# CO2e Emissions Analysis in Life Cycle Inventory Using Knowledge Graph Modeling

Laurensia Simanihuruk
*Department of Informatics*
*Institut Teknologi Sepuluh Nopember*
Surabaya, Indonesia
6025241050@student.its.ac.id

Kelly Rossa Sungkono
*Department of Informatics*
*Institut Teknologi Sepuluh Nopember*
Surabaya, Indonesia
kelly@its.ac.id

Riyanarto Sarno
*Department of Informatics*
*Institut Teknologi Sepuluh Nopember*
Surabaya, Indonesia
riyanarto@its.ac.id

Abdullah Faqih Septiyanto
*Department of Informatics*
*Institut Teknologi Sepuluh Nopember*
Surabaya, Indonesia
7025221022@student.its.ac.id

Rizqy Ahsana Putri
*Department of Informatics*
*Institut Teknologi Sepuluh Nopember*
Surabaya, Indonesia
7025242013@student.its.ac.id

Aziz Fajar
*Faculty of Advanced Technology and Multidiscipline Universitas Airlangga*
Surabaya, Indonesia
azizfajar@ftmm.unair.ac.id

Siti Halimah Larekeng
*Faculty of Forestry, Universitas Hasanuddin*
Makassar, Indonesia
sittihalimah@unhas.ac.id

Muhammad Suzuri Hitam
*Faculty of Computer Science and Mathematics*
*Univeristi Malaysia Terengganu*
Terengganu, Malaysia
suzuri@umt.edu.my

***Abstract*—Life Cycle Inventory (LCI) is widely used to evaluate the environmental impacts of products. The complex structure of LCI data often creates challenges for effective analysis, especially in the context of climate change mitigation. This study presents a knowledge graph-based approach to improve the structural representation and streamline the analysis of LCI data. Greenhouse gas emissions are converted to carbon dioxide equivalents ($CO_2e$) using the GWP100 metric, and the data are then categorized according to $CO_2e$ quartiles. The constructed knowledge graph maps relationships and properties among entities and enables efficient querying. The analysis identifies three emission categories: high (average 185.95 kg $CO_2e$), medium (0.58 kg $CO_2e$), and low (0.03 kg $CO_2e$). These results offer a structured foundation for pinpointing high-impact processes and prioritizing emission reduction strategies in sectors with significant environmental burdens, leveraging the strengths of graph-based modeling and structured data representation.**

***Keywords—$CO_2e$, GWP100, knowledge graph, life cycle inventory***

## I. Introduction

Life Cycle Assessment (LCA) is an analytical approach used to evaluate the environmental impacts of a product or process throughout its entire life span, from raw material extraction and production to use and disposal [1]. The Life Cycle Inventory (LCI) phase provides the input and output data for each stage, including material flows, energy consumption, and emissions. The quality of LCI data strongly influences the accuracy of the final results of an LCA study.

The management of LCI data poses a major challenge as industrial processes involve complex and interconnected material and energy flows. Relational databases with fixed tabular structures cannot represent these relationships effectively, limiting data integration and analytical depth [2]. Similar integration challenges occur in data-intensive systems such as microservice architectures, where rule-based extraction is needed to manage interrelated components efficiently [3]. This underscores the need for flexible and structured modeling techniques to handle complex datasets. Traditional LCI approaches often aggregate emissions across life cycle phases, obscuring hotspots and hindering the identification of high-impact processes. Such aggregation masks temporal and process-specific variations, leading to missed mitigation opportunities and less informed decisions [4]. Similar challenges in bioinformatics illustrate the need for structured analytical frameworks to improve precision and connectivity in complex LCI datasets [5].

A graph-based data model provides an alternative paradigm for representing life cycle systems in a more flexible and interconnected manner. Similar graph-based approaches have been applied in business process analysis to capture relationships and dependencies between entities effectively [6], demonstrating the capability of graph structures to represent complex system interactions. In this configuration, processes are represented as nodes and material or energy flows as relationships [7], [8], [9]. This structure enables tracing the distribution of resources among processes and analysing the interactions within industrial systems. The model enhances structural representation by embedding contextual meaning in each connection, making it a promising analytical approach for investigating greenhouse gas emissions in complex production environments.

Greenhouse gas (GHG) emissions constitute a principal contributor to global climate change. In 2018, global emissions were estimated at approximately 58 gigatonnes of $CO_2e$. The industrial sector accounted for roughly 24% of direct emissions and up to 35% when indirect energy use was included [10]. A separate study involving 866 product carbon footprints across 30 industries demonstrated that, on average, a single product generates emissions equivalent to 6.3 times its own weight in $CO_2e$ over its life cycle [11]. These empirical observations underscore the necessity of analytical frameworks that can identify and prioritize processes with the greatest emission intensities.

This study introduces a graph-based framework for greenhouse gas impact assessment within the LCA context. The proposed framework integrates knowledge graph modeling with carbon equivalence standardization utilizing the Global Warming Potential (GWP100) metric derived from the Intergovernmental Panel on Climate Change (IPCC) Sixth Assessment Report (AR6). All greenhouse gas emissions are converted into $CO_2e$, allowing for consistent comparison across processes. Each process is then classified into three emission categories: high, medium, and low. Graph queries in Neo4j are applied to identify emission patterns, reveal structural dependencies, and improve transparency within the LCA dataset.

The proposed approach enhances the structural representation of life cycle data and improves the understanding of emission distribution. This framework enables the prioritization of processes with substantial environmental burdens to support the development of adaptive and data-driven environmental assessment systems.

## II. Literature Review

This section presents a review of relevant studies and theoretical foundations related to Life Cycle Assessment, knowledge graph modeling, and greenhouse gas impact evaluation.

### *A. Life Cycle Inventory*

The Life Cycle Inventory (LCI) represents a crucial phase in the Life Cycle Assessment methodology, focusing on the collection and quantification of all resource inflows and outflows, along with the environmental emissions linked to a product system across its entire life cycle. ISO 14044 defines that the LCI stage encompasses the processes of gathering data and performing input-output calculations, which include energy consumption, water usage, raw material utilization, and emissions released into the air, water, and soil [12].

The accuracy and completeness of data at the LCI stage are crucial in determining the reliability of environmental impact assessment results in the subsequent stages, as any gaps or errors can lead to significant deviations in the final outcomes. Therefore, various databases such as Ecoinvent and USLCI have been developed to provide standardized, transparent, and traceable LCI data across various industrial sectors, ensuring consistency and comparability in LCA studies. These databases are continuously updated to reflect technological advancements and regional variations, further enhancing their applicability in global sustainability assessments.

### *B. Knowledge Graph*

A knowledge graph is a structured representation of data that captures relationships between entities, enabling flexible querying and contextual analysis. This approach has been widely used to address the challenges of data integration and analysis on a large scale because it can unify various dispersed information sources into a coherent and easily navigable data structure [13].

A knowledge graph consists of three core components: nodes, relationships, and properties [14]. In this study, nodes represent entities such as processes, flows, locations, and exchanges, while relationships connections (e.g., USES_FLOW, EMITS_ELEMENTARY_FLOW). Both nodes and relationships include key-value properties that store relevant attributes. This structure enables flexible querying and analysis, supporting deeper insights into process dependencies and emission pathways within the LCI system [15].

### *C. Global Warming Potential and $CO_2$e Standardization*

Global Warming Potential is an indicator for assessing the effects of a certain greenhouse gas on global warming from the effects of carbon dioxide ($CO_2$). This evaluation is conducted by the IPCC, with the values obtained being standardised over designated temporal periods. This metric serves to evaluate the cumulative radiative effects of various greenhouse gases over specified timeframes, typically categorized into three distinct periods, 20 years (GWP20), 100 years (GWP100), and 500 years (GWP500) [16].

TABLE I. Example of $CO_2$e Standardization Using GWP100

| Gas | Chemical Formula | GWP100 |
|---|---|---|
| Carbon Dioxide | $CO_2$ | 1 |
| Methane | $CH_4$ | 27 |
| Nitrous Oxide | $N_2O$ | 273 |
| Nitrogen Trifluoride | $NF_3$ | 17400 |

In this study, GWP100 is used because it is the most widely adopted standard in international environmental policies such as the Kyoto Protocol and the United Nations Framework Convention on Climate Change (UNFCCC) reporting framework [17]. GWP100 offers a balance between short-term sensitivity and long-term relevance, and is used consistently in national emissions inventories, LCA databases, and environmental software. The use of GWP100 also provides a comparable basis for categorizing different types of emission streams in the context of life cycle evaluation, making it easier to classify industrial processes based on their emissions intensity [17].

The standardization of $CO_2$ equivalents, as presented in TABLE I, was carried out by converting all greenhouse gas (GHG) emissions into $CO_2$e units using the GWP100 values from the IPCC AR6, which are officially available through the U.S. Government Data Portal [16]. AR6 was chosen as it represents the most recent advancements in climate science, offering updated and relevant GWP values.

## III. Methodology

The methodology of this study, as illustrated in Fig. 1, consists of five sequential stages: data collection, data preprocessing, knowledge graph construction in Neo4j, graph query analysis, and $CO_2$e categorization. Fig. 2 presents the graph schema that defines the underlying structure used to represent interconnected entities and relationships within the LCI dataset.

### *A. Data Collection*

The first stage of this study involved extracting data from the USLCI database, which is provided in JSON format, with each of the 861-unit processes stored as a separate file identified by a unique ID [17]. These files include detailed information about material and energy flows, emissions, and relevant metadata for each industrial activity.

This integration not only simplifies data management but also allows explicit representation of relationships between entities, forming the basis for constructing a knowledge graph using a node and relationship-based approach and metadata into a consistent format. The unified dataset ensures that each process, flow, and emission record can be systematically linked and analyzed in the subsequent stages of graph modeling.

### *B. Preprocessing Data*

Prior to graph construction, the consolidated dataset underwent preprocessing to ensure data consistency and compatibility for graph-based modeling. This step included standardizing flow names, units of measurement, and process identifiers, as well as removing incomplete or duplicate entries. Metadata fields were aligned to maintain uniform structure across all records.

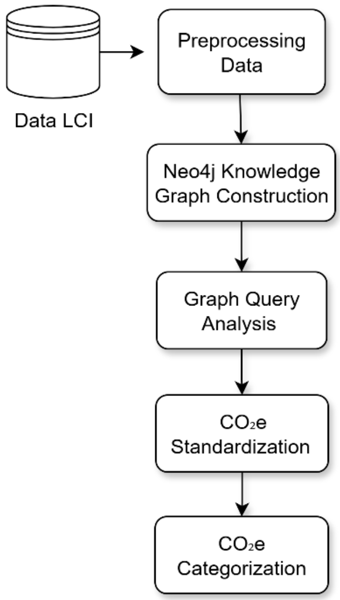


Fig 1. Workflow Methodology

The cleaned dataset was then converted into a property graph-ready format, where each industrial process, material flow, and emission factor was assigned as a distinct entity. Numerical attributes such as flow quantity and measurement units were retained to preserve quantitative accuracy. This preprocessing phase ensured that the data was fully structured and ready for integration into the Neo4j graph database.

### *C. Neo4j Knowledge Graph Construction*

The knowledge graph construction process in this study consisted of three main stages: knowledge extraction, knowledge integration, and knowledge application.

In the knowledge extraction stage, the consolidated dataset was analyzed to identify key entities and their relationships. Each entity, such as processes, flows, or locations, was represented as a node, while their interactions, such as USES_FLOW, EMITS_FLOW, or HAS_LOCATION, were represented as relationships. This mapping produced a set of triples (entity, relation, entity) that formed the structural foundation of the graph.

In the knowledge integration stage, the extracted information was implemented into a Neo4j graph database using the labeled property graph model. Nodes and edges were constructed to preserve both structural and topological coherence, enabling accurate representation of interconnections among processes within the production system.

In the knowledge application stage, the resulting knowledge graph was used to perform graph-based querying and $CO_2e$ emission analysis. These queries supported efficient navigation through the interconnected data, allowing the identification of emission pathways and the relationships between processes, flows, and their environmental impacts.

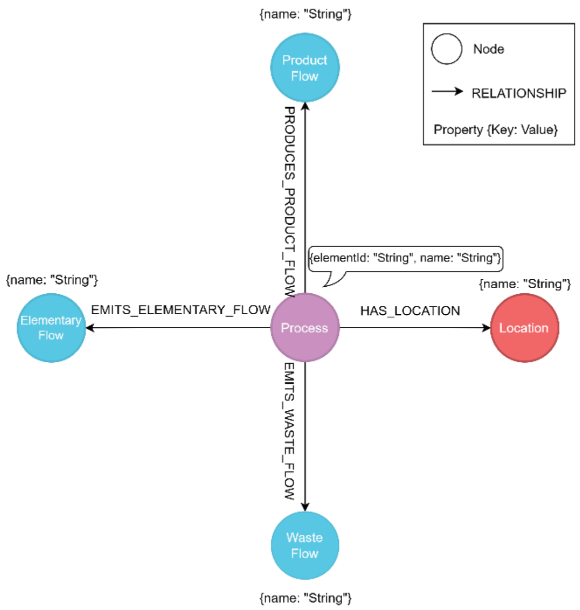


Fig 3. Graph Schema for LCI Data Modeling

Fig. 3 illustrates the overall structure of the developed knowledge graph.

### *D. Graph Query Analysis*

Graph query analysis was applied to retrieve and examine information based on the relational structure of entities within the knowledge graph. This procedure enabled efficient exploration of the interconnections among processes, material and energy flows, and emissions in the LCI system, allowing for systematic assessment of data relationships relevant to environmental performance.

In this stage, queries were executed in Neo4j using the Cypher query language to analyze emission relationships and trace linkages between processes and associated flows. The queries were designed to identify processes that produce or use specific flow types and to map emission pathways across the production network.

### *E. $CO_2e$ Standardization*

All greenhouse gas emissions in the dataset were standardized to $CO_2e$ to ensure comparability across different gas types. The conversion was performed using the 100-year GWP100 factors provided in the IPCC AR6. Each emission record in the dataset, including gases such as $CH_4$ and $N_2O$, was converted to $CO_2e$ using its respective GWP100 value.

The conversion process was conducted by multiplying the emission quantity of each gas, expressed in kilograms, by its corresponding GWP100 factor. For instance, 1 kg of $CH_4$ equals 27 kg $CO_2e$, and 1 kg of $N_2O$ equals 273 kg $CO_2e$ according to AR6. The resulting $CO_2e$ values were then assigned as attributes within the knowledge graph, allowing direct comparison of emissions between processes.

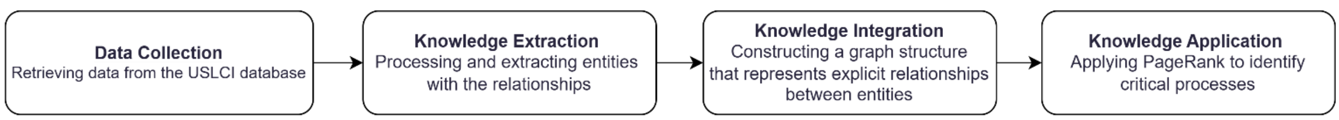


Fig 2. Process of Knowledge Graph Construction

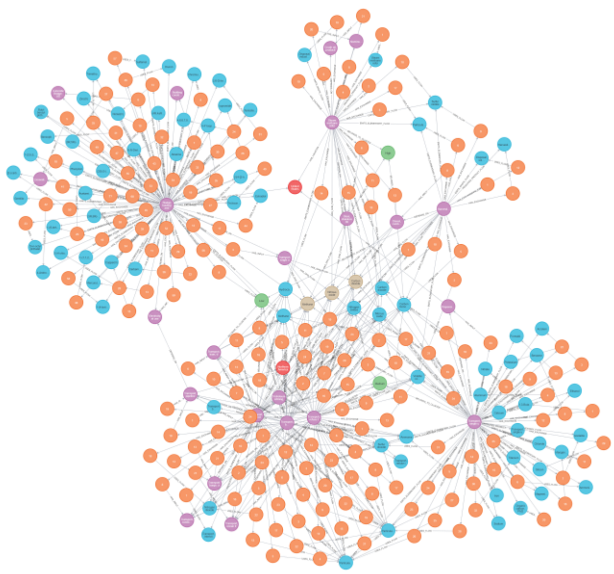

Fig 4. Visualization of Completed LCI Knowledge Graph

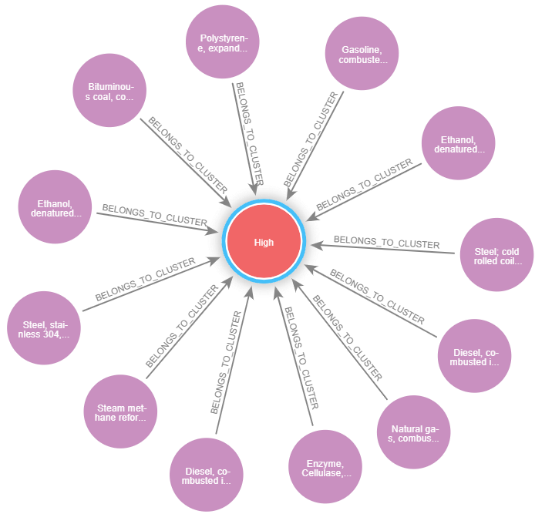


Fig 7. High $CO_2$e emission category

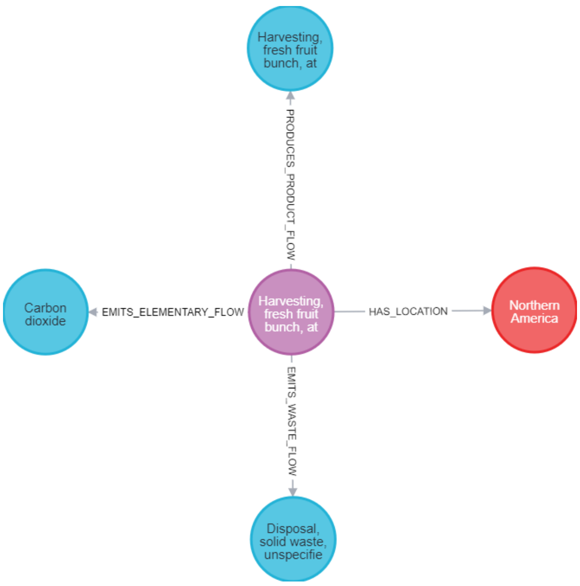


Fig 5. Visualization of a single process node and its associated flows and location in the LCI graph

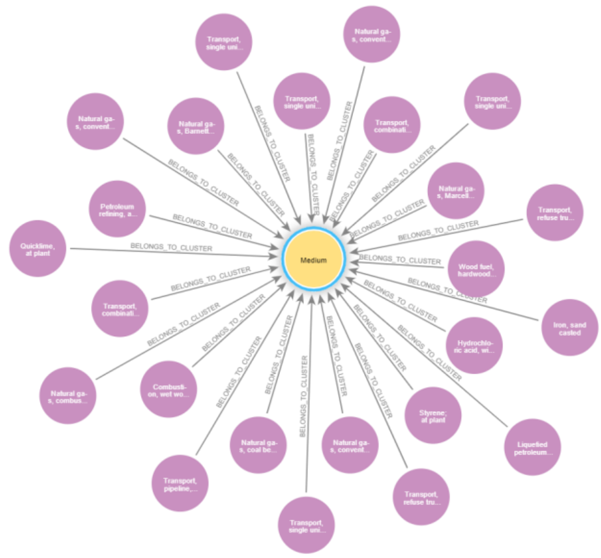


Fig 8. Medium $CO_2$e emission category


Fig 6. Visualization of processes emitting Carbon Dioxide


Fig 9. Low $CO_2$e emission category

This standardization ensured that all emission data was expressed in a unified metric, providing a consistent foundation for emission analysis and visualization within the LCI system.

### F. $CO_2e$ Categorization

To facilitate comparative analysis of emission intensity, each process was grouped into three $CO_2e$ categories: high, medium, and low, based on the statistical distribution of total $CO_2e$ values. This descriptive categorization was derived from the quartile ranges of the dataset.

Processes with $CO_2e$ values above the third quartile were classified as high emission, those between the first and third quartiles as medium emission, and those below the first quartile as low emission. This grouping provides a straightforward and interpretable means of distinguishing emission levels among processes without applying any clustering or machine learning algorithms.

## IV. Results and Discussion

This section presents the results obtained from each stage of the methodology, including knowledge graph construction, graph query analysis, $CO_2e$ emission standardization, and process categorization based on emission intensity.

### A. Result

As illustrated in Fig. 4, the LCI knowledge graph was modeled in Neo4j, where each node represents a key entity such as a process, flow, location, or exchange. The relationships between these entities, summarized in TABLE II, establish structured and consistent connections across the dataset. Fig. 5 visualizes the process Harvesting fresh fruit bunch, at plantation (ID: 0375c701-8d66-3b47-af2e-26b0c43c722f), showing its links to location, product flow, waste flow, and carbon dioxide emissions. This structure enables tracing both physical outputs and environmental impacts, demonstrating how the graph framework unifies diverse data types for detailed analysis of material and emission relationships.

Fig. 6 highlights the network of processes connected to carbon dioxide as an elementary flow. The central emission node links to multiple processes, revealing their $CO_2$ contributions and interactions through input and output exchanges. This structure helps identify emission hotspots and key processes that can be prioritized for reduction strategies. TABLE III shows the Cypher query used for quartile-based categorization of industrial processes according to their total $CO_2e$ emissions. The query calculates the first (Q1) and third (Q3) quartiles and assigns each process to a low, medium, or high emission category, storing the result as a label connected to the process node.

TABLE II. Relationship Types Used in the LCI Knowledge Graph

| Relationship Type | Description |
|---|---|
| HAS_LOCATION | Links a process to its geographic location |
| HAS_EXCHANGE | Connects a process with its associated exchanges |
| PRODUCES_PRODUCT_FLOW | Indicates the output of a process in the form of a product flow |
| EMITS_ELEMENTARY_FLOW | Represents the emission of elementary flows such as $CO_2$ |
| EMITS_WASTE_FLOW | Represents the emission of waste flows from a process |
| DEPENDS_ON | Indicates process dependency on other flows or processes |
| USES_FLOW | Denotes the input material or energy flow used by a process |

In Fig. 7, the high emission category of industrial processes is visualized based on the results of quartile-based classification. The central node, labeled as "High," corresponds to 12 processes in which total $CO_2e$ emissions exceed the third quartile threshold, including energy-intensive activities such as fossil fuel combustion, steel production, and synthetic material manufacturing. Statistical analysis shows that this category has an average emission of 185.95 kg of $CO_2e$ and a median of 3.41 kg, with a range from 1.97 to over two million kg. This concentration of high-emission processes highlights the disproportionate impact of a few activities and the importance of targeted intervention for maximum mitigation efficiency.

Fig. 8 presents the medium emission category, comprising 24 processes mainly related to natural gas combustion, transport, and logistics, with an average of 0.58 kg $CO_2e$ and a median of 0.17 kg. These represent processes with optimisation potential that could yield measurable reductions.

Fig. 9 shows the low emission category, including 12 processes dominated by transport and waste treatment, averaging 0.03 kg $CO_2e$ and a median of 0.027 kg. Although their individual impact is minor, these processes collectively help maintain system efficiency and support balanced emission management.

### B. Discussion

The knowledge graph approach successfully represents the relationships among processes, flows, and emissions within the LCI in a structured manner. This model enhances analytical capability by visualizing interconnections that are often difficult to identify in conventional tabular databases [7], [8]. The standardization of greenhouse gas emissions into $CO_2e$ using the GWP100 factors from the IPCC enables consistent comparison across different gas types and provides a quantitative basis for assessing global warming potential [8].

TABLE III. Cypher Query for Clustering Processes Based On $CO_2e$

| | |
|---|---|
| 1 | **CALL** {<br>**MATCH** (p:Process)<br>**WHERE** p.total_co2e **IS NOT NULL**<br>**RETURN**<br>percentileCont(p.total_co2e, 0.25) **AS** q1,<br>percentileCont(p.total_co2e, 0.75) **AS** q3<br>} |
| 2 | **WITH** q1, q3<br>**MATCH** (p:Process)<br>**WHERE** p.total_co2e **IS NOT NULL**<br>**WITH** p, q1, q3,<br>**CASE**<br>**WHEN** p.total_co2e <= q1 **THEN** "Low"<br>**WHEN** p.total_co2e <= q3 **THEN** "Medium"<br>**ELSE** "High"<br>**END** AS cluster<br>**MERGE** (c:CO2Cluster {level: cluster})<br>**MERGE** (p)-[:**BELONGS_TO_CLUSTER**]->(c)<br>**SET** p.clusterLevel = cluster |

The categorization results show that a small number of high-emission processes contribute most of the total $CO_2e$, consistent with the industrial emission distribution reported by Xie et al. [18] and Liang et al. [19]. This finding highlights the importance of focusing mitigation strategies on high-emission processes to achieve substantial reductions in total emissions. The integration of knowledge graph modeling and $CO_2e$ standardization improves traceability, consistency, and interpretability in LCI analysis. The developed framework not only visualizes complex interconnections among processes and emissions but also provides a quantitative foundation for emission mitigation planning and sustainable industrial development in the future.

## V. CONCLUSION

The complexity of LCI data makes it difficult to identify and prioritise emission-intensive industrial processes. Traditional methods often lack structural detail and analytical flexibility, reducing the effectiveness of LCI-based approaches for supporting climate change mitigation. This study demonstrates that combining knowledge graph modelling with GWP100-based $CO_2e$ standardisation and quartile-based categorisation effectively represents the relationships among entities within an LCI system.

The approach groups industrial processes by emission intensity, where high-, medium-, and low-emission categories average 185.95, 0.58, and 0.03 kg $CO_2e$, respectively. This categorisation provides a structured foundation for prioritising high-impact processes and guiding targeted mitigation strategies. Future research could incorporate machine-learning techniques to automate decarbonisation recommendations and support adaptive, data-driven decision-making in sustainable industrial systems.


## ACKNOWLEDGMENT

This work was funded by the Indonesian Endowment Fund for Education (LPDP) on behalf of the Indonesian Ministry of Higher Education, Science and Technology and managed under the EQUITY Program (Contract No 4299/B3/DT.03.08/2025 & No 3029/PKS/ITS/2025); the Directorate General of Higher Education, Research and Technology of the Republic of Indonesia under Penelitian Sinergi Hilirisasi Inovasi Komersial; Institut Teknologi Sepuluh Nopember (ITS) under Riset Kolaborasi Indonesia, Penelitian Kemitraan A, and Penelitian Departemen; and Ethereum Foundation under Academic Grants Round 2025.


## REFERENCES


[1] X. Chen, P. Shi, and H. Wang, "Life cycle environmental impact assessment of bridges based on comparison of different construction schemes," *Environment Development Sustainability*, 2025, doi: 10.1007/s10668-024-05895-6.

[2] P. Kotiranta, M. Junkkari, and J. Nummenmaa, "Performance of Graph and Relational Databases in Complex Queries," *Applied Sciences (Switzerland)*, vol. 12, no. 13, Jul. 2022, doi: 10.3390/app12136490.

[3] R. A. Putri, M. Nevin, D. B. Ansori, A. F. Septiyanto, and R. Sarno, "Automated Extraction of Microservice Architecture Using a Rule-Based Approach," in *Proceedings - 2024 2nd International Conference on Technology Innovation and Its Applications, ICTIIA 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/ICTIIA61827.2024.10761417.

[4] S. Shoaib-ul-Hasan, M. Roci, F. M. A. Asif, N. Salehi, and A. Rashid, "Analyzing temporal variability in inventory data for life cycle assessment: Implications in the context of circular economy," *Sustainability*, vol. 13, no. 1, pp. 1–12, 2021, doi: 10.3390/su13010344.

[5] A. F. Septiyanto, Y. F. C. Cahyono, R. Sarno, F. Taufany, S. Larekeng, K. R. Sungkono, K. Syahruddin, E. G. Lestari, Sholiq, "An Improved Method for Prioritizing Polymerase Chain Reaction (PCR) Primer Design in Sanger Sequencing," in *Proceeding - 6th International Conference on Information Technology, Information Systems and Electrical Engineering: Applying Data Sciences and Artificial Intelligence Technologies for Environmental Sustainability, ICITISEE 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 443–448. doi: 10.1109/ICITISEE57756.2022.10057821.

[6] K. R. Sungkono, R. Sarno, M. C. Salsabila, and C. P. Dewi, "A Graph-based Method for Merging Business Process Models by Considering Semantic Similarity," *International Journal of Intelligent Engineering and Systems*, vol. 16, no. 2, pp. 166–175, 2023, doi: 10.22266/ijies2023.0430.14.

[7] M. Saad, Y. Zhang, J. Tian, and J. Jia, "A graph database for life cycle inventory using Neo4j," *J. Clean. Prod.*, vol. 393, Mar. 2023, doi: 10.1016/j.jclepro.2023.136344.

[8] T. Peng, L. Gao, R. S. K. Agbozo, Y. Xu, K. Svynarenko, Q. Wu, C. Li, and R. Tang, "Knowledge graph-based mapping and recommendation to automate life cycle assessment," *Adv. Eng. Informatics*, vol. 62, Oct. 2024, doi: 10.1016/j.aei.2024.102752.

[9] I. Waspada, R. Sarno, E. S. Astuti, H. N. Prasetyo, and R. Budiraharjo, "Extending Graph-Based Process Discovery for Hierarchical Block Detection and Incomplete Concurrent Relationship Handling," *IEEE Access*, vol. 11, pp. 123382–123391, 2023, doi: 10.1109/ACCESS.2023.3324742.

[10] W. F. Lamb, T. Wiedmann, J. Pongratz, R. Andrew, M. Crippa, J. G. J. Olivier, D. Wiedenhofer, G. Mattioli, A. A. Khourdajie, J. House, S. Pachauri, M. Figueroa, Y. Saheb, R. Slade, K. Hubacek, L. Sun, S. K. Ribeiro, S. Khennas, S. de la Rue du Can, L. Chapungu, S. J. Davis, I. Bashmakov, H. Dai, S. Dhakal, X. Tan, Y. Geng, B. Gu, and J. Minx, "A review of trends and drivers of greenhouse gas emissions by sector from 1990 to 2018," Jul. 01, 2021, *IOP Publishing Ltd*. doi: 10.1088/1748-9326/abee4e.

[11] C. J. Meinrenken, D. Chen, R. A. Esparaza, V. Iyer, S. P. Paridis, A. Prasad, E. Whillas, "Carbon emissions embodied in product value chains and the role of Life Cycle Assessment in curbing them," *Scientific Reports*, vol. 10, no. 1, Dec. 2020, doi: 10.1038/s41598-020-62030-x.

[12] M. Finkbeiner, A. Inaba, R. B. H. Tan, K. Christiansen, and H. J. Klüppel, "The new international standards for life cycle assessment: ISO 14040 and ISO 14044," 2006, *Springer Verlag*. doi: 10.1065/lca2006.02.002.

[13] M. Natarajan, "From Graph To Knowledge Graph A Short Journey To Unlimited Insights," 2020.

[14] F. Han, Y. Deng, Q. Liu, Y. Zhou, J. Wang, Y. Huang, Q. Zhang, J. Bian, "Construction and application of the knowledge graph method in management of soil pollution in contaminated sites: A case study in South China," *Journal of Environmental Management*, vol. 319, Oct. 2022, doi: 10.1016/j.jenvman.2022.115685.

[15] S. Ji, S. Pan, E. Cambria, P. Marttinen, and P. S. Yu, "A Survey on Knowledge Graphs: Representation, Acquisition, and Applications," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 2, pp. 494–514, Feb. 2022, doi: 10.1109/TNNLS.2021.3070843.

[16] Wesley Ingwersen, "IPCC AR4, AR5, and AR6 20-, 100-, and 500-year GWPs," U.S. EPA Office of Research and Development (ORD).

[17] J. Fuglestvedt J. Rogelj, R. J. Millar, M. Allen, O. Boucher, M. Cain, P. M. Forster, E. Kriegler, D. Shindell, "Implications of possible interpretations of'greenhouse gas balance' in the Paris Agreement," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 376, no. 2119, May 2018, doi: 10.1098/rsta.2016.0445.

[18] Z. Xie, H. Jia, Y. Cai, Z. Huang, T. Luo, and J. Cheng, "Emission accounting, sectoral contributions and gridded spatial distribution of greenhouse gases in a typical petrochemical district of Shanghai," *Atmospheric Pollution Research*, vol. 14, no. 6, 2023, doi: 10.1016/j.apr.2023.101776.

[19] W. Liang, T. Gan, and W. Zhang, "Dynamic evolution of characteristics and decomposition of factors influencing industrial carbon dioxide emissions in China: 1991–2015," *Structural Change and Economic Dynamics*, vol. 49, pp. 93–106, 2019, doi: 10.1016/j.strueco.2018.09.009.