



InfoSphere : **NAVIGATING THE WORLD** **OF INFORMATION**

Preserving the Past, Valuing Present, Enriching the Future



**UNIVERSITI
TEKNOLOGI
MARA**

Fakulti
Sains Maklumat

InfoSphere:

Navigating the World of Information

Copyright © 2025 Faculty of Information Science

Published by Universiti Teknologi MARA Cawangan Johor

**Universiti Teknologi MARA, Cawangan Johor
Faculty of Information Science
Jalan Universiti, KM. 12, Off Jalan Muar,
85000 Segamat, Johor Darul Ta'zim, Malaysia**

All rights reserved. No part of this book may be reproduced, distributed, or transmitted in any form or by any means without prior written permission from the publisher, except for brief quotations in reviews. Published content does not necessarily reflect the publisher's views. For questions, please contact the publisher.

ISBN 978-629-7647-07-4

Disclaimer All content in this publication has been prepared accordingly with the utmost efforts of the author and publisher. The author and publisher make no representation or warranties with respect to the accuracy, applicability, fitness or completeness of the contents of this book.

Preface

It gives me great pleasure to present *InfoSphere: Navigating the World of Information*, a collective effort that brings together insightful discussions, empirical findings, and critical reflections from academics in information management, library management, records management, and information and communication technology (ICT).

In today's digital era, the vast and complex landscape of information continues to expand at an unprecedented pace. The convergence of technology and information systems has transformed how data is created, organized, stored, and utilized. This book aims to navigate that dynamic "infosphere", a term that reflects the interconnected environment in which information flows seamlessly across platforms, disciplines, and contexts.

The chapters compiled here explore diverse perspectives and contemporary issues shaping the management of information resources and services. From emerging trends in digital librarianship and knowledge governance to innovations in ICT applications and recordkeeping practices, each contribution highlights the growing importance of integrating technology, policy, and human expertise in managing information effectively.

This publication also serves as a reflection of our faculty's commitment to advancing scholarship and practice in the information domain. It showcases the intellectual depth and interdisciplinary collaboration among academics who continue to shape the future of information work.

As the Chief Editor, I wish to express my sincere appreciation to all the contributing authors for their dedication and scholarly rigor, and to the editorial team for their meticulous effort in ensuring the quality and coherence of this volume. Special thanks are also extended to the Faculty of Information Science and Universiti Teknologi MARA Cawangan Johor for their continuous support and encouragement in realizing this publication.

I hope that *InfoSphere: Navigating the World of Information* will inspire readers among students, educators, researchers, and practitioners alike to engage critically with the evolving information landscape and to contribute meaningfully to its advancement.

Azura Abdul Jamil @ Kamarudzzaman

Chief Editor

InfoSphere: Navigating the World of Information

2025

Table of Contents

Sections 1 - Information Management

Early History of Information Governance (IG) and its implementation in Malaysia <i>Nor Diana Abd Rahman, Rohayu Ahmad, Azura A.Jamil@Kamarudzzaman, Dr Siti Nuur-Ila Mat Kamal</i>	1
Information Skills in Students Learning and Development <i>Nurfaizah Kamarudin, Ahmad Fuzi Md Ajis, Siti Hajar Baharin, and Siti Nur Shahira Dahari</i>	7
Digital Transformation in Information Management <i>Nor Rashidah Mohamed and Maziah Mustapa</i>	12
Understanding Social Media for Information Professionals <i>Rabiatul Adawiyah Kamarulzaman, Nurul Huda Izzati Jamil, Razifah Othman and Suhaila Osman</i>	19
Introduction to Metadata <i>Siti Hajar Baharin², Siti Nur Shahira Dahari¹, Nurfaizah Kamarudin³ and Ahmad Fuzi Md Ajis</i>	28

Section 2 - Library Science

Information Professionals in Library Context <i>Siti Nur Shahira Dahari¹, Siti Hajar Baharin², Nurfaizah Kamarudin³ and Ahmad Fuzi Md Ajis</i>	34
--	----

Section 3 - Records Management

Records Management Role in Organizational Memory <i>Azura Abdul Jamil @ Kamarudzzaman, Nor Diana Abd Rahman, Rohayu Ahmad & Dr Siti Nuur-Ila Mat Kamal</i>	39
Significance of Electronic Records in Higher Learning Institutions <i>Rohayu Ahmad, Nor Diana Abd Rahman, Azura A.Jamil@Kamarudzzaman, Dr Siti Nuur-Ila Mat Kamal</i>	43

Section 4 - Information System Management

Internet of Things (IoT): The role of an information literacy skills <i>Faten Elina Kamaruddin, Isma Ishak and Ahmad Fuzi Md Ajis</i>	46
Cloud Computing in Information Management <i>Isma bin Ishak, Faten Elina binti Kamaruddin</i>	53

Section 5 - Information Analytics

Information Analytics: influences of crucial Data Pre-Processing Technique <i>Ahmad Fuzi Md Ajis, Nurfaizah Kamarudin, Siti Hajar Baharin, Siti Nur Shahira Dahari</i>	60
---	----

INFORMATION ANALYTICS : INFLUENCE OF CRUCIAL DATA PRE-PROCESSING PHASES.

By Ahmad Fuzi Md Ajis, Nurfaizah Kamarudin, Siti Hajar Baharin, Siti Nur Shahira Dahari
Email: ahmadfuzi@uitm.edu.my

1.0 Introduction

Information analytics is one crucial process in any organization on obtaining insights based on existing data to facilitate data-driven decision making (Koukaras & Tjortjis, 2025). While data and information were crucially influencing the sustainability of organizations operation and productivity, the whole process of data would so much rely on how such data is being processed (Martins et al., 2025). However, while many look towards the information or data processing, the pre-processing stage influences the quality of data to be prepared for the actual data processing phase (Lopez-Miguel, 2021).

Pre-processing phases is an operation whereby raw data is prepared to be structured in enabling it to be analyzed (Martins et al., 2025). This phase happened before the process of the ETL (extract, transform and load) stage. It involves major pre-processing operations such as Data Cleaning, Data Integration, Data Transformation and Data Reduction (Koukaras & Tjortjis, 2025).

2.0 Influence of Pre-processing Phase towards Information Analytics

The pre-processing phase started earlier before ETL stage and the accuracy, timeliness and fitness to be used by organizations during analysis would influence the findings from the analytics process (Martins et al., 2025). Table 1 provides the goal and key activities of the pre-processing phase which highlight the major focus of each step in structuring the raw data before engaging any data analysis process (Lopez-Miguel, 2021).

Data Cleaning

Essential focus on the data cleaning is to focus on accurate and consistent data. This is the initial stage where we structure the datasets appropriately to guarantee its quality and reliability. To achieve this, the key activities are the crucial elements which involve identifying and rectifying errors, removing inconsistencies, and handling missing values (Côté et al., 2023). Some analysts may use popular software like PowerBi and Google Refine to execute data cleaning activities before engaging in any data analysis (Martins et al., 2025).

Table 1

Goal and Key activities of the pre-processing phase.

Pre-processing Phase	Primary Goal / Purpose	Key Activities / Focus	Important Notes / Tools
1. Data Cleaning	To achieve accurate and consistent data .	Error detection and correction, handling missing values and outliers, consistency checks, label / feature cleaning (for ML) (Côté et al., 2023)	Tool performance matters (e.g. PowerBi, OpenRefine, PyJanitor, Great Expectations) when data is large and domains vary (Martins et al., 2025)
2. Data Integration	To create a unified data set .	Combining various data sets.	This step is essential when working with multiple sources of raw information. An example is combining date, month, and year into a single, cohesive unit.
3. Data Transformation	To ensure data is adequately prepared to "provide and generate valuable data and understandable data ".	Converting the data, adjusting the format of the data , standardizing the units, and normalizing the variable.	This process standardizes and adjusts the format of the variables.
4. Data Reduction	To reduce the dimensionality of the data while ensuring that all of its essential information is retained.	Trying to describe emptiness in the data, addressing irrelevant data, and performing structural reduction.	This process does not mean deleting or changing the original data. It manages the volume and relevance of the data.

Data Integration

Once a cleansed data is achieved, unified and compiled datasets are very important when the information or data that you have coming from multiple sources and a holistic perspective are required to be churned out from the data. However, the challenge comes in the aspect of harmonizing the data into cohesive and logical structure. Part of the challenge is the schema matching where data from multiple sources may have a variety of data schemas and structures which harmonization of data would become difficult to prevent misintegration of data. This is commonly due to the known cause called entity identification where data from different systems, schemas and structures could possibly represent similar data representation or vice versa (Wang et al., 2022).

Data Transformation

After the unification of the data completed, standardization and adjustments for the appropriate data format should be executed for optimum processing. Data transformation ensures the objective of preparing data for the processing stage with intention of generating valuable and insightful findings is achievable. Data will be converted or adjusted to the format of the required operation, standardizing the units, and normalizing the variable to ensure optimization of data preparedness for major analysis (Koukaras & Tjortjis, 2025).

Data Reduction

While maintaining the essential information retained in the data, dimension of the data requires reduction which intends to reduce and manage the data complexity and computational cost. As the volume of data explodes, the complexity of analysis may increase and prohibit analysis to be done thoroughly with less time consuming. Although reduction may be defined as deleting or changing data, but instead of deleting, data reduction can be rectified as summarizing information without deleting important aspects of original data. Besides, its major focus is to describe emptiness in the data, addressing irrelevant data, and performing structural reduction (Hancock et al., 2024; Zhang et al., 2023).

3.0 Conclusions

In data science, the pre-processing stage is the “behind-the-scene” operation that lies beneath any effective and accurate information analytics. The quality information derives from the final insights entirely depends on these four initial steps whereby ignoring part of them would expose the data to bias and inaccurate or flawed analysis. These foundational works are essentially required to avoid data chaos to generate reliable analytical information (Koukaras & Tjortjis, 2025; Martins et al., 2025).

References

- Côté, M., Nikanjam, N., Ahmed, I., Humeniuk, K., & Khomh, F. (2023). Benchmarking data cleaning tools for large messy datasets. *IEEE International Conference on Big Data (Big Data)*, 1456–1465. <https://doi.org/10.1109/BigData59066.2023.10372543>
- Hancock, J., Wang, Y., Khoshgoftaar, T. M., & Napolitano, A. (2024). Data reduction techniques for highly imbalanced Medicare big data. *Journal of Big Data*, 11(1), 22. <https://doi.org/10.1186/s40537-023-00869-3>
- Koukaras, G., & Tjortjis, C. (2025). Data preprocessing for machine learning: A review and best practices. *Data*, 6(10), 257. <https://doi.org/10.3390/data6100257>
- Lopez-Miguel, A. (2021). A review of preprocessing techniques in big data analytics. *International Journal of Computer Applications*, 183(45), 1–8. <https://doi.org/10.5120/ijca2021921609>
- Martins, L. M., Cardoso, T. C., Váz, C., Silva, L. J., & Abbasi, A. (2025). Scalable data preprocessing for information analytics: A framework review. *Information Systems Frontiers*. <https://doi.org/10.1007/s10796-025-10331-2>
- Wang, Z., Li, S., & Wang, H. (2022). Sudowoodo: Contrastive self-supervised entity matching with semantic type detection and error correction. *arXiv preprint*. <https://arxiv.org/abs/2207.04122>
- Zhang, C., Zhao, Y., Yang, C., & Tan, B. (2023). On the effectiveness of dimensionality reduction techniques on high-dimensionality datasets. In *Proceedings of the International Conference on Machine Learning Applications (ICMLA)*. https://doi.org/10.1007/978-3-031-25344-7_15



UNIVERSITI
TEKNOLOGI
MARA

Fakulti
Sains Maklumat

ISBN 978-629-7647-07-4



9 786297 647074