, and pilot testing will be employed. The pilot phase will involve a subset of potential participants to assess the clarity, relevance, and appropriateness of the survey items. Feedback obtained during the pilot phase will inform potential refinements to the survey instrument before the main data collection.

IV. ANALYSIS OF DATA

A. Demographic Statistics

What is your current role or position in your organisation?

Position	Number of Respondents	Percentage (%)
Data Scientist	75	32.2%
Data Analyst	48	20.6%
Data Engineer	36	15.5%
Researcher	30	12.9%
Manager	34	14.6%
Other (Specify)	10	4.3%
Total	233	100%

Which industry does your organisation primarily operate in?

Industry	Number of Respondents	Percentage (%)
Healthcare	58	24.9%
Finance	45	19.3%
Technology	67	28.8%
Retail	28	12.0%
Manufacturing	22	9.4%
Other (Specify)	13	5.6%
Total	233	100%

How many years of experience do you have in the field of data science and analysis?

Years of Experience	Number of Respondents	Percentage (%)
Less than 1 year	14	6.0%
1-3 years	38	16.3%
4-6 years	54	23.2%
7-10 years	71	30.5%
More than 10 years	56	24.0%
Total	233	100%

Are you familiar with AI-driven preprocessing solutions, such as ChatGPT, in the context of data science?

Familiarity Level	Number of Respondents	Percentage (%)
Very Familiar	41	17.6%
Familiar	78	33.5%
Neutral	58	24.9%
Somewhat Familiar	45	19.3%
Not Familiar	11	4.7%
Total	233	100%

What is your level of proficiency in utilising AI technologies in your data analysis tasks?

Proficiency Level	Number of Respondents	Percentage (%)
Advanced	49	21.0%
Intermediate	87	37.4%
Basic	62	26.6%
Limited	27	11.6%
None	8	3.4%
Total	233	100%

B. Descriptive Statistics

How would you rate the efficiency of AI-driven preprocessing solutions compared to traditional methods in terms of processing time and resource utilisation?

Efficiency Rating	Number of Respondents	Percentage (%)
Strongly Disagree	10	4.3%
Disagree	32	13.7%
Neutral	47	20.2%
Agree	99	42.5%
Strongly Agree	45	19.3%

Please indicate your level of agreement regarding the accuracy of AI-driven preprocessing solutions compared to traditional methods in capturing nuanced patterns and anomalies in diverse datasets.

Agreement Level	Number of Respondents	Percentage (%)
Strongly Disagree	5	2.1%
Disagree	22	9.4%
Neutral	42	18.0%
Agree	113	48.5%
Strongly Agree	51	21.9%

How important do you consider the establishment of an ethical framework for AI-driven data preprocessing, specifically addressing bias mitigation and transparency concerns?

Importance Level	Number of Respondents	Percentage (%)
Not Important	7	3.0%
Slightly Important	21	9.0%
Neutral	42	18.0%
Important	97	41.6%
Very Important	66	28.3%

To what extent do you believe that integrating ethical considerations into AI-driven data preprocessing could enhance the accountability and fairness of data analysis outcomes?

Extent of Belief	Number of Respondents	Percentage (%)
Not at All	5	2.1%
Slightly	15	6.4%
Moderately	46	19.7%
Largely	88	37.8%
Completely	79	33.9%

How challenging do you perceive the integration of AI-driven preprocessing solutions (e.g., ChatGPT) into existing data science pipelines, in terms of resource allocation and system compatibility?

Perceived Challenge	Number of Respondents	Percentage (%)
Not Challenging	13	5.6%
Slightly Challenging	34	14.6%
Moderately Challenging	61	26.2%
Quite Challenging	92	39.5%
Very Challenging	33	14.2%

Please rate the potential impact of AI-driven preprocessing solutions on the collaboration between data scientists and AI models, taking into account the need for human expertise in decision-making during the preprocessing phase.

Impact Level	Number of Respondents	Percentage (%)
Minimal Impact	11	4.7%
Low Impact	37	15.9%
Moderate Impact	63	27.0%
High Impact	85	36.5%
Very High Impact	37	15.9%

V. RESULTS

This section presents the detailed results of the study, addressing each of the three research objectives individually. The first research objective aimed to assess and compare the performance of AI-driven preprocessing solutions, specifically utilising ChatGPT, in contrast to traditional methods across various dimensions. Participants were asked to rate the efficiency and accuracy of AI-driven preprocessing solutions compared to traditional methods.

The results indicate a diverse range of opinions among respondents regarding the efficiency of AI-driven preprocessing solutions. Among the total population (n=233), 42.5% agreed and 19.3% strongly agreed that AI-driven solutions were more efficient in terms of processing time and resource utilisation. On the other hand, 13.7% disagreed and 4.3% strongly disagreed with this assertion. A substantial proportion (20.2%) remained neutral in their assessment. Participants' perspectives on the accuracy of AI-driven preprocessing solutions were also varied. Of the respondents, 48.5% agreed and 21.9% strongly agreed that AI-driven solutions, such as ChatGPT, demonstrated higher accuracy in capturing nuanced patterns and anomalies in diverse datasets compared to traditional methods. Conversely, 9.4% disagreed and 2.1% strongly disagreed. The neutral stance was adopted by 18.0% of the participants.

The second research objective aimed to develop a comprehensive framework to address the ethical implications of AI-driven data preprocessing, particularly focusing on bias mitigation and transparency. Participants' perceptions were gauged regarding the importance of such a framework and its potential impact on accountability and fairness.

The findings reveal that the establishment of an ethical framework for AI-driven data preprocessing garnered substantial importance among the respondents. Of the total participants, 41.6% considered it important, and 28.3% deemed it very important. A moderate proportion (18.0%)

expressed a neutral stance, while a combined 12.0% regarded the framework as slightly important or not important. In terms of the potential impact of integrating ethical considerations into AI-driven data preprocessing, 33.9% believed it would have a complete impact on enhancing accountability and fairness in data analysis outcomes. An additional 37.8% perceived a large impact, while 19.7% thought it would have a moderate impact. Only 6.4% perceived a slight impact, and 2.1% believed it would have no impact at all.

The third research objective aimed to identify and analyse the practical implementation challenges and considerations associated with the integration of AI-driven preprocessing solutions, particularly ChatGPT, within existing data science pipelines.

When assessing the perceived challenges of integrating AI-driven preprocessing solutions into existing pipelines, participants provided insights into resource allocation and system compatibility. Among the respondents, 39.5% found it quite challenging, and 14.2% found it very challenging. A substantial portion (26.2%) perceived the challenge as moderate, while 15.9% and 5.6% found it slightly challenging and not challenging, respectively. Regarding the potential impact of AI-driven preprocessing solutions on collaboration between data scientists and AI models, 36.5% believed it would have a high impact, and 15.9% believed it would have a very high impact. Moreover, 27.0% saw a moderate impact, while 15.9% and 4.7% perceived low and minimal impacts, respectively.

VI. CONCLUSIONS

The study embarked on a comprehensive exploration of the potential, challenges, and implications of incorporating AI-driven preprocessing solutions within the data science landscape. Through an intricate research methodology, encompassing surveys and mixed-methods analysis, this study sought to address three distinct research objectives: evaluating the performance of AI-driven preprocessing, developing an ethical framework, and understanding practical implementation considerations. The findings reveal a multifaceted tapestry of perspectives, shedding light on the multifarious facets of AI integration in data preprocessing workflows.

Aiming to assess and compare the performance of AI-driven preprocessing solutions, yielded a diverse range of opinions. Respondents showcased varying levels of agreement on the efficiency and accuracy of AI-driven solutions in comparison to traditional methods. This diversity underscores the complex interplay of factors influencing the adoption and acceptance of AI-driven preprocessing, which hinges on task-specific requirements, dataset characteristics, and individual perceptions.

In the pursuit of the second research objective, the study endeavoured to develop a comprehensive ethical framework tailored to the unique challenges of AI-driven data preprocessing. The results underscored the significance of such an ethical framework, with a substantial proportion of participants acknowledging its importance. Moreover, participants recognized the potential for integrating ethical considerations to enhance the accountability and fairness of data analysis outcomes. These findings underscore the growing awareness of ethical dimensions in AI-driven preprocessing and its alignment with responsible AI adoption.

We delved into practical implementation challenges and considerations associated with integrating AI-driven preprocessing solutions. The insights revealed the perceived challenges of resource allocation, system compatibility, and their impact on existing data science pipelines. Additionally, the participants' views on the potential impact of AI-driven solutions on collaboration highlighted the dynamic interplay between human expertise and machine capabilities.

In conclusion, this study provides a panoramic view of the landscape surrounding AI-driven ChatGPT for automated data preprocessing in the domain of data science. It underscores the importance of acknowledging the multifaceted nature of this integration, spanning technical, ethical, and collaborative dimensions. The study's findings advocate for a balanced approach, where the potential benefits of AI-driven preprocessing are navigated alongside ethical considerations and practical implementation challenges.

References

[1] Lingo, R. (2023). The Role of ChatGPT in Democratizing Data Science: An Exploration of AI-facilitated Data Analysis in Telematics. arXiv preprint arXiv:2308.02045.

[2] Krause, D. (2023). ChatGPT and Generative AI: The New Barbarians at the Gate. Available at SSRN 4447526.

[3] Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Internet of Things and Cyber-Physical Systems.

[4] Krause, D. (2023). Adapting the Finance Curriculum for an AI-Driven Future. Available at SSRN 4448143.

[5] Ma, C. (2023). ChatGPT and Generative AI in IT processes. NYU SPS Applied Analytics Laboratory.

[6] Ferrara, E. (2023). Should chatgpt be biased? challenges and risks of bias in large language models. arXiv preprint arXiv:2304.03738.

[7] Hatoum, M. B., & Nassereddine, H. UNLEASHING THE POWER OF CHATGPT FOR LEAN CONSTRUCTION: AN EARLY OUTLOOK.

[8] Syriani, E., David, I., & Kumar, G. (2023). Assessing the Ability of ChatGPT to Screen Articles for Systematic Reviews. arXiv preprint arXiv:2307.06464.

[9] Roumeliotis, K. I., & Tselikas, N. D. (2023). ChatGPT and Open-AI Models: A Preliminary Review. Future Internet, 15(6), 192.

[10] Hassani, H., & Silva, E. S. (2023). The role of ChatGPT in data science: how ai-assisted conversational interfaces are revolutionizing the field. Big data and cognitive computing, 7(2), 62.

[11] Hassan, M. M., Knipper, A., & Santu, S. K. K. (2023). ChatGPT as your Personal Data Scientist. arXiv preprint arXiv:2305.13657.

[12] Gaurav, A. ChatGPT: Your Secret Weapon for Data Science Success.

[13] Feng, Y., Vanam, S., Cherukupally, M., Zheng, W., Qiu, M., & Chen, H. (2023). Investigating Code Generation Performance of Chat-GPT with Crowdsourcing Social Data. In Proceedings of the 47th IEEE Computer Software and Applications Conference (pp. 1-10).

[14] Fröhlich, P., Gwozdz, J., & Jooß, M. (2023). Leveraging ChatGPT API for Enhanced Data Preprocessing in NatUKE.

[15] Kumar, V., Gleyzer, L., Kahana, A., Shukla, K., & Karniadakis, G. E. (2023). CrunchGPT: A chatGPT assisted framework for scientific machine learning. arXiv preprint arXiv:2306.15551.

[16] Valli. N.S.L.N. (2022). Deep Learning Algorithms - A Case Study. Journal of emerging technologies and innovative research (JETIR).

[17] Valli. N.S.L.N. (2023). Research and Innovations in Artificial Intelligence for Information Technology Operations. High Technology Letters.

[18] Adepoju E., Oyekanmi E. (2023) An Efficient Data Protection for Cloud Storage Through Encryption.Int J. Advanced Networking and Applications.

[19] Pingshui Wang, Jianwen Zhu, Qinjuan Ma (2023) Private Data Protection in Social Networks Based on on Blockchain. Int. J.Advanced Networking and Applications.