



**National Conference on Recent Trends in Engineering, Science,  
Humanities and Management (NCRTESHM – 2023)**

29<sup>th</sup> January, 2023, West Bengal, India.

**CERTIFICATE NO : NCRTESHM /2023/C0123219**

**ENHANCING XML DATA RETRIEVAL PERFORMANCE WITH  
CLUSTERING AND INDEXING**

**WANJARI RAVINDRA SHANKAR**

Research Scholar, Department of Computer Science  
Kalinga University, Naya Raipur, Chhattisgarh, India.

**ABSTRACT**

Using SQL Server 2005 and Berkeley DBXML (BDBXML) as case studies, this research assesses how well a similarity-based clustering method retrieves XML data. If you're using SQL Server 2005, you can speed up data retrieval speeds by combining clustering and indexing. Applying clustering and indexing simultaneously, for instance, decreases the retrieval time for 10,000 items from 0.88 seconds to 0.46 seconds. The most efficient retrieval for 10,000 entries utilizing both clustering and indexing was 2.389 seconds in BDBXML, which demonstrates better retrieval speeds overall. With BDBXML demonstrating quicker retrieval speeds than SQL Server 2005 for big datasets, the results demonstrate that indexing and clustering improve performance. For big datasets utilized in similarity-based clustering tasks, this study shows that improving XML data retrieval by clustering and indexing is effective.

***Keywords: Similarity, Indexing, Retrieval, Database, Server***