Original Article

Exploring Popular ETL Testing Techniques

Dhamotharan Seenivasan

Project Lead-Systems, Mphasis, Texas, USA.

Abstract *- ETL (Extract, Transform, Load) testing is an essential process in ensuring the accuracy, completeness, and consistency of data throughout the ETL process. In this article, we provide an overview of the ETL testing process and explore various ETL testing techniques, including data completeness, accuracy, integrity, transformation, performance, user acceptance, and regression testing. We also discuss best practices for ETL testing, common challenges, and the importance of ETL testing in maintaining high-quality data. By implementing effective ETL testing techniques, organizations can ensure the reliability and validity of their data, leading to better business decisions and outcomes.*

Keywords *- Data warehouse, ETL testing, Extract Transform and Load (ETL), ETL performance, ETL optimization.*

1. Introduction

Extract, Transform, Load (ETL) is a vital process for organizations to manage their data effectively. It involves extracting data from various sources, transforming the data into a consistent format, and then loading it into a destination system for further analysis. As data volumes continue to increase, organizations increasingly rely on ETL processes to manage their data efficiently. However, the ETL process can introduce errors, leading to inaccurate and inconsistent data. ETL testing is critical for ensuring data accuracy, completeness, and consistency as it moves through the ETL process. In this article, we explore various ETL testing techniques that can help organizations ensure their data's reliability.

2. Literature review

Extract, Transform, and Load (ETL) testing techniques have gained significant attention in recent years due to the growing importance of data quality and the need for organizations to optimize data processing efficiency. ETL testing is a critical aspect of the ETL process. Organizations may encounter significant data quality issues without effective ETL testing, leading to inaccurate reporting, decision-making, and, ultimately, negative business outcomes. This section will review some existing literature on ETL testing techniques and the best practices for ETL testing.

According to a study by Theobald (2007), data completeness testing is a crucial aspect of ETL testing that helps ensure that all required data has been successfully extracted from source systems and loaded into the target system. The study also highlights the importance of data accuracy testing, data transformation testing, and performance testing.

In their article "Automated ETL testing on the data quality of a data warehouse", Dakrory, Mahmoud and Ali discussed the challenges and benefits of automated ETL testing. They emphasize the importance of automation in ensuring the accuracy and consistency of data during the ETL process and provide practical advice on implementing the automated testing.

In another study, "Data Warehouse/ETL Testing: Best Practices", Mookerjea and Malisetty (2008) analyzed the challenges and best practices associated with ETL testing in data warehousing. They identified various challenges, including complex data environments, lack of test data, ETL tool limitations, and resource constraints. They provided best practices for addressing these challenges, such as using test automation, monitoring performance, and documenting the testing process.

In their research paper "Data warehouse testing", Golfarelli and Rizzi (2011) proposed a comprehensive framework for ETL testing that included eight steps, from requirements gathering to testing execution. They also identified several testing types, such as data completeness, data consistency, and data transformation testing, and provided guidelines for conducting each type of test. The study concluded that the proposed framework could help organizations improve the effectiveness and efficiency of their ETL testing processes.

In his book, "Data Warehouse ETL Toolkit," Kimball (2016) provided a practical guide for designing and testing ETL processes in data warehouse systems. He emphasized the importance of using real-world data for testing and provided guidelines for developing a comprehensive testing plan that includes testing data completeness, accuracy, consistency, and transformation.

In a study by Vandana and Sujatha (2013), ETL testing helped an organization identify and fix data quality issues leading to inaccurate reporting and analysis. The study demonstrates how organizations can use ETL testing techniques such as source-to-target data counts, source data sampling, data profiling, and business rule validation to ensure data accuracy and completeness.

The existing research on ETL testing highlights the different types of ETL testing techniques, tools and methodologies used for ETL testing and best practices for ETL testing. The literature also provides guidance on conducting ETL testing effectively, including using test automation, monitoring performance, and documenting the testing process.

3. ETL Testing Process

The ETL testing process involves several key steps, including extraction testing, transformation testing, and loading testing.

Extraction testing involves verifying that the data has been extracted from the source systems correctly. This involves checking that all the data has been extracted and is consistent with the source system.

Transformation testing involves ensuring that the data has been transformed correctly. This includes verifying that data has been cleansed, standardized, and enriched as required and that the business rules and logic have been applied correctly.

Loading testing involves verifying that the data has been loaded into the target system correctly. This includes checking that the data has been loaded completely, that it is consistent with the target system, and that there are no data validation errors.

Throughout the ETL testing process, it is essential to ensure that the data quality is maintained and that any errors are identified and resolved promptly. By following a robust ETL testing process, organizations can ensure that their data is reliable and accurate, which is critical for making informed decisions.

4. ETL Testing Techniques

In this section, we will explore the different types of ETL testing techniques that organizations can use to ensure data accuracy, completeness, and consistency.

4.1. Data Completeness Testing

Data completeness testing is a crucial component of ETL testing, and it ensures that all the required data has been extracted from the source system. In this testing, testers verify that all the expected data has been extracted from the

source system and that there are no missing records or data fields. To perform data completeness testing, testers must first understand the source system's data model and the data that needs to be extracted. They should compare the expected data with the data that has been extracted to identify any missing data fields or records.

For instance, consider a scenario where an organization needs to extract customer data from a Customer Relationship Management (CRM) system. The expected data fields may include the customer's name, address, phone number, email address, and other relevant information. In data completeness testing, testers would verify that all the required fields have been extracted for each customer record. Testers must identify the root cause and take the necessary corrective actions if any data fields or records are missing. The corrective actions may include modifying the ETL process or the source system to ensure that all the required data is extracted.

Here are some techniques that organizations can use for data completeness testing:

4.1.1. Source-to-target Data Counts

This technique involves comparing the number of records in the source system to the number of records loaded into the target system. If the counts do not match, it indicates that some data may be missing or incomplete.

4.1.2. Source Data Sampling

This technique involves randomly selecting a sample of data from the source system and verifying that it has been successfully loaded into the target system.

4.1.3. Business Rule Validation

This technique validates that all data required to meet the business rules and requirements have been successfully loaded into the target system. This helps ensure that the ETL process has successfully extracted and loaded all required data.

4.2. Data Accuracy Testing

Data accuracy testing is another essential component of ETL testing, and it ensures that the data extracted from the source system is accurate. In this testing, testers verify that the extracted data values are correct and that there are no data type or data format errors. Data accuracy testing is a crucial aspect of the ETL process, ensuring that data has been accurately transformed and loaded into the target system. This testing type helps ensure that data is consistent with the source systems and that calculations and aggregations have been performed correctly. To perform data accuracy testing, testers need to understand the source system's data model and the data that needs to be extracted. They should compare the extracted data with the expected data values to identify any discrepancies.

For example, consider a scenario where an organization needs to extract sales data from the point of Sale (POS)system. The expected data fields may include product name, price, quantity, and total sales amount. Testers would verify that the extracted data values match the expected values for each sales transaction in data accuracy testing. If there are any discrepancies between the extracted and expected data values, testers must resolve the issue.

Below are some techniques that organizations can use for data accuracy testing:

4.2.1. Rule-based Testing
This technique involves testing the data against the business rules and requirements to ensure that the data has been transformed and loaded correctly. For example, if the business rule requires a certain percentage of data to be filtered out, this technique will verify that the percentage has been met.

4.2.2. Data Profiling
This technique involves analyzing the data to identify any outliers or anomalies indicating data quality issues. By identifying and correcting these issues, data accuracy can be improved.

4.2.3. Aggregation Testing
This technique involves testing the accuracy of aggregations and calculations performed on the data, such as sums and averages. This ensures that the data has been transformed accurately and the correct calculations.

4.3. Data Consistency Testing
Data consistency testing ensures that the data extracted from the source system is consistent with the source system. In this testing, testers verify that the extracted data is complete and accurate and that there are no data duplication or data redundancy issues. Testers should compare the extracted data with the source system to identify any data inconsistencies. Data consistency testing is an essential aspect of the ETL process that ensures that data is consistent across all systems and platforms. This type of testing helps to identify any data discrepancies or inconsistencies that may exist between the source and target systems.

Here are some techniques used for data consistency testing:

4.3.1. Referential Integrity Testing
This technique involves testing the relationships between different data elements to ensure that the data is consistent across all systems. For example, if a customer record has a relationship with an order record, this technique will verify that the relationship exists in both the source and target systems.

4.3.2. Historical Data Validation
This technique involves testing the historical data to ensure consistency across all systems. This helps to ensure that the data in the target system is consistent with the data in the source system over time.

4.3.3. Duplicate Record Identification
This technique involves identifying and eliminating duplicate records that may exist in the source or target system. This helps to ensure that the data is consistent across all systems and that there are no redundant records.

4.4. Data Integrity Testing
Data integrity testing ensures that the data extracted from the source system maintains its integrity throughout the ETL process. In this testing, testers verify that the data is not lost, corrupted, or altered during the ETL process. To perform data integrity testing, testers need to understand the ETL process and the data that needs to be extracted. They should identify the potential points where the data can be lost or corrupted and verify the data at each point to ensure its integrity.

For example, consider a scenario where an organization needs to extract data from a legacy system to a modern data warehouse. In data integrity testing, testers would verify that the data is not lost or corrupted during the extraction, transformation, and loading process. If there are any data integrity issues, testers must identify the root cause and resolve it.

Some of the integrity testing techniques are listed below.

4.4.1. Source System Validation
This technique involves validating the data in the source system to ensure that it is accurate and complete. This helps to ensure that the data is not corrupted or altered before it is transformed and loaded into the target system.

4.4.2. Data Mapping Validation
This technique involves verifying that the data has been correctly mapped from the source system to the target system. This helps ensure that the data is not lost or corrupted during the transformation process.

4.4.3. Data Validation and Reconciliation
This technique verifies that the source and target systems data are consistent and reconciled. This helps to ensure that the data is not lost or corrupted during the ETL process and that the data is accurate and complete.

These techniques help to ensure that the data is trustworthy and can be used for reporting, analysis, and decision-making purposes.

4.5. Data Transformation Testing

This technique ensures that the data is transformed correctly during the ETL process. This type of testing focuses on verifying that the data has been transformed according to the business rules and requirements. To perform transformation testing, testers need to understand the ETL process and the business rules that need to be applied during the transformation. They should identify the transformation rules, data dependencies, and data relationships and ensure the data transformation rules are correctly applied.

For example, consider a scenario where an organization needs to transform sales data into a format suitable for analysis. The transformation rules may include aggregating sales data by region, product, and date. In transformation testing, testers would verify that the transformation rules are correctly applied and that there are no data inconsistencies or data loss issues.

Below are some of the data transformation testing techniques used in organizations.

4.5.1. Input/output testing

This technique involves comparing the input and output data to ensure that the data has been transformed correctly. It helps to ensure that the data is consistent with the business rules and requirements.

4.5.2. Data Transformation Validation

This technique involves validating that the data has been correctly transformed during the ETL process.

4.5.3. Error Handling Testing

This technique involves testing the ETL process to ensure errors are handled correctly during the transformation process. This helps ensure that the data is not lost or corrupted during the ETL process.

4.5.4. Transformation Rule Testing

This technique involves testing the transformation rules to ensure they are correctly implemented during the ETL process. This helps to ensure that the data is transformed according to the business rules and requirements.

4.6. Performance Testing

Performance testing ensures that the ETL process meets the performance requirements. The performance of the ETL process can impact the overall system performance, and any bottlenecks or performance issues in the ETL process can lead to delays and increased processing times.

Performance testing focuses on measuring the performance and scalability of the ETL system. This testing type helps identify any performance bottlenecks, errors, or issues that may affect the overall system performance.

Here are some techniques used for performance testing:

4.6.1. Load Testing

It involves testing the ETL process under normal and peak load conditions to identify the system's maximum processing capacity and any performance issues under different loads.

4.6.2. Stress Testing

Stress testing involves testing the ETL process under extreme load conditions to identify any system failures, bottlenecks, and other performance issues. Stress testing helps organizations identify the system's breaking point and take the necessary corrective actions.

4.6.3. Volume Testing

It involves testing the ETL process with large volumes of data to ensure that the ETL process can handle large data volumes without any performance issues. Volume testing helps organizations identify any data-related issues that can impact the system's performance and take the necessary corrective actions.

4.6.4. End-to-end Testing

This technique involves testing the entire ETL process end-to-end to measure the performance and scalability of the system. This helps identify performance bottlenecks, errors, or issues affecting the overall system performance.

By using these techniques, organizations can effectively test the performance and scalability of the ETL system and ensure that the system can handle large volumes of data and perform well under stress conditions. This helps to ensure that the ETL system is performing optimally and can be used for reporting, analysis, and decision-making purposes.

4.7. User Acceptance Testing

User Acceptance Testing (UAT) ensures that the ETL process meets the user requirements and is fit for purpose. UAT involves testing the ETL process with a sample of end-users to ensure that it meets their expectations, is user-friendly, and meets the business requirements. UAT is performed in a controlled environment to ensure that the system's performance, functionality, and user interface meet the user requirements. Testers can use scenarios, test cases, and user stories to test the ETL process and ensure it meets the user requirements.

During UAT, testers can collect end-user feedback and incorporate it to improve the system's performance and functionality. UAT helps identify any issues, errors, or gaps in the ETL process that must be addressed before the system is released to production.

UAT can also help organizations to ensure that the ETL process meets the user requirements, reduces the risk of user

dissatisfaction, and increases the system's adoption rate. By involving end-users in the testing process, organizations can ensure that the system is fit for purpose and meets the business requirements, leading to improved decision-making and overall business performance.

4.8. Regression Testing

Regression testing ensures that any changes or updates to the ETL process do not affect the existing functionality or performance of the system. Regression testing involves testing the entire ETL process, including all the associated data sources, transformations, and destinations, to ensure that any changes or updates have not introduced any new issues or errors. Testers can use automated testing tools to re-run the existing test cases and compare the results with the previous test runs to perform regression testing. Any differences or inconsistencies in the results can help identify any issues or errors introduced due to the changes or updates.

Regression testing can help organizations to reduce the risk of system failures, performance issues, and data errors caused by changes or updates to the ETL process. By performing regular regression testing, organizations can ensure that the ETL process meets the user requirements, is fit for purpose, and performs optimally, leading to improved business performance and decision-making.

Some of the regression testing techniques are listed below:

4.8.1. Re-testing

This technique involves re-testing the ETL system after updates or changes to ensure that the system still operates correctly and has not introduced any new issues.

4.8.2. Automated Testing

This technique involves creating automated regression tests that can be run quickly and easily after updates or changes to the ETL system. This helps to ensure that the system still operates correctly and has not introduced any new issues.

4.8.3. Impact Analysis Testing

This technique involves analyzing the impact of updates or changes to the ETL system to identify any areas that may be affected and require testing.

4.8.4. Version Control Testing

This technique involves testing different versions of the ETL system to ensure that they still operate correctly and have not introduced any new issues.

4.9. Production Validation Testing

Production validation testing ensures that the ETL process works correctly in the production environment.

Production validation testing involves testing the ETL process with real-time or production data to ensure it works correctly and meets the user requirements and business objectives.

Production validation testing is typically performed after the ETL process has been deployed to the production environment. Testers can use scenarios, test cases, and user stories to test the ETL process and ensure it performs correctly in the production environment.

During production validation testing, testers can verify that the data is correctly extracted, transformed, and loaded into the destination systems. They can also verify the data quality, accuracy, completeness, and consistency, ensuring that the data is fit for purpose.

Production validation testing can help organizations to reduce the risk of data errors, system failures, and performance issues caused by issues in the ETL process. By validating the ETL process in the production environment, organizations can ensure that the system works correctly and meets the business objectives and user requirements, leading to improved decision-making and business performance.

4.10. Application Migration Testing

Application migration testing is an ETL testing technique that organizations can use when they are migrating an application from one system to another. It involves testing the ETL process to ensure that the data is correctly extracted from the source system, transformed to meet the requirements of the destination system, and loaded into the destination system without any issues or errors. During application migration testing, testers can use various testing techniques, such as data completeness testing, data accuracy testing, data consistency testing, data integrity testing, and performance testing, to ensure that the ETL process meets the requirements of the destination system. Application migration testing can help organizations to ensure a smooth and seamless transition from one system to another. By identifying and addressing any issues or errors in the ETL process, organizations can ensure that the data is correctly migrated to the destination system, reducing the risk of data loss, system failures, or performance issues.

Moreover, application migration testing can help organizations to ensure that the data is correctly formatted, transformed, and loaded into the destination system, ensuring that it meets the business objectives and user requirements. By performing thorough application migration testing, organizations can minimize the risk of any data-related issues and ensure a successful migration to the new system.

Here are some techniques that organizations can use for application migration testing:

4.10.1. Compatibility Testing

It involves testing the ETL system for compatibility issues with the new environment or platform to ensure it operates correctly.

4.10.2. Configuration Testing

It involves testing the ETL system's configuration in the new environment or platform to ensure it has been set up correctly.

4.10.3. Data Migration Testing

This technique involves testing the ETL system's ability to migrate data to the new environment or platform to ensure that it operates correctly and produces accurate results.

4.11. Incremental Testing

This technique is used to test the ETL process as it is being developed incrementally rather than waiting until the entire process is complete before testing. Incremental testing involves testing small increments or batches of data as they are being extracted, transformed, and loaded into the destination systems.

Incremental testing allows organizations to identify and address any issues or errors in the ETL process as soon as they occur, reducing the risk of any major issues or errors occurring later in the development process. It also allows organizations to test the ETL process iteratively, making it easier to identify and fix issues as they arise.

Moreover, incremental testing can help organizations to reduce the time and cost of testing, as they can identify and address any issues or errors in the ETL process early in the development process rather than waiting until the entire process is complete before testing. By using incremental testing, organizations can ensure that the ETL process works correctly and meets user requirements and business objectives.

5. Best Practices for ETL Testing

To ensure the success of ETL testing, organizations can follow several best practices:

5.1. Define and Document Test Objectives

It is important to clearly define and document the test objectives to ensure that the testing is focused and effective.

5.2. Test Data Preparation

Use realistic test data that covers a variety of scenarios and edge cases to ensure that the ETL system is thoroughly tested.

5.3. Create Test Cases

Create detailed test cases that cover all aspects of the ETL system, including data completeness, accuracy, consistency, and performance.

5.4. Automate Testing

Automate as much of the testing process as possible to improve efficiency and reduce the risk of human error.

5.5. Perform End-to-End Testing

Perform end-to-end testing of the entire ETL process to ensure that the system operates correctly and produces accurate results.

5.6. Perform Regression Testing

Perform regression testing after changes or updates to the ETL system to ensure they have not introduced any new issues.

5.7. Collaborate with Stakeholders

Work closely with stakeholders, such as developers, business analysts, and end users, to ensure that the ETL system meets their requirements and produces accurate results.

5.8. Document and Report Issues

Document and report any issues or defects found during testing to ensure that they are addressed and resolved in a timely manner.

5.9. Use Real-world Data

Using real-world data in ETL testing can help ensure that the ETL process works correctly and meets the user requirements and business objectives.

5.10. Monitor performance

Monitoring performance during ETL testing can help organizations identify performance issues, such as slow data processing or data overloading, and address them before they impact the production environment.

By following these best practices, organizations can ensure that their ETL process is tested thoroughly and meets the user requirements and business objectives, leading to improved decision-making and business performance.

6. Challenges of ETL Testing

ETL testing can be challenging due to various factors, including:

6.1. Data Volume and Complexity

ETL systems process large volumes of complex data, which can be challenging to test thoroughly.

6.2. Data Quality Issues

ETL systems can be impacted by data quality issues, such as missing or incorrect data, which can be challenging to identify and resolve.

6.3. Performance Issues

ETL systems can be impacted by performance issues, such as slow load times, which can be challenging to identify and optimize.

6.4. Integration with Other Systems

ETL systems often integrate with other systems, such as data warehouses or reporting tools, which can be challenging to test and ensure that the integration operates correctly.

6.5. Managing Test Data

ETL testing requires managing large volumes of test data, which can be challenging to create, maintain, and manage.

6.6. Lack of Skilled Resources

ETL testing requires skilled resources with experience in data testing, which can be challenging to find and train.

6.7. Version Control

ETL processes may involve multiple versions of data sources, making it challenging to ensure that the correct data version is used during the testing process.

6.8. Test Environment Setup

Setting up the test environment for ETL testing can be challenging. It involves creating a replica of the production environment and ensuring that all the required data sources, tools, and systems are available for testing.

6.9. Time and Cost

ETL testing can be time-consuming and expensive, particularly when testing large volumes of data and complex ETL processes.

By understanding these challenges and developing strategies to address them, organizations can improve the effectiveness of their ETL testing and ensure that their ETL process is robust and reliable.

7. Conclusion

ETL testing is a critical process that organizations must undertake to ensure that their data is accurate and consistent. Various testing techniques, such as data completeness testing, data accuracy testing, data consistency testing, data integrity testing, data transformation testing, performance testing, user acceptance testing, regression testing, production validation testing, application migration testing,

and incremental testing, can be used to test the ETL system thoroughly.

To ensure that the ETL system operates correctly and produces accurate results, organizations can follow best practices, such as defining clear test objectives, creating detailed test cases, automating testing, performing end-to-end testing, collaborating with stakeholders, and documenting and reporting issues. Despite the challenges of ETL testing, organizations can overcome them by adopting best practices, using tools and technologies, and investing in training and development programs. By thoroughly testing the ETL system and ensuring that it operates correctly and produces accurate results, organizations can make better-informed decisions, improve operational efficiency, and gain a competitive advantage.

The field of ETL testing is continuously evolving, with new technologies and techniques emerging to improve the efficiency and effectiveness of the process.

Some future directions for ETL testing include the following:

Machine learning and artificial intelligence can be used to identify patterns and anomalies in data, making it easier to detect errors and identify potential issues in the ETL process. Cloud-based ETL testing can provide organizations greater scalability, flexibility, and cost-effectiveness than traditional on-premises ETL testing.

Test automation is becoming increasingly popular in ETL testing, with organizations using tools and frameworks to automate the testing process, leading to greater speed and efficiency.

Data virtualization is an emerging technology that allows organizations to access and use data from different sources without having to physically move the data, reducing the need for ETL testing.

Collaborative testing involves involving multiple stakeholders in the ETL testing process, such as developers, testers, and business users, leading to improved collaboration, communication, and decision-making.

By adopting these and other future directions, organizations can improve the efficiency and effectiveness of their ETL testing processes, leading to improved business outcomes and competitive advantage.

References

[1] Theobald J., "Strategies for Testing Data Warehouse Applications," *Information Management*, vol. 17, no. 6, p. 20, 2007.

[2] Mahendra A Sethi, and Santosh S Lomte, "An Interactive Dynamic Elearning Framework for Visual and Verbal Learners," *SSRG International Journal of Computer Science and Engineering*, vol. 4, no. 7, pp. 3-13, 2017. *Crossref*, https://doi.org/10.14445/23488387/IJCSE-V4I7P102

[3] Sara B. Dakrory, Tarek M. Mahmoud, and Abdelmgeid A. Ali, "Automated ETL Testing on the Data Quality of a Data Warehouse," *International Journal of Computer Applications*, vol. 131, no. 16, pp. 9-16, 2015. *Crossref*, http://dx.doi.org/10.5120/ijca2015907590

[4] Vandana KV., and Sujatha V., *ETL Testing in Datawarehousing*, 2013

[5] Baljit Singh, "Enterprise Reporting on SAP S/4HANA Using Snowflake as Cloud Datawarehouse," *International Journal of Computer Trends and Technology*, vol. 71, no. 1, pp. 45-49, 2023. *Crossref*, https://doi.org/10.14445/22312803/IJCTT-V71I1P107

[6] Mookerjea A, and Malisetty P, *Data Warehouse/ETL Testing: Best Practices*, Proc. Test (Test Excellence Through Speed And Technology), New Delhi, India. 2008

[7] Ralph Kimball, and Joe Caserta, *The Data Warehouse ETL Toolkit*, John Wiley & Sons, 2004

[8] Matteo Golfarelli, and Stefano Rizzi, "Data Warehouse Testing," *International Journal of Data Warehousing and Mining (IJDWM)*, vol. 7, no. 2, pp. 26-43, 2011.

[9] Mitesh Athwani, "*A Novel Approach to Version XML Data Warehouse*," *SSRG International Journal of Computer Science and Engineering*, vol. 8, no. 9, pp. 5-11, 2021. *Crossref*, https://doi.org/10.14445/23488387/IJCSE-V8I9P102

[10] Dhamotharan Seenivasan, "ETL (Extract, Transform, Load) Best Practices," *International Journal of Computer Trends and Technology,* vol. 71, no. 1, pp. 40-44, 2023. *Crossref*, https://doi.org/10.14445/22312803/IJCTT-V71I1P106

[11] Sonali Vyas, and Pragya Vaishnav, "A comparative study of various ETL process and their testing techniques in data warehouse," *Journal of Statistics and Management Systems,* vol. 20, no. 4, 2017. *Crossref*, https://doi.org/10.1080/09720510.2017.1395194

[12] [Online]. Available: https://www.researchgate.net/publication/368300555_ETL_for_Data_Warehousing

[13] Dhamotharan Seenivasan, "ETL (Extract, Transform, Load) Best Practices," *International Journal of Computer Trends and Technology*, vol. 71, no. 1, pp. 40-44, 2023. *Crossref*, https://doi.org/10.14445/22312803/IJCTT-V71I1P106

[14] [Online]. Available:https://www.zyxware.com/articles/5152/what-are-the-challenges-of-etl-testing

[15] [Online]. Available: https://www.tutorialspoint.com/etl_testing/etl_testing_challenges.htm

[16] [Online]. Available: https://www.guru99.com/utlimate-guide-etl-datawarehouse-testing.html

[17] [Online]. Available: https://www.softwaretestinghelp.com/etl-testing-data-warehouse-testing/

[18] [Online]. Available: https://www.cigniti.com/blog/conquering-the-challenges-of-data-warehouse-etl-testing/

[19] [Online]. Available: https://www.altexsoft.com/blog/etl-testing/

[20] Dhamotharan Seenivasan, Improving Performance of the ETL Jobs, IIOAB, 2023

[21] [Online]. Available: https://hevodata.com/learn/etl-testing/

[22] [Online]. Available: https://www.datagaps.com/data-testing-concepts/etl-testing/

[23] [Online]. Available: https://www.matillion.com/resources/blog/what-are-the-basics-of-etl-testing

[24] [Online]. Available: https://www.testingxperts.com/blog/etl-testing-guide

[25] Mr. Sujit Prajapati, and Mrs. Sarala Mary, "Study of ETL Process and Its Testing Techniques," *Ijraset Journal for Research in Applied Science and Engineering Technology,* vol. 10, no. 6, 2022. *Crossref*, https://doi.org/10.22214/ijraset.2022.43931

[26] [Online]. Available: https://dataq.io/etl-testing/

[27] [Online]. Available: https://impactqa.medium.com/etl-test-automation-for-better-business-operations-749d768929bc

[28] [Online]. Available: https://www.javatpoint.com/etl-testing

[29] [Online]. Available: https://www.tricentis.com/blog/key-considerations-data-warehousing-etl-test-automation

[30] [Online]. Available: https://www.testing-whiz.com/blog/importance-of-etl-testing-process-challenges-and-future/

[31] [Online]. Available: https://solutionsreview.com/data-integration/the-best-etl-testing-tools-for-data-integration-success/

[32] [Online]. Available: https://www.prolifics-testing.com/testing-services/projects/etl-testing

[33] [Online]. Available: https://dataintegrationinfo.com/data-validation-testing/