



## DESIGN OF EFFICIENT FRAMEWORK OF INFORMATION AGGREGATION FROM VARIOUS WEBSITES USING WEB CONTENT MINING TECHNIQUES

DOI: <https://doi.org/10.5281/zenodo.23162610>

Dr. Narendra M. Jathe<sup>1</sup>, Ms. Lavanya S. Gahale<sup>2</sup>

\*Corresponding Author: <sup>1</sup>Assistant Professor, Department of Computer Science, Smt. Narsamma Arts, Commerce and Science College, Amravati, Maharashtra, India  
[njathe@gmail.com](mailto:njathe@gmail.com)

ORCID : <https://orcid.org/0000-0003-2927-7882>

<sup>2</sup>Research Scholar, Department of Computer Science, Smt. Narsamma Arts, Commerce and Science College, Amravati, Maharashtra, India  
[lavanyagahale22@gmail.com](mailto:lavanyagahale22@gmail.com)

ORCID: <https://orcid.org/0009-0009-5551-3608>

### ABSTRACT

The exponential growth of data generated from conventional databases and web-based sources has created significant challenges in effective data management, integration, and utilization. With the continuous evolution of digital technologies and online services, organizations are increasingly adopting adaptive strategies to extract meaningful insights from diverse and distributed data sources. Information aggregation has emerged as a crucial service in this context, enabling the consolidation, analysis, and presentation of relevant data from multiple platforms to enhance accessibility and support decision-making.

Traditionally, intermediaries such as newspapers, magazines, and professional journals served as aggregators. However, the emergence of web portals and digital platforms has significantly transformed this role by enabling seamless and large-scale integration of information. Information aggregation plays a vital role in the development of knowledge-based systems across various domains, including healthcare, economics, applied sciences, artificial intelligence, robotics, and machine learning.

Despite the vast amount of information available on the web, users often encounter difficulties due to data dispersion, heterogeneous formats, and geographical distribution. Aggregating such information involves addressing complex challenges related to data comparison, relationships, and both intra- and inter-organizational integration. Although existing techniques such as screen scraping, direct data feeds, and XML-based encoding are widely used, there is still considerable scope for improving aggregation methodologies. Universities represent a significant yet underexplored domain for information aggregation. While academic institutions publish extensive public information on their websites, this content is often fragmented across multiple sources. Therefore, this research proposes an efficient framework for aggregating information from various websites using web content mining techniques, with the aim of enhancing data accessibility, usability, and value creation for diverse user groups.

**KEYWORDS:** Information Aggregation, Web Content Mining, Knowledge-Based Systems, Data Integration, Web Data Extraction, Decision Support Systems, Digital Platforms.