Invited Paper
DOLAP: A 25 Year Journey Through Research Trends and Performance*

Tatsawan Timakum[1,†], Soobin Lee[2,†], Haotian Hu[3,†], Il-Yeol Song[4,*,†] and Min Song[2,†]

[1] *Department of Information Sciences, Chiang Mai Rajabhat University, Chiang Mai, Thailand*
[2] *Department of Library and Information Science, Yonsei University, Seoul, Korea*
[3] *School of Information Management, Nanjing University, Nanjing 210023, China*
[4] *College of Computing and Informatics, Drexel University, Philadelphia, PA, USA*

Abstract
The DOLAP workshop celebrated its 25th anniversary in 2023. This significant milestone reflects the enduring impact and importance of DOLAP as a platform for advancing research in topics covered in the forum. DOLAP, founded in 1998 as an ACM workshop alongside CIKM, served as a pioneering academic forum for research in Data Warehousing and OLAP. Since 2017, DOLAP has been co-located with EDBT, and it has evolved into the International Workshop on Design, Optimization, Languages and Analytical Processing of Big Data. This study aims to analyze the 25-year evolution (1998-2023) of DOLAP research. By analyzing 328 full-text DOLAP papers through the DMR topic modeling, the research employs bibliometrics and social network analysis for research performance and collaboration networks. The paper presents evolution of thematic trends, a variety of citation statistics, and collaboration networks among DOLAP authors.

Keywords
DOLAP, Data warehousing, OLAP, Big data, Topic modeling, Collaboration networks

1. Introduction

DOLAP (International Workshop on **D**ata Warehousing and **OLAP**) was founded in 1998 as an ACM workshop in conjunction with CIKM (The ACM International Conference on Information and Knowledge Management). DOLAP was founded as the first academic forum for research in data warehousing and OLAP areas. From 1998 to 2015, DOLAP has been a pioneering and representative forum where researchers and practitioners exchange knowledge and insights related to data warehousing and OLAP. DOLAP addressed a wide range of research topics such as data warehouse schema design, OLAP processing and patterns, materialized view processing, indexing, query optimization in data warehouses ETL, data integration issues with non-relational data such as XML or GIS, spatial data warehouse, XML data warehouses, security in data warehouses, etc. DOLAP's contributions significantly advanced the field of data warehousing and provided valuable insights for both researchers and practitioners.

Over the years, research in data warehouses has expanded to include big data, and research on OLAP has evolved to encompass data analytics. In response to the changing research landscape, DOLAP underwent a transformation and adopted a new name of the International Workshop on Design, Optimization, Languages and Analytical Processing of Big Data from 2017. (The new name was suggested by Carlos Ordonez, Univ of Houston.) The new DOLAP has been held in

DOLAP 2024: 26th International Workshop on Design, Optimization, Languages and Analytical Processing of Big Data, co-located with EDBT/ICDT 2024, March 25, 2024, Paestum, Italy
*Corresponding author.
†These authors contributed equally.

tatsawan_tim@cmru.ac.th (T. Timakum); bini122@yonsei.ac.kr (S. Lee); hhtdlam@gmail.com (H. Hu); songiy@drexel.edu (I. Song); min.song@yonsei.ac.kr (M. Song)
0000-0002-9877-0323 (T. Timakum); 0000-0002-7515-0199 (S. Lee); 0000-0002-4112-8719 (H. Hu); 0000-0001-7706-959X (I. Song); 0000-0003-3255-1600 (M. Song)

conjunction with EDBT (The International Conference on Extending Database Technology) conferences.

The DOLAP forum celebrated its 25th anniversary in 2023. This significant milestone reflects the enduring impact and importance of DOLAP as a platform for advancing research in data warehousing, OLAP, big data, and data analytics. Over these 25 years, DOLAP has published numerous innovative research findings and significant results. However, quantitative research assessment of its overall influence and significance has been lacking. To address this gap, this paper aims to analyze the 25-year evolution (1998-2023) of DOLAP research. We present research trends, performance metrics, and citation analysis of all papers published in DOLAPs between 1998 to 2023. By analyzing the 328 full-text DOLAP papers through the DMR topic modeling, we employ bibliometrics and social network analysis for research performance and collaboration networks. The paper highlights evolution of thematic trends, key DOLAP authors, and impactful papers that have contributed significantly to the field. Additionally, we visualize co-authorship networks to depict collaboration patterns among researchers.

This study provides valuable insights for researchers and practitioners interested in research areas on data warehousing, OLAP, big data, data analytics. and related domains. It enriches our understanding of the growth, impact, and collaborative dynamics within the DOLAP field.

2. Methods

2.1. Data collection

The dataset of full-text papers in PDF file format was collected from two sources. The ACM digital library (www.dl.acm.org/conference/dolap) provided papers published during the DOLAP workshops from 1998 to 2015. Papers from the years 2017 to 2023 were downloaded from the DBLP (https://dblp.org/db/conf/dolap/index.html). DOLAP 2016 was not held and thus no DOLAP 2016 papers exist. In total, the dataset comprises 328 records, which include regular papers, invited talk papers, short papers, and keynote papers.

The dataset containing bibliography information, including abstracts, citation counts, and cited references, was imported from the Scopus database (www.scopus.com). Additionally, we gathered citation data from Google Scholar and the ACM Digital Library to assess the performance of the DOLAP conference. These datasets were gathered in July, 2023.

2.2. Data processing and analysis

2.2.1. Data preprocessing

This study used the **PyPDF2** library to extract text from PDF full-text paper files and saved it to a plain text file. Afterward, the dataset was preprocessed for topic modeling analysis. First, unnecessary punctuation marks or symbols were removed from the sentences. Next, the dataset was split into sentences and then tokenized based on spaces. Later, the data was POS-tagged, and only the nouns were selected, retaining only the words. Lastly, the data with stop words removed was used for topic analysis.

2.2.2. Research metrics analysis

This process involves reviewing the statistics to provide an overview of the publication performance of the DOLAP workshops. We examined the influence of papers and authors using citation data. Our assessment included considerations of authorship, paper titles, and the total number of citations received. We consolidated all citation data sources into an Excel spreadsheet and visualized the research performance through a bar chart.

2.2.3. Topic modeling analysis

We used Dirichlet-Multinomial Regression (DMR) topic modeling method to estimate topic and word distributions, assuming that all documents contain potential topics [1]. The method explores topics while incorporating metadata such as publication year or author information [2]. In this study, we use the expanded process of the DMR topic model that includes a log-linear prior on document-topic distributions.

In this study, the year of publication was associated with each full text as metadata, and an experiment was conducted by dividing the full text into five sentences for meaningful analysis. The total number of years considered in the experiment was 24, with the number of topics set to 20, including the list of top words in the probability distribution (30 top words). We employed the Python package *tomotopy* for DMR topic modeling analysis. Additionally, topics were automatically labeled using Pointwise Mutual Information (PMI), which measures the similarity of pairs of words [3].

2.2.4. Co-authorship network analysis

In co-authorship network analysis, authors are represented as nodes in a network. We create a link (edge) between nodes when authors co-authored a paper together. Analyzing the resulting co-authorship network can provide insights into collaboration patterns, the structure of research communities, and the impact of individual researchers or institutions. In our

study, we utilized cited publication information from the Scopus dataset, specifically those cited by the DOLAP scholar community.

We performed author name disambiguation based on the Scopus "Author(s) ID" field. Subsequently, we constructed the co-authorship network and used degree and betweenness centrality measures to identify and analyze the core nodes. In network analysis, degree centrality measures the extent of a node's direct connections with other nodes in the network.

The overview of research processes is shown in Figure 1.

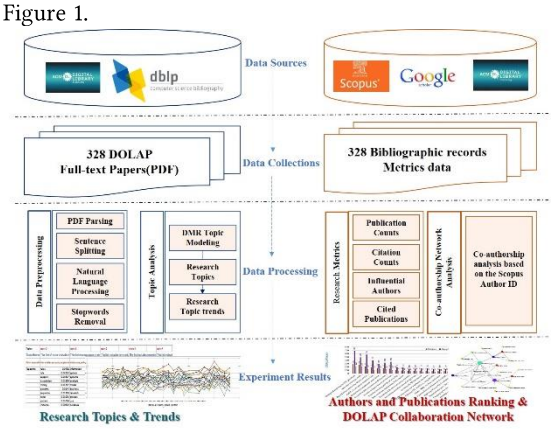


Figure 1: Research flow overview.

3. Results and Discussions

3.1. DOLAP research topics over 24 years

The experimental results identified 20 topics along with point mutual information (PMI) and 30-word distributions. Following the analysis of the topic results, we interpreted the PMI and 30-word distributions to label each topic, excluding non-meaningful wordings in terms of the DOLAP research area. In Table 1, we report the total 20 topics, encompassing Topic 0 to Topic 19. Each topic displays the label and only the top 5 word distributions. The topic labeling is used to demonstrate the theme of each topic.

3.2. Topic trends analysis

This section presents the analysis results of DOLAP research topics across three distinct time stages: 1998–2005, 2006–2015, and 2017–2023. Furthermore, the trends of topics over the twenty-five-year span (1998–2023) are identified as increasing, constant, and declining trends. These results are obtained through DMR topic modeling analysis.

3.2.1. DOLAP research topics across three distinct time stages

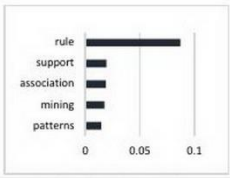
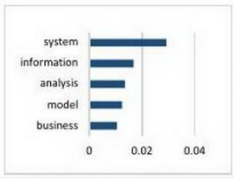
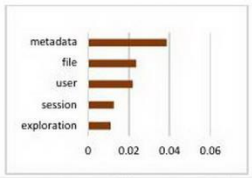
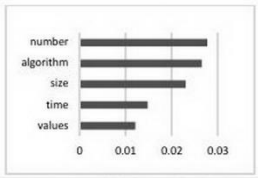
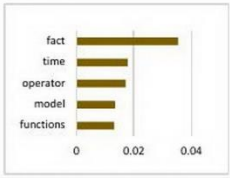
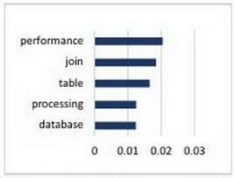
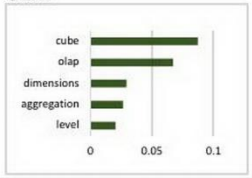
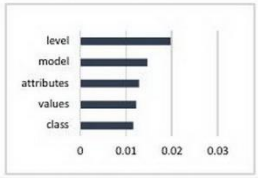
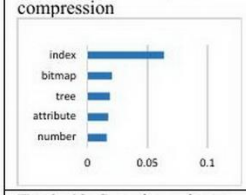
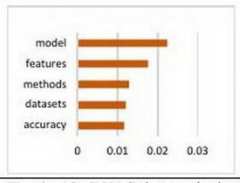
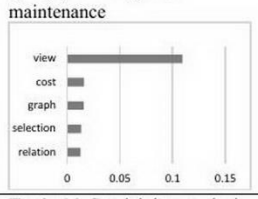
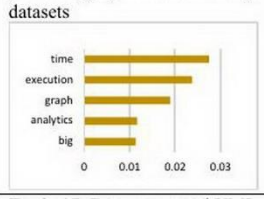
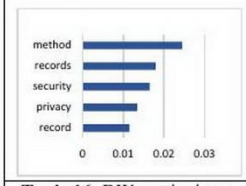
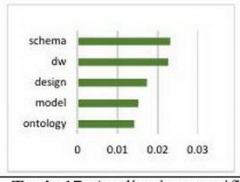
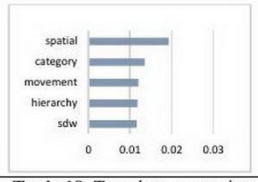
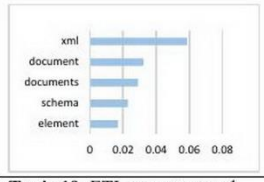
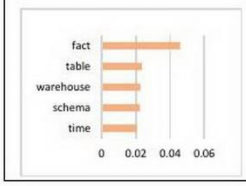
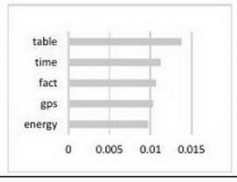
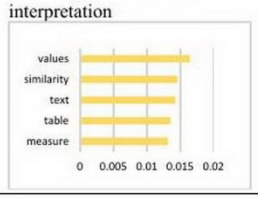
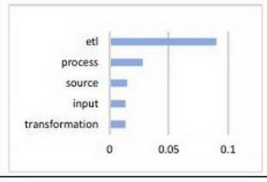
In Figure 2, we present the relative frequency rankings of each topic. Notably, Topics 11, 15, and 3 consistently captured the interest of researchers across three distinct time stages. As described in Table 1, these topics revolve around big data analytics and processing (Topic 11), document and XML data processing & management (Topic 15), and sparse data, algorithms & computational processes (Topic 3). However, it is worth noting that the top five research topics exhibit variations over the three distinct periods.

Between 1998 and 2005, in addition to Topics 11, 15, and 3, the two additional topics emerged with high frequencies are Topics 1 and 16. Topic 1 includes data warehouse requirements, business models and processes. Topic 16 covers data warehouse versioning and schema evolution. Two additional notable topics are Topics 6 and 17, which represent multidimensional data analysis with OLAP and application areas. This period represents the early stage of DOLAP research landscape, covering DW requirements, business processes, multidimensional data analysis & OLAP, schema design and evolution, and DW applications.

Between 2006 to 2015, in addition to Topics 11, 15, and 3, the two additional topics emerged with high frequencies are Topics 4 and 8. Topic 4 includes various proposals on algebra and aggregation. Topic 8 includes physical database techniques and performance optimization such as indexing and compression techniques. Two additional notable topics are again Topics 6 and 17, which represent multidimensional data analysis with OLAP and DW application areas. This period marked the rise of big data analytics, performance optimization, and document data processing.

Figure 2: DOLAP research topics across three distinct time stages.

Table 1
Overview of the DOLAP research topics

Topic Label & Top Terms	Topic Label & Top Terms	Topic Label & Top Terms	Topic Label & Top Terms
Topic 0: Data mining topics 	**Topic 1**: System requirements, business models and process 	**Topic 2**: Metadata management, workload, and quality 	**Topic 3**: Sparse data, algorithms and computational processes 
Topic 4: Algebra and aggregation 	**Topic 5**: Query optimization and performance 	**Topic 6**: Multidimensional data analysis with data cube and OLAP 	**Topic 7**: Multidimensional data modeling and design 
Topic 8: Indexing, physical data representation, and compression 	**Topic 9**: Machine learning topics 	**Topic 10**: Materialized Views: Query processing and maintenance 	**Topic 11**: Big data analytics and processing, optimization of large datasets 
Topic 12: Security, privacy, and encryption 	**Topic 13**: DW Schema design, integration, and ontology 	**Topic 14**: Spatial data analysis and GIS 	**Topic 15**: Documents and XML data management 
Topic 16: DW versioning and schema evolution 	**Topic 17**: Application-specific representation and analysis 	**Topic 18**: Text data processing, similarity, evaluation & interpretation 	**Topic 19**: ETL processes and tools 

Lastly, between 2017 to 2023, in addition to Topics 11, 15, and 3, the two additional topics emerged with high frequencies are Topics 18 and 1. Topic 18 is related to text processing, similarity measurement, classification, evaluation, and interpretation (Topic 18), while Topic 1 includes requirements, business models and processes. Two additional notable topics are Topics 9 and 6, which represent machine learning approach and multidimensional data analysis. This period indicates the popularity of text data processing and the rise of machine learning approach in the DOLAP research landscape.

3.2.2. Topic trends over the twenty-five-year span

Next, we used the *publication time* as a variable in our DMR analysis to identify the research trends across 20 topics. These trends were categorized into *increasing*, *constant*, and *declining trends*. We elaborate on each trend using the topics listed in Table 1. These trends

reflect the evolving landscape of research and exploration in the field. The results are presented as follows:

Increasing topic trends

Figure 3 depicts the upward trends of topics, showing the relative distribution ratio of each topic over the last 25 years from 1998 to 2023. The graph reveals seven increasing topic trends, with the predicted lines for each topic trend showing a subtle rise. Notably, we observed upward trends in topics related to algebra and aggregation (Topic 4), indexing and physical data representation techniques (Topic 8), machine learning approaches (Topic 9), security, privacy, and encryption (Topic 12), documents and XML data processing and management (Topic 15), application-specific topics (Topic 17), as well as text data processing, similarity measurement, and evaluation (Topic 18).

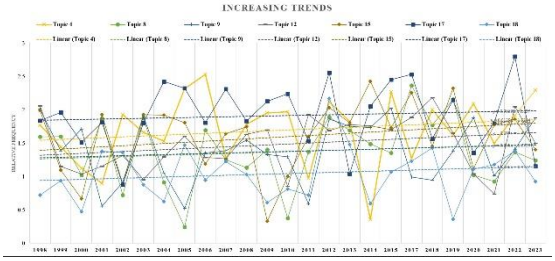


Figure 3: Increasing trends.

Constant topic trends

Figure 4 shows the prediction of constant trends, illustrating four themes that DOLAP researchers have consistently explored from 1998 to 2023. These enduring themes include the data mining techniques (Topic 0), sparse data, algorithms, and computational processes (Topic 3), multidimensional data analysis with data cubes and OLAP (Topic 6), and big data analytics & processing framework, and optimization of large datasets (Topic 11). These research areas have remained steadfast over the 25 years, contributing to the field's stability and growth.

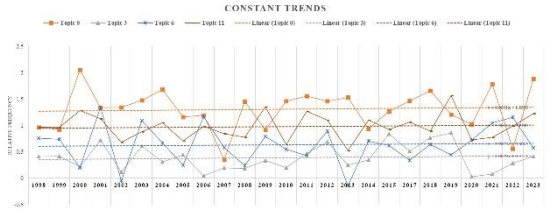


Figure 4: Constant trends.

Declining topic trends

Figure 5 presents prediction of declining trends, highlighting nine topics that exhibit a slight decrease in popularity. These topics include system requirements, business processes (Topic 1), metadata management, workload, and quality (Topic 2), query optimization and performance (Topic 5), and multidimensional data modeling and design framework (Topic 7). Furthermore, other declining topics include materialized view processing and maintenance (Topic 10), data warehousing schema design and ontology (Topic 13), spatial data analysis and GIS framework (Topic 14), data warehousing versioning, and schema evolution (Topic 16), and ETL processes and tools (Topic 19). These trends show that the research on schema design, query optimization, query processing with materialized views, versioning, and ETL are relatively well studied.

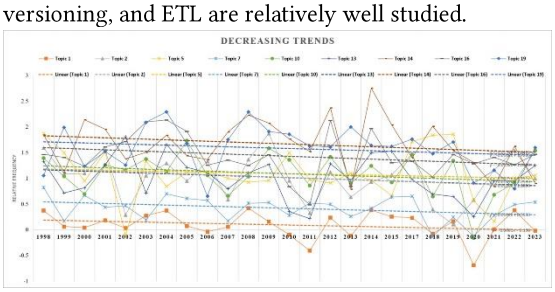


Figure 5: Declining trends.

3.3. DOLAP paper metrics

In this section, we delve into the performance of DOLAP authors and their publications. First, we describe the performance of DOLAP workshops using information from those academic databases. Second, we spotlight the productive DOLAP authors.

3.3.1. DOLAP paper performance analysis

We collected citation counts for DOLAP papers from Scopus and summed up the total citation counts for each DOLAP proceeding. In Figure 6, the results reveal that over 24 years, a total of 328 papers were published in DOLAP workshops, accumulating 5472 citations. The count encompasses regular papers, invited talk papers, short papers, and keynote papers. Noteworthy observations are that, in 2012, a total of 18 papers were published, which makes the largest number of papers, followed by the years 2007 and 2010, each with 16 papers.

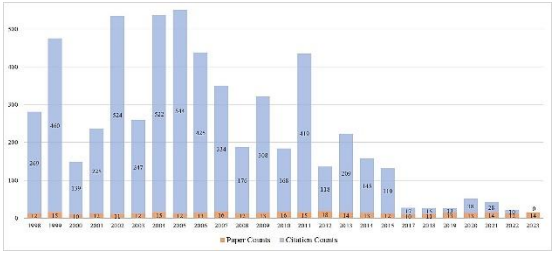


Figure 6: DOLAP papers and citation counts.

The most influential conference among the scholarly community was in 2005, with 544 citation counts, despite only 12 papers being published. This was

followed by the year 2002 with 524 citation counts and the year 2004 with 522 citations.

Next, we examined the citation records of each DOLAP paper by collecting data from three sources: Scopus, Google Scholar, and ACM. However, the ACM database only provided DOLAP paper data from 1998 to 2015.

Figure 7 shows the top 20 most cited papers, and Table 2 reveals information about the top 10 papers. It is important to note that, in this result, we ranked the papers by Scopus citation counts. The most dominant paper is titled "Conceptual modeling for ETL processes," by Vassiliadis P., Simitsis A., and Skiadopoulos S. [4], which received 281 citations by Scopus, 682 by Google Scholar, and 223 by ACM. Following this is the paper entitled "Analytics over large-scale multidimensional data: The big data revolution!" by Cuzzocrea A., Song I.-Y., and Davis K.C. [5] with 258, 490, and 204 by Scopus, Google Scholar, and ACM, respectively. The 3rd paper is entitled "Beyond data warehousing: What's next in business intelligence?" by Golfarelli M. and Rizzi S.; Cella I. [6], with citation counts of 207, 574, and 133 by the three sources, respectively.

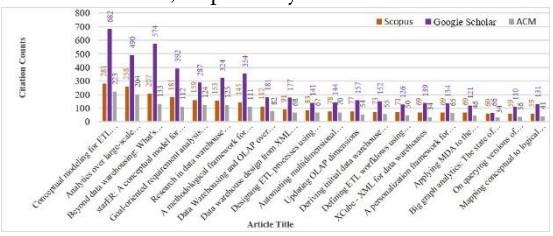


Figure 7: DOLAP papers and citation counts.

Table 2
Top 10 most cited DOLAP papers

Title/Authors	Citation Counts		
	Scopus	Google Scholar	ACM
Conceptual modeling for ETL processes / Vassiliadis P.; Simitsis A.; Skiadopoulos S. [4]	**281**	**682**	**223**
Analytics over large-scale multidimensional data: The big data revolution! / Cuzzocrea A.; Song I.-Y.; Davis K.C. [5]	**258**	**490**	**204**
Beyond data warehousing: What's next in business intelligence? / Golfarelli M.; Rizzi S.; Cella I. [6]	**207**	**574**	**133**
starER: A conceptual model for data warehouse design / Tryfona N.; Busborg F.; Christiansen J.G.B. [7]	**181**	**392**	**112**
Goal-oriented requirement analysis for data warehouse design / Giorgini P.; Rizzi S.; Garzetti M. [8]	**159**	**287**	**124**
Research in data warehouse modeling and design: Dead or alive? / Rizzi S.; Abelló A.; Lechtenbörger J.; Trujillo J. [9]	**153**	**324**	**125**
A methodological framework for data warehouse design / Golfarelli M.; Rizzi S. [10]	**145**	**354**	**111**
Data Warehousing and OLAP over Big Data: Current challenges and future research directions / Cuzzocrea A.; Bellatreche L.; Song I.-Y. [11]	**112**	**181**	**82**
Data warehouse design from XML sources Golfarelli M.; Rizzi S.; Vrdoljak B. [12]	**91**	**177**	**68**
Designing ETL processes using semantic web technologies / Skoutas D.; Simitsis A. [13]	**83**	**141**	**67**

3.3.2. Productive DOLAP authors

While analyzing the Scopus dataset, we discovered that Stefano Rizzi was the most productive among DOLAP authors who published papers in the DOLAP workshops, contributing a total of 23 papers over 24 years. His research areas cover business intelligence and data warehouse modeling and design. This is followed by Patrick Marcel, Alberto Abelló, Torben Bach Pedersen, Panos Vassiliadis, and Matteo Golfarelli, as shown in Table 3.

Table 3
Top 10 productive DOLAP researchers

Name	Affiliation	DOLAP Papers
Rizzi, Stefano	University of Bologna, Italy	23
Marcel, Patrick	Université F. Rabelais 3, France	19
Abelló, Alberto	U. Politecnica de Catalunya, Spain	17
Pedersen, Torben Bach	Aalborg University, Denmark	16
Vassiliadis, Panos	National Technical Univ. of Athens, Greece	15
Golfarelli, Matteo	University of Bologna, Italy	14
Ordonez, Carlos	University of Houston, United States	14
Trujillo, Juan	Universidad de Alicante, Spain	12
Bellatreche, Ladjel	LISI ENSMA Téléport, France	10
Song, Il-Yeol	Drexel University, United States	10

3.4. Co-authorship network

We used bibliographic information from Scopus to construct the co-authorship networks. The process included name disambiguation, which relies on the "Author(s) ID" field to accurately identify and distinguish authors with the same name. For network analysis, we utilized degree and betweenness centrality measures to characterize the network's structure and identify key researchers.

The DOLAP co-authorship network analysis uncovered a network comprising 589 nodes (authors) and 1032 edges (co-authorship links). It consists of 93 disconnected sub-networks. The largest subnetwork, depicted in Figure 8, contains 237 nodes, accounting for 40.24% of the total. Figure 9 (9-A to 9-D) shows four sub-networks ranked 2nd to 5th in terms of size. Within each sub-network, nodes are color-coded based on their degree centrality. Thicker edges represent closer collaborations. To enhance clarity, the graphs display only nodes with significant degree centrality (influential authors).

Within the dominant network depicted in Figure 8, several authors exhibited extensive connections. They are Rizzi, Stefano, Marcel, Patrick; Vassiliadis, Panos; Peralta, Verónika; and Labroche, Nicolas. Their numerous collaborations highlight their prominence within the network. Notably, Rizzi, Stefano shared a particularly strong partnership with Golfarelli, Matteo. Figure 9 reveals additional hubs within individual sub-networks: Abelló, Alberto; Trujillo, Juan; Ordonez, Carlos; and Lehner, Wolfgang emerged as central figures in their respective clusters, demonstrating high connectivity and influence.

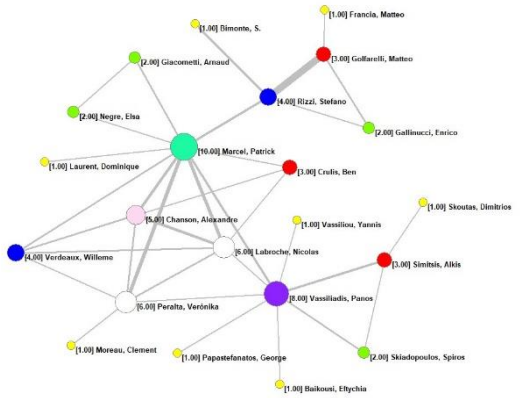


Figure 8: Giant sub-network of collaborations.

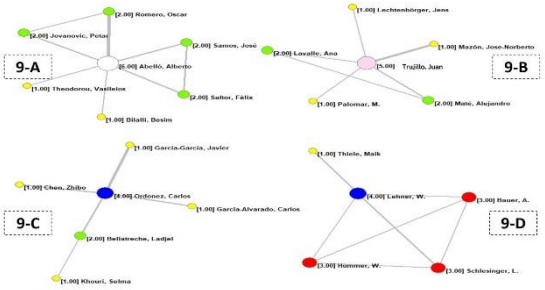


Figure 9: Sub-networks of collaborations (9A-9D).

Table 4 presents the top 10 most connected authors within the entire network, ranked by degree centrality. This metric reveals the number of direct connections a node (author) has with others in the co-authorship network. Noteworthy observations include: Rizzi, Stefano emerged as the most connected author, exhibiting a degree centrality of 39 connections; Marcel, Patrick closely follows with 37; Pedersen, Torben Bach holds the third place with 25. The table also includes betweenness centrality values for each top-ranked author. This measure assesses an individual's importance as a bridge or intermediary within the network, indicating their potential influence over information flow.

Table 4
Top 10 most connected authors in the DOLAP workshops

No	Author	Degree	Betweenness Centralization
1	Rizzi, Stefano	39	0.05604
2	Marcel, Patrick	37	0.04206
3	Pedersen, Torben Bach	25	0.02842
4	Abelló, Alberto	24	0.04067
5	Vassiliadis, Panos	24	0.03298
6	Peralta, Verónika	22	0.00650
7	Bellatreche, Ladjel	22	0.02868
8	Trujillo, Juan	18	0.03238
9	Labroche, Nicolas	18	0.00605
10	Song, Il-Yeol	17	0.02642

4. Conclusion

In this study, we delved into the realm of DOLAP research, spanning a quarter-century from 1998 to 2023. Our investigation encompassed 328 full-text DOLAP papers, drawing insights from diverse sources including DBLP, Scopus, Google Scholar, and the ACM Digital Library.

We examined evolution of research themes, key research performance metrics, and the co-authorship networks of the DOLAP scholarly community. DMR topic modeling and social network analysis were employed to explore research topics and collaboration patterns, respectively.

The key findings reveal a dynamic landscape of persistent research interests, upward trends, and declining interests. Over the past 25 years of the DOLAP workshops, the consistent research topics encompass topics related to data mining, various algorithmic approaches, big data analytics, and multidimensional analysis. Topics that exhibited upward trends include areas related to text data processing, indexing and physical data representation, machine learning, and security. Topics with declining interests include requirement analysis, business processes, data warehousing design frameworks, schema evolution, view processing, ETL, spatial data analysis, and GIS frameworks.

DOLAP 2005 marked the most impactful DOLAP workshop, gathering a remarkable 544 citations. This year stood out as a peak of influence within the DOLAP history. Notably, the paper entitled "Conceptual modeling for ETL processes" by Vassiliadis et al. [4] marked as the most cited paper, racking up 281 citations on Scopus, significantly influencing the field. Among all DOLAP authors, Stefano Rizzi stood out with a remarkable 23 papers published in DOLAPs. The top 10 most productive authors and the top 10 most influential authors were also identified.

This study's outcome can contribute to a more comprehensive understanding of the evolution of research landscape and collaborative dynamics within the DOLAP field. These findings will provide valuable insights for researchers, practitioners, and future DOLAP organizers for continued growth and impact.

Acknowledgements

This work was partially supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2022R1A2B5B02002359).
Supporting data can be found in GitHub Repository at https://github.com/tatsawan/DOLAP-paper-analysis

References

[1] D. M. Blei, A. Y. Ng, M. I. Jordan, Latent Dirichlet Allocation, Journal of Machine Learning Research, 3(2003) 993-1022.

[2] D. M. Mimno, A. McCallum, Topic models conditioned on arbitrary features with Dirichlet-multinomial regression, UAI 24 (2008) 411-418.

[3] P. D. Turney, Mining the web for synonyms: PMI-IR versus LSA on TOEFL, in: European conference on machine learning, Berlin, Heidelberg, Springer Berlin Heidelberg, 2001, pp. 491-502.

[4] P. Vassiliadis, A. Simitsis, S. Skiadopoulos, Conceptual modeling for ETL processes, in: Proceedings of the 5th ACM International Workshop on Data Warehousing and OLAP, 2002, pp. 14-21.

[5] A. Cuzzocrea, I. Y. Song, K. C. Davis, Analytics over large-scale multidimensional data: the big data revolution!, in: Proceedings of the ACM 14th International Workshop on Data Warehousing and OLAP, 2011, pp. 101-104.

[6] M. Golfarelli, S. Rizzi, I. Cella, Beyond data warehousing: what's next in business intelligence?, in: Proceedings of the 7th ACM International Workshop on Data Warehousing and OLAP, 2004, pp. 1-6.

[7] N. Tryfona, F. Busborg, J.G. Borch Christiansen, starER: A conceptual model for data warehouse design, in: Proceedings of the 2nd ACM International Workshop on Data Warehousing and OLAP, 1999, pp. 3-8.

[8] P. Giorgini, S. Rizzi, M. Garzetti, Goal-oriented requirement analysis for data warehouse design, in: Proceedings of the 8th ACM International Workshop on Data Warehousing and OLAP, 2005, pp. 47-56.

[9] S. Rizzi, A. Abelló, J. Lechtenbörger, J. Trujillo, Research in data warehouse modeling and design: dead or alive?, in: Proceedings of the 9th ACM International Workshop on Data Warehousing and OLAP, 2006, pp. 3-10.

[10] M. Golfarelli, S. Rizzi, A methodological framework for data warehouse design, in: Proceedings of the 1st ACM International Workshop on Data Warehousing and OLAP, 1998, pp. 3-9.

[11] A. Cuzzocrea, L. Bellatreche, I. Y. Song, Data warehousing and OLAP over big data: current challenges and future research directions, in: Proceedings of the Sixteenth International Workshop on Data Warehousing and OLAP, 2013, pp. 67-70.

[12] M. Golfarelli, S. Rizzi, B. Vrdoljak, Data warehouse design from XML sources, in: Proceedings of the 4th ACM International Workshop on Data Warehousing and OLAP, 2001, pp. 40-47.

[13] D. Skoutas, A. Simitsis, Designing ETL processes using semantic web technologies, in: Proceedings of the 9th ACM International Workshop on Data Warehousing and OLAP, 2006, pp. 67-74.