Scholastic Seed Inc.

www.gjeis.com

global Journal of ENTERPRISE INFORMATION SYSTEM

Data Cleaning: Paving a Way for Accurate and Clean Data

– Laxmi Ahuja*
Dy. Director, AIIT, Amity University, Noida
laxmiahuja@gmail.com *https://orcid.org/0000-0002-4486-3081*

– Bhoomika Singh
Student, AIIT, Amity University, Noida
bhoomikasingh.1611@gmail.com *https://orcid.org/0009-0004-3780-6865*

– Rajbala Simon
Addl. Superintendent examination, AIIT, Amity University, Noida
rajbalasimon@gmail.com *https://orcid.org/0000-0002-7204-3486*

ARTICLE HISTORY

Paper Nomenclature: Theme Based Paper (TBP)
Paper Code: GJEISV16I1JM2024TBP1
Submission at Portal (www.gjeis.com): 04-Jan-2024
Manuscript Acknowledged:16-Jan-2024
Originality Check: 25-Jan-2024
Originality Test (Plag) Ratio (Drill BIT): 07%
Author Revert with Rectified Copy: 29-Jan-2024
Peer Reviewers Comment (Open): 31-Jan-2024
Single Blind Reviewers Explanation: 14-Feb- 2024
Double Blind Reviewers Interpretation: 19-Feb-2024
Triple Blind Reviewers Annotations: 28-Feb-2024
Author Update (w.r.t. correction, suggestion & observation): 02- Mar-2024
Camera-Ready-Copy: 19-Mar-2024
Editorial Board Excerpt & Citation: 22-Mar-2024
Published Online First: 31-Mar-2024

ABSTRACT

Purpose: Data cleaning plays one of the most important roles to ensure the quality and reliability of data that has been used for various purposes such as data analysis, artificial intelligence, decision making, etc. With the ever-increasing amount of data in this digital age, it becomes very significant to address the problem of data inconsistency, duplication, incompleteness and inadequacy.

Design/Methodology/Approach: With the help of various other research papers available online, different point of views regarding the data cleaning and various datasets available as a result of data cleaning using various techniques.

Findings: The research paper first discusses data cleansing, its steps and the significance of data cleansing in various fields. It also specifies key dimensions of data quality such as completeness, correctness, consistency, accuracy and uniqueness. The paper also covers various data cleaning techniques including ETL and text mining techniques such as NTLK and NLP techniques. Additionally, this paper covers the various challenges associated with data cleansing in RDBMS. It explores emerging trends and various advances in data cleansing during OLTP. The conclusion of the study emphasizes the need for a systematic approach to data cleaning and the importance of evaluating and proposing data cleaning. And Major technological improvement in this area..

Originality/Value: This paper will help us understand the current technologies and further advancements that can be made in the data cleaning field.

Paper Type: Theme Based Paper.

KEYWORDS: ETL (Extract Transform Load) | NTLK (Natural Language Toolkit) | NLP (Natural Language Processing) | RDBMS (Relational Database Management System) | OLTP (Online Transaction Processing System) | ML (Machine Learning)

***Corresponding Author** (Laxmi Et. Al)*

- Present Volume & Issue (Cycle): Volume 16 | Issue-1 | Jan-Mar 2024
- International Standard Serial Number: Online ISSN: 0975-1432 | Print ISSN: 0975-153X
- DOI (Crossref, USA) https://doi.org/10.18311/gjeis/2024
- Bibliographic database: OCLC Number (WorldCat): 988732114
- Impact Factor: 3.57 (2019-2020) & 1.0 (2020-2021) [CiteFactor]
- Editor-in-Chief: Dr. Subodh Kesharwani
- Frequency: Quarterly
- Published Since: 2009
- Research database: EBSCO https://www.ebsco.com
- Review Pedagogy: Single Blind Review/ Double Blind Review/ Triple Blind Review/ Open Review

- Publishers: Scholastic Seed Inc. and KARAM Society
- Place: New Delhi, India.
- Repository (figshare): 704442/13

Introduction

In today's digital era where everything is dependent on the digital data; every organisation faces the challenge of integrating the data from diverse sources, in different formats, structures and quality without any inconsistency. Extract, Transform, Load i.e., (or ETL) process is used to clean and process the data. However, the success of these processes heavily depends upon the data quality and cleanliness of the integrated data. Therefore, Data cleansing has an important role in ensuring the consistency, reliability and accuracy of the integrated data (Ilyas, 2019). Data cleaning, also known as "data cleansing" or "cleaning", is one of the most important steps in data integration process such as ETL. This includes finding and correcting errors and inconsistencies in data to improve data quality. Data quality issues like typographical errors, grammatical errors, missing information, redundant data, as well as invalid data exist in any data collection such as files and databases, and when the entire data collection needs to be integrated into a data warehouse or anywhere then Data cleaning becomes more and more important for those databases and systems (Sreemathy, 2021).

Data warehouse needs data cleansing support. They load and data from various sources. In addition, data warehouses are used for decision-making, so data should be correct to avoid wrong conclusions and decisions. Data cleaning is a hectic and tedious process most of the time, but it's definitely worth it to get the best out of your data, with consistent and accurate results (Tang, 2014). Explanation on a scale of 1 to 100: It costs $1 to prevent bad data, $10 to fix bad data, and $100 to fix problems caused by bad data. Therefore, it is important to perform data cleaning for best results.

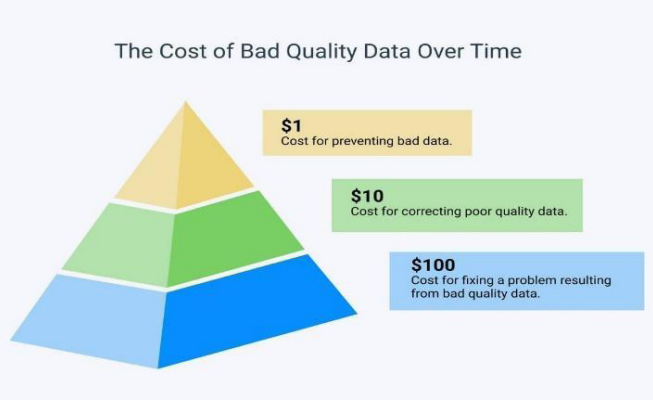

Figure1: Image of the cost of Bad Quality Data Over Time

In machine learning, accurate data is more crucial than the most powerful algorithm, because ML models only operate on the data they were trained on. If we train ML models with bad data, the final results will be detrimental to our organization. Therefore, data cleaning will save our time, money and make our organization more efficient.

Literature Review

Everyday increase in the amount of data and data dependency, the focus on data accuracy, consistency and reliability has increased the research in the field of Data Cleaning. This literature gives us the review of nuances of data cleaning and its problems, approaches and improvements that can address the challenges in future. This research has identified the process of ETL in streamlining the Data Cleaning workflows. Furthermore, the literature also covers the proactive ETL tools and strategies to prevent errors without compromising the quality of the data. Improved solutions leverage AI and NLP to detect and rectify anomalies. In this research paper, the existing knowledge of Data cleaning provides the insights into evolving methodologies of Data Cleaning.

Proposed Framework

This proposed framework consists of several key components discussing the current technologies used to clean the data and the future improvements that can be done in these techniques. Imputation methods are commonly used methods which basically predicts the missing values based on available information. We have several online tools like Oracle Data Integrator, Fivetran, Databricks, Apache Airflow etc. We use text mining such as SpaCy, NTLK nowadays. But all of these approaches do not rectify the data completely. Every approach has some loopholes associated. This research paper introduces us to some improvements by integrating AI to the Data Cleaning technologies, through Real-Time Data Processing, version-control systems and further improvements are discussed later.

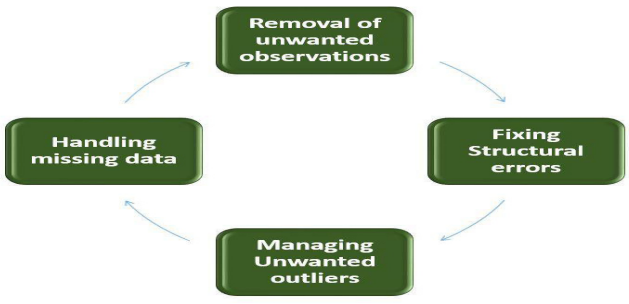

Figure 2: Steps Involved in Data Cleaning

Data Cleaning Methodologies

- **Delete Irrelevant Data:**

Example: If you are analyzing Volkswagen car owners, but your dataset contains data on Hyundai car owners..

We should also remove hashtags, URLs, emoticons, etc. as they are unnecessary parts of the analysis.

• Deduplication:

Example: If two datasets have the same information and results, only one dataset should be considered, otherwise our ML model will produce duplicate results every time.

• Correct Structural Errors Such as Spelling Mistakes, Wrong Words, etc.:

Example: If we are analyzing different data sets and one data set has the column "male" and the other has the column "men", then we need to normalize the headings.

• Handle Missing Data:

Scan your data for missing cells, text gaps, missed survey responses, and more.

• Filter Data Outliers:

Example: If you calculate the average score for the entire class and a student answered 0 questions, his score will have a large effect on the average score. Therefore, you should remove his data points so that the result is closer to the average (Ilyas, 2019).

• Check Your Data:

Check that your data structure is correct and that nothing is missing or inaccurate because wrong data will surely negatively affect organisation's result. So, Data Cleaning will save our time, and money and will make our organisation more efficient.

Outlier Detection:

Outliers are the outlying values which means the values which are the values that are either too large or too small as compared to other values. These values can disrupt the whole result. Therefore, it is primarily important to detect and remove the outliers. Various Outliers Detection methods like Z-Score detection, Linear Regression Models, High Dimensional Outlier Detection Methods are used to detect outliers. Outlier detection also faces some challenges: It is strenuous to know what normal data pattern is and to consider what is the outlier in the data as outliers may vary from data to data (Ilyas, 2019).

Flowchart for Data Cleaning Process:

From the below flowchart, we can infer that the raw data is being cleaned by the data cleaning tools in the transformation stage of ETL Process. In the first step, the tool accepts the raw data and checks its format and structure. Then it checks the constraints and various integrity checks used in the data. If any error is found, this cycle begins again from checking the constraints and removing the errors till the time comes when no error is left in the data. The algorithms of the tools check whether the data accuracy is above the satisfactory level or not; if the data accuracy is not above the satisfactory level, the data cleaning tool starts with changing the structure, format and constraints of the data to make them accurate and above the satisfactory level. When the data accuracy comes above the satisfactory level, the clean dataset is obtained by the organisation in data warehouse (Makarov, 2023).

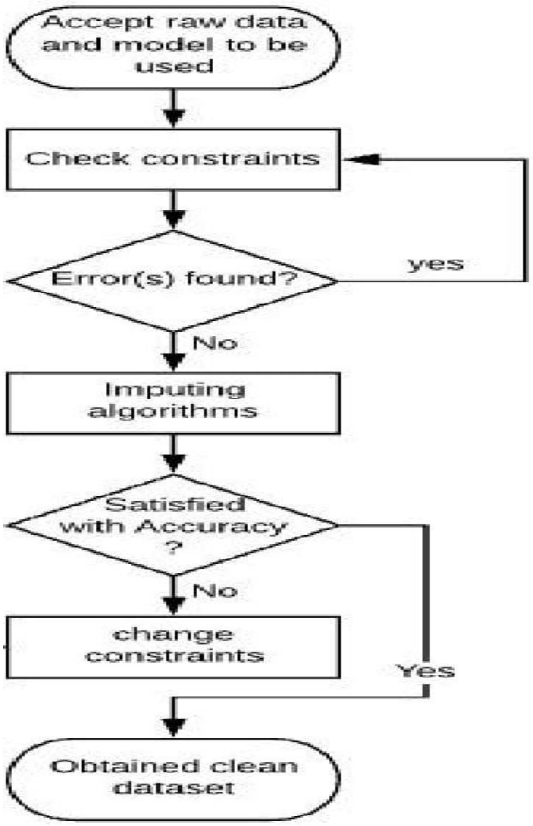

Figure 3: Flowchart of Data Cleaning Process

Data Duplication:

Data duplication is the most common error that occurs in data. These duplicate data need to be removed from the spreadsheets or any type of data you are working on because it can also generate wrong conclusions and hence lead to wrong decision making. Data Deduplication, also known as Duplicate Detection or Record Linkage refers to identifying the duplicates tuples and removing them. In Ms-Excel and Google Sheets, we can apply various methods like Sort & Filter and Conditional Formatting to detect the Duplicate values. Some more methods that we can use to detect duplicates includes hashing, matching, clustering or anomaly detection (Chu, 2016)

Data Warehouse:

A Data Warehouse is a computer system that stores, organize, transforms and analyze an organization's data in Online Transaction Processing System (OLTP). It is a collection of data designed to support decisions of any organisation. It consists of the data collected by the organization to draw various decisions and conclusions from it (Tang, 2014). To facilitate the extraction of useful information, the data needs to be pre-processed and then the processed data is stored in a data warehouse. These data warehouses are used for analytical data processing. So, for data processing we have several integrated parts of data warehouse like ETL process.

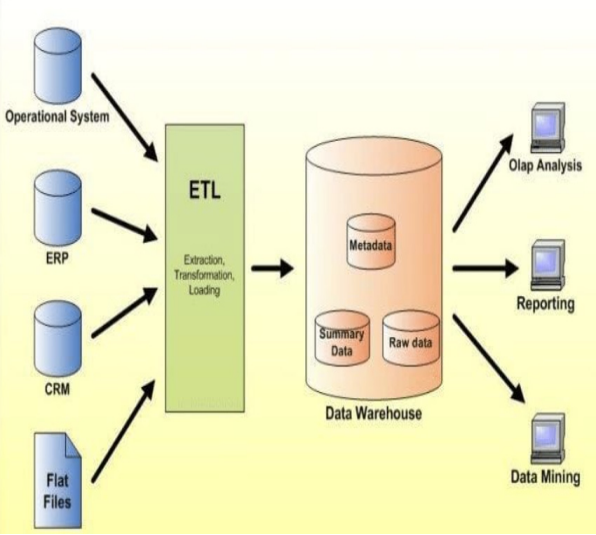

Figure 4: Data Warehouse Architectural diagram

Data Pre-Processing in Data Warehouse:

Irrelevant and inconsistent data always increases the cost of the data warehousing. So, Data Pre-processing is needed to improve the performance of system. There are different kinds of pre-processing required for different types of data.

- Classical Preprocessing: It includes Data Fusion Phase, Data Cleaning Phase, Data Structuration Phase
- Advanced Preprocessing: It includes Data Summarization Phase (Sreemathy, 2021).

ETL-Driven Cleaning:

ETL stands for Extract, Transform and Load. In this process, ETL tools (Oracle, MarkLogic, Segment, etc.) extracts data of different format from large number of RDBMS sources, transforms the data by applying computations, concatenations, etc., on the data and finally load the processed data into the data warehouse system.

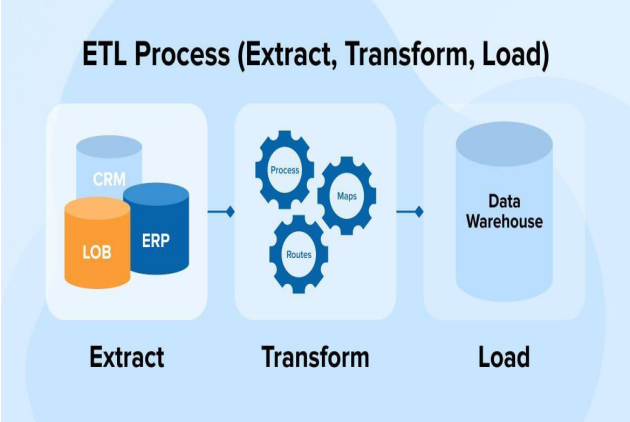

Figure 5: Extract, Transform, Load (ETL) Process

Extract: As in the image given above, the first step of ETL Process is to extract the important data from the source system and sending it to the transform stage where the possible transformations in data are done. The data can be extracted from various sources such as database, flat files and various other sources.

There are 3 ways to extract data

- Full extraction
- Partial Extraction - No update notification
- Partial extraction – with update notification (Sreemathy, 2021)

Transform: The Transform phase converts the data into a format that is compatible with the data warehouse. Data is transformed using various techniques such as data cleaning, data filtering, data transformation, data modelling. Data cleansing is part of the transformation phase of the ETL process and is performed before transforming the data into the required format using data cleansing tools (Sreemathy, 2021).

Load: In this phase, the data is loaded into the warehouse in a short period of time. Therefore, in the case of performance, the loading process must be optimized. In the case of a load failure, recovery must start from the point of failure without losing data integrity.

There are different types of loading:

- Initial loading: Populates all data warehouse tables.
- Incremental loading: apply ongoing changes as needed.
- Full refresh: deletes the contents of the (Sreemathy, 2021) tables and loads them with new contents

ETL Tools for Data Cleaning:

There are various ETL Tools available in the market. Some of them are:

- **MarkLogic:**

It is an ETL tool that uses a set of advanced features to make the data integration process faster and easier. It requests different types of data such as documents, metadata, big data, etc.8]]

Figure 6: MarkLogic logo

- **Oracle Data Integrator:**

ODI is an industry-leading ETL tool which offers us a large range of Data Warehouse solutions for both in the cloud and on-premises. It comes with connectors for many databases like Hadoop, ODBC, JDBC, JSON, XML etc

Figure 7: ODI logo

- **Amazon RedShift**

It is very simple, user-friendly and cost-effective tool for analyzing different types of data. It uses standard SQL to analyze data (Singh, 2020).

Figure 8: Amazon Redshift logo

- **Etleap:**

Etleap is the perfect ETL tool for enterprise teams that need to pipeline the data quickly. With Etleap, we can use SQL to phrase, structure and clean data. Pipelining data with Etleap is very easy, and as with all issues we encounter, the Etleap support staff quickly identifies and resolves the issues (Singh, 2020).

Figure 9: Etleap logo

- **Segment by Twilio:**

Acquired by Twilio in 2020, Segment is a platform for collecting events from mobile applications. Its powerful feature is data capture and delivery, which helps us collect data from customer touch points and load it into various warehouses. It has a free plan.

Figure 10: Segment by Twilio logo

- **Fivetran**

Fivetran is another ETL tool that is great for beginners and non-technical teams that need to connect together and clean data. However, the platform does not provide any data quality control features, which is its biggest drawback.

Figure 11: Fivetran logo

- **Stitch**

Stitch is a simple but very powerful ETL tool used by teams integrating over 130 data sources. Various enterprise teams are using Stitch for tighter security, management and analytics.

Figure 12: Stitch logo

- **Databricks SQL**

It is a serverless data warehouse running on the Databricks Lakes platform. We can run all SQL and BI applications. It works with SQL and other powerful tools like Fivetran, Tableau and Power BI. Since it is a serverless data warehouse, there is no need to maintain the infrastructure of cloud services, which is its biggest advantage.

Figure 13: Databricks logo

- **Apache Airflow**

It is an open source ETL tool for monitoring workflows and data. It consists of user and command interfaces that are characterized by processing data from various sources (Singh, 2020).

Figure 14: Apache Airflow logo

Problems Associated with Data Cleaning

Data cleaning is one of the essential processes in today's era which involves data identification, correction of errors and inconsistencies in the data. Due to increasing data day by day, several problems arises during the data cleaning process:

- **Missing Data:** Incomplete data and records, even a missing value or variable in the data that is to be used for the data analysis can be harmful for the analysis and can affects the quality of the data analysis and modelling
- **Redundancy:** Repeated and duplicate data is increasing day by day which also makes data analysis difficult and confusing. It also causes memory wastage.
- **Inconsistent Data:** Data coming from different data sources have different formats; these inconsistent representations of the data can become a major challenge to combine the data. It can arise due to the system errors during data extraction and collection and then merging of the data can absolutely introduce some errors.
- **Major Outliers:** Outliers are the values that are much smaller or much larger than most of the other values in the data set and it can distort the statistical analyses and machine learning models
- **Different Datatypes:** Since the data assembled in the warehouse comes from the various sources which uses different datatypes to create the data, this difference in the datatype can make the data inconsistent and difficult to understand and merge (Calabrese, 2018).

Approaches Associated with the Data Cleaning Problems

- Some techniques such as imputation methods such as mean/median imputation, regression imputations or multiple imputations can be used to fill the missing values.
- Another approach is to eliminate the records that contain the missing data, but it can introduce distortions in the results so, it's better to predict the missing values using above mentioned methods.
- To remove the inconsistencies present in the data, we can convert the whole dataset into a single data format and units. Also, we can apply predefined rules to correct the errors like spelling errors or value errors.
- Use of text mining such as NTLK, SpaCy, Apache OpenNLP etc and NLP techniques such as tokenization to identify and correct the inconsistencies hidden in the text of the data.
- We can even detect the outliers using statistical methods such as z-scire, quartiles or boxplots and after detecting we can either remove them or correct them. We can also apply the clustering algorithms to identify the outliers based on their deviations.
- Use of the Quality Detection tools such as OpenRefine, Trifacta Wrangler, DataCleaner etc can be used to detect the quality of the data based on the predefined criteria. Not even that we can also employ the experts for the manual inspection of the dataset for critical datasets (Wang, 2019).
- The number of data sources and their degree of dirtiness also determine data transformation and cleaning steps.

Improvements that Can be Done in Data Cleaning Tools

Various improvements can be done in data cleaning tools to ensure that the data is cleaned effectively and efficiently. Here are some few improvements suggested:

- Various algorithms of data cleaning can be integrated with the artificial intelligence as artificial intelligence can learn automatically from the patterns of data formats. This can automate the process of data cleaning and hence can make it easier (Makarov, 2023).
- Data Cleaning tools with the limited storage should be extended to unlimited storage so that large amounts of the data can be assembled and can be cleaned simultaneously; saving the time of data cleaning.
- Data cleaning tools can be improved and converted to real time data monitoring software which can detect and correct the errors in real time (Makarov, 2023).
- Data Cleaning tools can be modified to allow the multiple users to collaborate with each other simultaneously and also implement version control systems in the tool to track the changes made previously so that the user can easily go back to the changes made if they want to.
- Some powerful error handling mechanisms can be implemented to handle any type of error occurring and provide the error reports and real-time troubleshooting helpline too.
- Data Cleaning tools can be improved by improving the UI of the application. Making a more user-friendly interface will help the user to understand the features of data cleaning tools easily (Makarov, 2023).

Future Thoughts

Even though there are many advances going on in data cleaning, still there are loopholes in every condition. There are multiple future research directions available:

- We have discussed various methods to detect and remove errors but still many errors remain undetected; therefore, in future more expressive integrity constraint languages needs to be designed (Ilyas, 2019).
- Involving humans in data cleaning process intelligently needs to be explored.
- Next question that needs to be solved is that even after so many advances in data cleaning technology; Is there any way to preserve the privacy of the secretive data that needs to be combined with other data sources during transformation? Can we create such technology which does not compromise the privacy of the data (Ilyas, 2019)?

Conclusion

Data cleaning plays an important role in maintaining the quality, accuracy, reliability and consistency of the data used in various fields. It is one of the most important areas in data analytics. It improves the decision making and insights of the organization. This Research paper not only explored data cleaning but also explored various tools of data cleaning. It highlights the problems that an analyst can face during the data cleaning and the current approaches towards it and some improvements that can be done in data cleaning tools to address the problem of the data quality issues. Data Cleaning tools can provide the organization with the authority to handle large amounts of day-to-day data, improve its quality and improve the decision-making process. Overall, this Research paper highlights the significance of data cleaning. It is necessary for the explorers to come up with more advanced technologies for data cleaning and to overcome the emerging challenges of big data.

References

- Ilyas, I. F., & Chu, X. (2019). *Data cleaning*. Morgan & Claypool
- Chu, X., Ilyas, I. F., Krishnan, S., & Wang, J. (2016, June). Data cleaning: Overview and emerging challenges. In *Proceedings of the 2016 International Conference on management of data* (pp. 2201-2206).
- Tang, N. (2014). Big data cleaning. In *Asia-Pacific Web Conference* (pp. 13-24). Cham: Springer International Publishing.
- Wang, X., & Wang, C. (2019). Time series data cleaning: A survey. *Ieee Access*, *8*, 1866-1881.
- Makarov, A., & Namiot, D. (2023). Overview of data cleaning methods for machine learning. *International Journal of Open Information Technologies*, *11*(10), 70-78.
- Kiebler, L., Moroff, N. U., & Jacobsen, J. J. (2022). Preliminary analysis on data quality for ML applications. In *Changing Tides: The New Role of Resilience and Sustainability in Logistics and Supply Chain Management–Innovative Approaches for the Shift to a New Era. Proceedings of the Hamburg International Conference of Logistics (HICL), Vol. 33* (pp. 207-236). Berlin: epubli GmbH.
- Calabrese, B. (2018). Data cleaning. *Encyclopedia of bioinformatics and computational biology: ABC of bioinformatics*, *472*.
- Singh, S. K., & Dwivedi, D. R. K. (2020). Data mining: dirty data and data cleaning. *Available at SSRN 3610772*.
- Sreemathy, J., Brindha, R., Nagalakshmi, M. S., Suvekha, N., Ragul, N. K., & Praveennandha, M. (2021, March). Overview of etl tools and talend-data integration. In *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)* (Vol. 1, pp. 1650-1654). IEEE.
- Patel, M., & Patel, D. B. (2020). Progressive growth of ETL tools: A literature review of past to equip future. *Rising Threats in Expert Applications and Solutions: Proceedings of FICR-TEAS 2020*, 389-398.
- Biswas, N., Sarkar, A., & Mondal, K. C. (2019). Empirical analysis of programmable ETL tools. In *Computational Intelligence, Communications, and Business Analytics: Second International Conference, CICBA 2018, Kalyani, India, July 27–28, 2018, Revised Selected Papers, Part II 2* (pp. 267-277). Springer Singapore.

Annexure 16.1.2

Submission Date	Submission Id	Word Count	Character Count
25-Jan-2024	1575302 (DrillBit)	3099	19271

Analyzed Document	Submitter email	Submitted by	Similarity
2.1 TBP1_Laxmi_GJEIS Jan to Mar 2024.docx	laxmiahuja@gmail.com	Laxmi Ahuja	07%

DrillBit

7 SIMILARITY % | 6 MATCHED SOURCES | A GRADE

B-Upgrade (11-40%)
C-Poor (41-60%)
D-Unacceptable (61-100%)

LOCATION	MATCHED DOMAIN	%	SOURCE TYPE
1	docplayer.net	2	Internet Data
2	www.guru99.com	2	Internet Data
3	fdokumen.id	1	Internet Data
4	plosjournal.deepdyve.com	1	Internet Data
5	dokumen.pub	<1	Internet Data
6	nludelhi.ac.in	1	Publication

Reviewers Memorandum

Reviewer's Comment 1: The article effectively highlights the importance of data cleaning in today's digital era where data plays a crucial role in decision-making processes. The inclusion of real-world examples and cost analysis adds practical value to the discussion. The proposed framework and methodologies provide a comprehensive overview of the data cleaning process.

Reviewer's Comment 2: While the article covers various aspects of data cleaning, it could benefit from deeper insights into emerging trends and technologies in the field. Exploring recent advancements such as machine learning-based data cleaning algorithms or blockchain-based data validation methods could enhance the relevance and comprehensiveness of the paper.

Reviewer's Comment 3: The article is well-structured and easy to follow. However, some sections could be further refined for clarity. For instance, the discussion on ETL-driven cleaning in the Proposed Framework section could be elaborated with clearer explanations or diagrams to aid understanding. Overall article is well written and easy to read.

Citation

Laxmi Ahuja, Bhoomika Singh and Rajbala Simon "Data Cleaning: Paving a Way for Accurate and Clean Data" Volume-16, Issue-1, Jan-Mar 2024. (www.gjeis.com)

https://doi.org/10.18311/gjeis/2024
Volume-16, Issue-1, Jan-Mar 2024
Online iSSN : 0975-1432, **Print iSSN** : 0975-153X
Frequency : Quarterly, Published Since : 2009

Google Citations: Since 2009
H-Index = 96
i10-Index: 964

Source: https://scholar.google.co.in/citations?user=S47TtNkAAAAJ&hl=en

Conflict of Interest: Author of a Paper had no conflict neither financially nor academically.

Editorial Excerpt

The article has 07% of plagiarism which is the accepted percentage as per the norms and standards of the journal for publication. As per the editorial board's observations and blind reviewers' remarks the paper had some minor revisions which were communicated on a timely basis to the authors (Laxmi, Bhoomika & Rajbala), and accordingly, all the corrections had been incorporated as and when directed and required to do so. The comments related to this manuscript are noticeably related to the theme "**Data Cleaning: Paving a Way for Accurate and Clean Data**" both subject-wise and research-wise. The article presents a thorough investigation of the importance and methodologies of data cleaning in ensuring the accuracy and reliability of integrated data. The inclusion of relevant figures and flowcharts enhances the visual appeal and understanding of the concepts discussed. However, to further enrich the paper, it would be beneficial to include a brief discussion on the potential ethical implications of data cleaning, particularly regarding privacy concerns. After comprehensive reviews and the editorial board's remarks, the manuscript has been categorized and decided to publish under the "**Theme Based Paper**" category.

Acknowledgement

The acknowledgement section is an essential component of academic research papers, as it provides due recognition to all those who contributed their hard work and effort towards the writing of the paper. The author/s (Laxmi, Bhoomika & Rajbala) express their sincere gratitude to all those who assisted in the research process and made this paper a possibility. Lastly, the reviewers and editors of GJEIS deserve recognition for their pivotal role in publishing this issue, without whom the dissemination of this valuable research would not have been possible.

Disclaimer

All views expressed in this paper are my/our own. Some of the content is taken from open-source websites & some are copyright-free to disseminate knowledge. Those some had been mentioned above in the references section and acknowledged/cited as when and where required. The author/s have cited their joint own work mostly, and tables/data from other referenced sources in this particular paper with the narrative & endorsement have been presented within quotes and reference at the bottom of the article accordingly & appropriately. Finally, some of the contents are taken or overlapped from open-source websites for knowledge purposes and mentioned in the references section. On the other hand, opinions expressed in this paper are those of the author/s and do not reflect the views of the GJEIS. The author/s have made every effort to ensure that the information in this paper is correct, any remaining errors and deficiencies are solely their responsibility.