ISBN: 978-628-96613-3-0. ISSN: 2414-6390. Digital Object Identifier: https://dx.doi.org/10.18687/LACCEI2026.1.1.2153

Artificial Intelligence in Chemical Engineering: A Bibliometric and Foresight Analysis with Insights from Peru

Marco Gusukuma[1]; Carlos Guillermo Hernández-Cenzano[2]
[1,2] Departamento de Ingeniería, Pontificia Universidad Católica del Perú, Lima, Perú, mgusukuma@pucp.edu.pe, carlos.hernandez@pucp.edu.pe

Abstract*—Chemical engineering has expanded from its traditional focus on unit operations and petrochemical processing to areas such as biotechnology, nanotechnology, advanced materials, environmental engineering, and sustainability. In parallel, artificial intelligence (AI) has gained relevance in process modeling, optimization, monitoring, and decision support for industrial systems. This study examines the relationship between AI and chemical engineering through a bibliometric review of Scopus-indexed publications from 2015 to 2024, complemented by a foresight-oriented analysis of emerging research directions. The bibliometric results identify five thematic clusters: adsorption and material characterization, process optimization and prediction, computational chemistry, neural networks and chemometrics, and chemical process automation. The temporal analysis shows a shift from early applications in neural networks, process control, and multivariate analysis toward recent developments in bioprocessing, pharmaceutical design, materials discovery, and agri-food quality assessment. The paper also relates these global trends to the Peruvian productive context, where dependence on primary exports coexists with opportunities for selective technological upgrading in mining analytics, agro-industry, pharmaceuticals, and bioprocessing. The findings indicate that data-driven methods can support chemical engineering education, research agendas, and sectoral innovation strategies, particularly when they are linked to human capital formation, data infrastructure, and sustainability-oriented industrial development.

***Keywords*—Artificial intelligence, Chemical engineering, Bibliometric analysis, Technology foresight, Peru**

I. Introduction

A. Evolution of Chemical Engineering

Chemical engineering emerged from the need to design, operate, and improve industrial processes, particularly in petrochemical production, reaction engineering, separations, and large-scale manufacturing. Over time, its scope has widened to include biotechnology, nanotechnology, environmental engineering, advanced materials, and computational methods [1], [2]. This expansion has been shaped by changes in industrial demand, progress in modeling and simulation, and growing pressure to improve resource efficiency, reduce emissions, and develop safer production systems [3], [4]. As a result, the field now contributes not only to the optimization of process units, but also to cleaner technologies, knowledge-intensive products, and sustainability-oriented industrial systems.

B. Artificial Intelligence in Engineering and Process Systems

Artificial intelligence (AI) has entered engineering through a gradual process of adoption. Early applications were based on expert systems, rule-based reasoning, and heuristic algorithms, mainly for diagnosis, control, and technical decision support [5], [6]. Later, machine learning methods expanded the ability to model nonlinear relationships, process large datasets, and identify operational patterns that are difficult to represent with conventional models [7], [8]. More recently, machine learning, deep learning, sensor networks, cloud computing, and industrial data platforms have supported new applications in process monitoring, predictive modeling, optimization, and automation [9].

These developments are especially relevant to chemical engineering, where many systems involve nonlinear behavior, multivariable interactions, uncertainty, and strict requirements for quality, safety, and environmental performance. AI-based methods can support soft sensing, fault detection, process optimization, chemometric analysis, and hybrid models that combine first-principles knowledge with data-driven components [10]–[12]. Their value lies not in replacing established engineering principles, but in extending the analytical capacity available for complex process systems.

C. Peru's Productive Context

The relevance of these technologies can also be examined in economies with limited productive diversification. Peru shows a marked gap between income and economic complexity. According to the Atlas of Economic Complexity, the country ranks 69th globally in income but 125th in economic complexity, reflecting a productive structure still closely linked to primary exports, particularly copper and gold [13]. National statistics also indicate that manufacturing remains concentrated in a reduced group of activities, including food production, metals, and chemicals, which together represent a substantial share of industrial output [14].

This structure creates constraints, but also areas for selective technological upgrading. Dependence on raw materials limits the development of higher value-added activities; however, mining, agro-industry, pharmaceuticals, and bioprocessing offer possible spaces for AI-enabled chemical engineering applications. Mineral feedstock characterization, process analytics, agri-food quality control, and computational support for product development could contribute to incremental improvements in productivity, quality, and value creation, provided that adoption is supported by human capital, data infrastructure, and sector-specific innovation policies [15]–[18].

D. Research Objective

This study examines the integration of artificial intelligence into chemical engineering research through a bibliometric analysis of Scopus-indexed publications from 2015 to 2024, complemented by a foresight-oriented interpretation of future trajectories. The analysis identifies thematic areas, emerging applications, and possible implications of data-driven chemical engineering for resource-based economies, with particular attention to the Peruvian context.

The study has three specific objectives: (1) to analyze recent trends in data-driven chemical engineering between 2015 and 2024, with emphasis on materials design, process optimization, chemometrics, and agri-food quality assessment; (2) to identify potential applications of data-driven approaches in Peruvian sectors such as mining, agro-exports, pharmaceuticals, and bioprocess industries; and (3) to propose an educational and training framework for data-driven chemical engineering, supported by a staged roadmap that links research capabilities, industrial needs, and future technological development.

II. Background and Literature Review

A. Evolution of AI in Industry

Artificial intelligence has been incorporated into industry through successive waves of technological adoption. Early applications in the 1980s were mainly based on expert systems, rule-based inference, and heuristic reasoning, which formalized specialized knowledge for diagnosis, control, and technical decision-making [5], [6]. During the 1990s and 2000s, greater availability of operational data and computing capacity favored the use of statistical learning, pattern recognition, neural networks, and data-driven models in engineering problems [7], [10].

More recently, AI applications have been supported by high-performance computing, industrial sensors, cloud platforms, and the Internet of Things (IoT). These technologies have contributed to digitally enabled production environments, often associated with *Industry 4.0* [8], [9]. In industrial contexts, AI is now used not only for automation, but also for predictive maintenance, process monitoring, quality control, logistics optimization, energy management, and decision support. Its relevance is particularly evident in systems characterized by nonlinear behavior, uncertainty, large volumes of heterogeneous data, and real-time operational constraints [11].

B. AI in Chemical Engineering

Chemical engineering has a long tradition of using computational methods, including process simulation, numerical optimization, computational chemistry, multivariate analysis, and chemometrics. These tools provided the basis for early AI applications in nonlinear process control, fault detection, predictive modeling, and quality monitoring [1], [10]. Expert systems and neural networks were initially used to support process troubleshooting, surrogate modeling, and control of reaction and separation processes. Later, the integration of spectroscopy, process analytical technology (PAT), and multivariate data analysis strengthened the role of chemometrics in real-time monitoring and product quality assessment.

Current developments point toward a broader use of hybrid and data-driven approaches in chemical engineering. Machine learning and deep learning are being applied to soft sensors, anomaly detection, process optimization, molecular design, materials discovery, and bioprocess modeling [12], [19]. Hybrid models are especially relevant because they combine first-principles knowledge with empirical data, allowing process behavior to be modeled without abandoning mechanistic understanding. This balance is important in industries where interpretability, safety, scalability, and regulatory compliance remain central requirements [11]. Recent applications also include pharmaceutical modeling, AI-assisted drug design, hyperspectral quality assessment in agri-food chains, and deep learning for materials development [17], [18], [20]. Table I summarizes these developments by decade.

TABLE I: Chronological evolution of AI in chemical engineering (1980s–2020s).

Period	Representative developments
1980s	Expert systems for process troubleshooting, rule-based reasoning, and early AI-assisted control applications [5].
1990s	Neural networks for process modeling, fuzzy logic control, and AI-supported concurrent engineering applications [6].
2000s	Statistical learning, pattern recognition, and chemometric methods applied to process optimization, spectroscopy, and PAT [7], [10].
2010s	Expansion of machine learning, deep learning, IoT-enabled monitoring, and data-driven approaches for process systems engineering [8], [11].
2020s	Hybrid modeling, AI-assisted molecular and materials design, bioprocess analytics, pharmaceutical modeling, and sustainability-oriented process optimization [12], [17], [18], [20].

C. Economic and Industrial Context in Peru

The Peruvian economy remains closely linked to primary commodities, particularly copper, gold, and other mineral exports [13], [14]. This export structure has supported economic growth, but it has also limited the development of more diversified and knowledge-intensive productive capabilities. The Atlas of Economic Complexity reports a significant gap between Peru's income ranking and its economic complexity ranking, which reflects a productive structure concentrated in activities with relatively low technological sophistication [13]. National statistics similarly show that manufacturing is concentrated in a limited group of activities, including food products, metals, and chemicals [14].

In this context, AI-enabled chemical engineering should not be presented as a general solution to structural constraints. Its potential lies in supporting selective technological upgrading in sectors where Peru already has productive capabilities. In mining, spectroscopy and machine learning can support feedstock characterization, blending decisions, and process monitoring [16]. In agro-industry, machine vision and hyper-spectral imaging can improve quality assessment, moisture prediction, traceability, and sorting processes [17]. In pharmaceutical and bioprocess activities, computational platforms and hybrid models can contribute to product development, process understanding, and candidate prioritization [18], [19]. These applications suggest possible pathways for value creation in resource-based economies, provided that adoption is accompanied by human capital formation, data infrastructure, and sector-specific innovation policies [15]. Figure 1 summarizes the relationship between Peru's export structure and potential AI-enabled value chain opportunities.

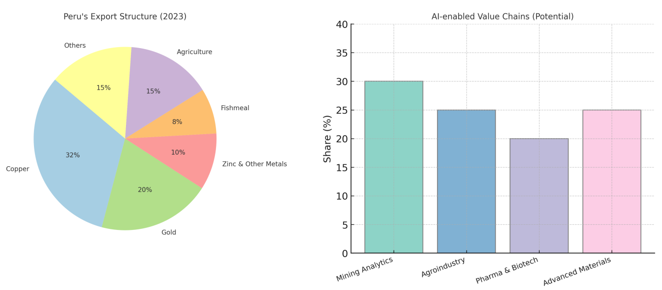

Fig. 1: Peru's export structure compared with opportunities from AI-enabled value chains.

III. Methodology

A. Research Design

The study follows a hybrid design that combines bibliometric analysis with foresight methods. Bibliometric analysis was used to examine the structure and evolution of scientific production on artificial intelligence (AI) in chemical engineering, including publication trends, keyword co-occurrence, thematic clusters, and collaboration patterns. This approach is appropriate for mapping research fields because it allows large volumes of scientific literature to be organized through systematic and replicable procedures [21]–[23].

The foresight component was incorporated to interpret the bibliometric results from a future-oriented perspective. Instead of treating the bibliometric maps only as a description of past publications, the study uses them as evidence for identifying emerging technological trajectories and possible development paths for data-driven chemical engineering. This combination is useful in fields shaped by rapid technological change, where scientific trends may anticipate future industrial capabilities, skills requirements, and policy challenges [24]–[26]. In this paper, the foresight analysis is applied with particular attention to the Peruvian industrial context, where AI-enabled chemical engineering may support selective technological upgrading in resource-based sectors.

B. Data Collection

The primary data source was Scopus [27], selected for its coverage of peer-reviewed literature in engineering, applied sciences, industrial technology, and related interdisciplinary fields. The analysis covered the period 2015–2024, which was selected to capture the most recent decade of development in AI-enabled chemical engineering. The search was limited to documents written in English and classified as articles, reviews, or conference papers.

The search strategy combined general AI terms with chemical engineering and process-related terms using Boolean operators. To make the query logic explicit, the equation was organized with parentheses separating the AI-related terms from the chemical engineering and application-related terms. The search parameters and filters applied in Scopus are summarized in Table II.

TABLE II: Search parameters and filters applied in Scopus.

Parameter	**Description**
Database	Scopus [27]
Period	2015–2024
Search equation	TITLE-ABS-KEY((*"artificial intelligence"* OR *"machine learning"* OR *"deep learning"*) AND (*"chemical engineering"* OR *chemometrics* OR *"process optimization"* OR *sustainability*))
Document types	Articles, reviews, and conference papers
Language	English

The search parameters were reported explicitly to support replication in Scopus. Detailed record-count traceability and multi-database comparison were considered outside the scope of this version and are addressed as limitations and future work.

C. Analysis Process

The analysis was conducted in three stages. First, the records exported from Scopus were reviewed to standardize author keywords, reduce spelling variations, and group equivalent terms where appropriate. This step was necessary because bibliometric maps are sensitive to inconsistencies in terminology, especially in interdisciplinary fields where related concepts may appear under different labels. The resulting dataset was then used to examine descriptive indicators, including annual publication output, thematic areas, and collaboration patterns.

Second, science mapping was performed using VOSviewer [28] and Bibliometrix [29]. VOSviewer was used to generate keyword co-occurrence networks and identify thematic clusters, while Bibliometrix supported the temporal analysis and complementary bibliometric indicators. Keyword co-occurrence was selected because it allows the intellectual structure of a field to be inferred from the frequency with which terms appear together in the same records [22]. The interpretation of clusters considered both the visual structure of the maps and the substantive meaning of the terms, avoiding a purely mechanical reading of the software output.

Third, the bibliometric results were used as inputs for the foresight analysis. The Three Horizons framework was applied to organize short-, medium-, and long-term trajectories for AI-enabled chemical engineering [26]. Morphological analysis was then used to structure the main drivers considered in the scenario construction: technology maturity, regulation and standards, market adoption, and education and skills. These drivers were derived from the bibliometric evidence, the literature review, and the discussion of Peru's productive context. The scenarios are exploratory rather than predictive; their purpose is to identify plausible development paths and strategic implications for research, education, and industrial upgrading. The overall methodological sequence is presented in Figure 2.

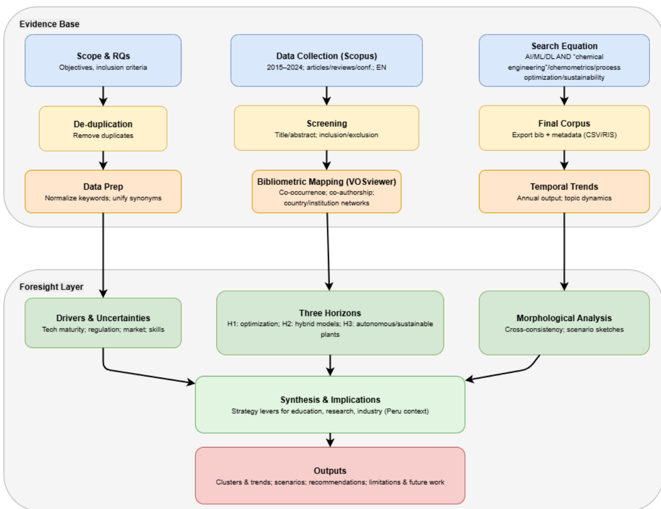

Fig. 2: Methodological framework integrating bibliometric mapping and foresight analysis.

IV. Results

A. Publication Trends

The annual distribution of publications shows sustained growth in research at the intersection of artificial intelligence and chemical engineering during the 2015–2024 period. This pattern is consistent with the broader incorporation of data-driven methods into process systems engineering, process analytics, and digitally supported industrial operations [10], [11]. The year-by-year counts obtained from the Scopus query described in Section III are presented in Fig. 3. The trend suggests a progressive consolidation of the field, particularly in the most recent years.

This growth can be associated with the increasing availability of industrial and experimental data, the maturation of machine learning methods for nonlinear modeling, and the expansion of sensing, computing, and data infrastructure in process industries [8], [9]. In chemical engineering, these conditions have supported applications in process monitoring, soft sensing, optimization, chemometrics, materials discovery, and bioprocess modeling [12], [19].

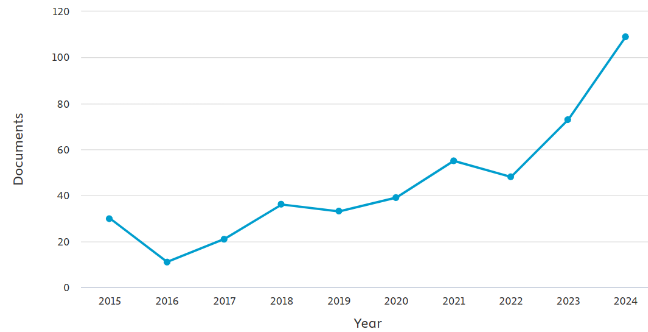

Fig. 3: Annual publications on AI and chemical engineering from 2015 to 2024 based on the Scopus query described in Section III.

B. Co-Occurrence Network

The keyword co-occurrence network identifies five thematic clusters that describe the main research areas in the corpus (Fig. 4; Table III). These clusters correspond to adsorption and material characterization, process optimization and prediction, computational chemistry, neural networks and chemometrics, and process automation and analytics. Their distribution indicates that AI applications in chemical engineering are spread across process systems, analytical chemistry, materials research, and digitally enabled operations.

The presence of terms such as process optimization, soft sensors, model predictive control, and PAT reflects the long-standing connection between chemical engineering and process systems engineering [10], [11]. The grouping of neural networks, chemometrics, spectroscopy, and imaging points to the use of machine learning in quality monitoring and data-rich experimental environments [17], [30]. In addition, the clusters associated with computational chemistry and materials characterization are consistent with recent developments in hybrid modeling and data-driven materials discovery [12], [20].

C. Temporal Evolution

The temporal analysis shows a gradual change in the dominant uses of AI in chemical engineering (Fig. 5). In the first stage, corresponding to early and foundational applications, AI was mainly associated with expert systems, rule-based reasoning, neural networks, and fuzzy control for process troubleshooting, surrogate modeling, and nonlinear control [5], [6]. These approaches established the initial connection between AI and process engineering problems.

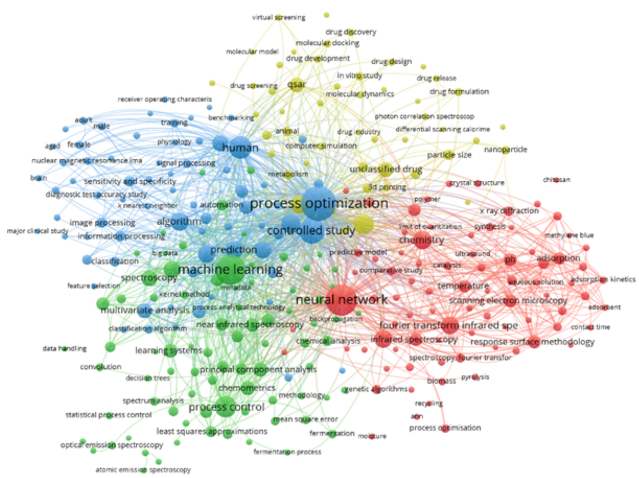

Fig. 4: Keyword co-occurrence network generated with VOSviewer. Colors denote clusters; node size reflects keyword frequency; and edge thickness indicates co-occurrence strength.

TABLE III: Cluster characteristics in the co-occurrence network.

Cluster (theme)	Representative keywords	Methods / references
Adsorption & material characterization	adsorption; porosity; surface area; spectroscopy; SEM/FTIR	Chemometrics and multivariate analysis [30]
Process optimization & prediction	optimization; soft sensors; MPC; quality; PAT	Machine learning, process analytics, and control [10], [11]
Computational chemistry	molecular modeling; thermodynamics; kinetics; simulation	Hybrid and first-principles/data-driven models [1], [12]
Neural networks & chemometrics	PCA/PLS; spectral analysis; imaging; classification	Deep learning and chemometric workflows [17], [30]
Process automation & analytics	sensors; IoT; anomaly detection; dashboards	Industrial analytics and digitally enabled operations [8], [9]

In the second stage, machine learning became more closely linked with process data analytics, spectroscopy, and chemometric workflows. This period reflects the growing use of multivariate analysis for process monitoring, quality assessment, and PAT applications [10], [30]. In the most recent stage, covering the 2020–2024 period, the field moves toward more specialized and data-intensive applications, including AI-assisted drug design, data-driven alloy discovery, hyperspectral quality assessment in agri-food systems, and hybrid modeling for chemical and biochemical processes [17]–[20]. This trajectory suggests a movement from isolated modeling applications toward more integrated data-driven workflows.

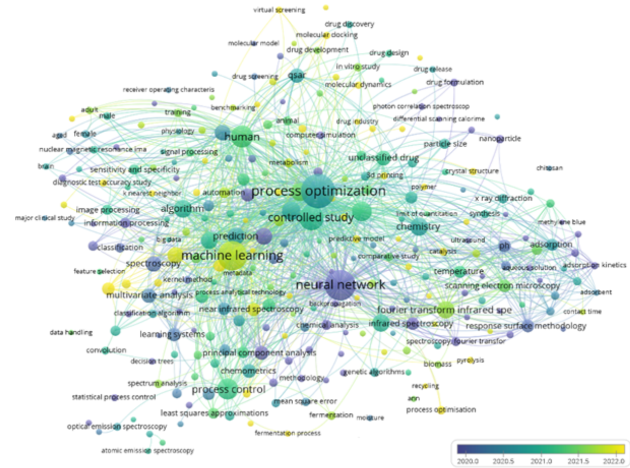

Fig. 5: Timeline of thematic evolution across stages derived from keyword bursts and annual co-occurrence overlays.

D. Thematic Map

The thematic map (Fig. 6) locates topics according to centrality and density. Centrality indicates the degree of connection of a theme with the broader research field, while density reflects its internal development. This representation allows the map to distinguish between motor themes, basic themes, niche themes, and declining or emerging themes [21], [22].

The motor area includes terms such as *machine learning*, *process control*, and *algorithm*, suggesting that these topics are both developed and connected to the rest of the field. *Process optimization* and *artificial neural networks* appear as basic themes, indicating broad relevance but a more dispersed internal structure. A niche area links *neural networks* with *chemistry* and *FTIR*, reflecting specialized applications in spectroscopic analysis and chemical characterization. Terms such as *optical emission spectroscopy* and *semiconductor device manufacture* appear in a less central position, which may indicate a narrower role within the corpus or a declining presence relative to the dominant AI–chemical engineering themes.

Overall, the thematic map is consistent with the co-occurrence clusters reported in Table III. It also supports the interpretation that the field has moved from early ANN-based and rule-based applications toward machine learning, chemometric integration, and hybrid modeling workflows [10], [12], [30].

E. Case Applications

Selected applications from the literature illustrate how the identified research themes are being translated into specific technical uses (Table IV). In extractive industries, spectroscopic characterization combined with machine learning can support feedstock assessment and process decision-making,

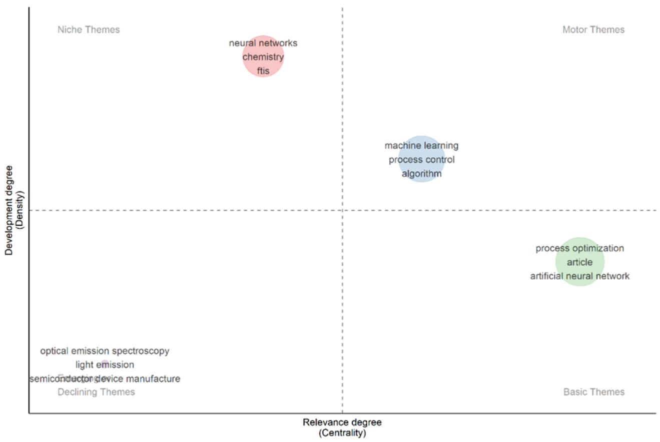

Fig. 6: Thematic map from keyword co-occurrence analysis. Centrality indicates relevance within the field, while density indicates the degree of internal development of each theme.

which is relevant for mineral processing and smelting operations [16]. In pharmaceutical research, AI-driven platforms can assist molecular design by integrating molecular evolution, pharmacokinetic simulations, and multi-parameter optimization [18]. In materials engineering, deep learning approaches are increasingly used to relate composition and structure with material properties, supporting alloy screening and discovery [20].

These cases are not presented as empirical validation of industrial adoption in Peru. Rather, they serve as reference examples for discussing the types of data, methods, and decision contexts through which AI can be connected to chemical engineering practice. They also provide input for the foresight analysis, especially in relation to mining analytics, pharmaceutical and bioprocess development, agri-food quality assessment, and advanced materials.

TABLE IV: Representative case applications and sectors impacted.

Case / domain	**Data / method**	**Outcome / reference**
Spectroscopy for mining analytics	Spectral features, Raman/HSI, chemometrics, and ML	Feedstock characterization and process decision support [16]
AI-driven drug design (AIDD)	Molecular evolution, PBPK simulation, and multi-parameter optimization	Candidate prioritization and design efficiency [18]
Deep alloys and materials discovery	Composition and structure features with deep learning models	Property prediction and alloy screening [20]
Hyperspectral agri-food quality assessment	Machine vision, hyperspectral imaging, and learning models	Moisture prediction and quality control in agricultural products [17]

V. Discussion

A. Strategic Insights

The results suggest that the integration of data and process models is becoming a relevant pathway for improving productivity, quality, safety, and environmental performance in process industries. In chemical engineering, this integration is reflected in applications such as soft sensing, process monitoring, fault detection, model predictive control, hybrid modeling, and data-assisted product development [10]–[12]. These developments do not replace established chemical engineering principles; rather, they extend their analytical scope by combining empirical data, mechanistic knowledge, and computational tools in decision-making.

For Peru, these trends should be interpreted in relation to the country's productive structure. The economy remains strongly linked to primary commodities, particularly mineral exports, while manufacturing and knowledge-intensive activities show limited diversification [13], [14]. In this context, AI-enabled chemical engineering may support selective upgrading in sectors where Peru already has productive capabilities. In mining, spectroscopy combined with machine learning can improve feedstock characterization, blending decisions, and process monitoring [16]. In agro-industry, machine vision and hyperspectral imaging can support quality control, sorting, moisture prediction, and traceability [17]. In pharmaceutical and bioprocess activities, computational design and hybrid modeling may contribute to candidate selection, process understanding, and product development [18], [19]. Materials discovery also represents a possible route toward higher value-added activities through data-driven alloy screening and property prediction [20].

These opportunities should not be interpreted as automatic outcomes of AI adoption. Their realization depends on reliable data, specialized human capital, sector-specific infrastructure, and collaboration among universities, firms, and public institutions. Therefore, the strategic implication is not limited to the introduction of AI tools; it also involves building the organizational and technical conditions required to use them in productive settings. This perspective is consistent with competitiveness approaches that emphasize the transition from resource extraction toward more sophisticated and knowledge-intensive value propositions [15].

B. Educational Implications

The bibliometric and thematic results suggest that data-driven methods should be incorporated into chemical engineering education without weakening disciplinary foundations such as thermodynamics, transport phenomena, reaction engineering, separations, and process design. A balanced curriculum would reinforce statistics, numerical methods, scientific computing, data analysis, and optimization, while connecting these areas with process systems engineering, process analytical technology, model predictive control, and sustainability assessment [10], [11].

Recent work on machine learning and hybrid modeling in chemical and biochemical engineering also indicates that future graduates will need to understand both data-driven and first-principles models, including their assumptions, limitations, validation requirements, and industrial risks [12], [19]. This requires learning experiences that combine theory, computational practice, laboratory data, and sector-based projects. Topics such as data governance, reproducibility, model maintenance, ethics, and lifecycle thinking should also be considered, particularly when AI systems support decisions related to safety, quality, or environmental performance [31].

For Peru, these educational implications are relevant because AI adoption in mining, agro-industry, pharmaceuticals, and bioprocessing will require professionals capable of translating data science techniques into process-specific solutions. Figure 7 outlines a program-level framework that connects foundational quantitative training with process systems, PAT and control, sustainability, and sectoral application studios. The framework also emphasizes industry partnerships and shared data and computing resources as enabling conditions for applied learning and technology transfer.

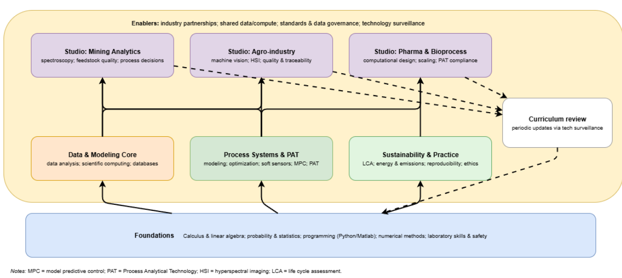

Fig. 7: Proposed framework for integrating data-driven methods into chemical engineering education in Peru. The structure connects core quantitative training to process systems, PAT and control, sustainability, and sectoral application studios, supported by data governance, shared infrastructure, and industry collaboration.

VI. Foresight Analysis

A. Morphological Drivers

The foresight analysis was developed to translate the bibliometric findings into plausible future trajectories for data-driven chemical engineering. The purpose is not to forecast a single outcome, but to organize uncertainty around drivers that may influence the adoption of AI-enabled methods in research, education, and industry. This approach is consistent with foresight methods that combine evidence from current trends with the structured exploration of alternative futures [24]–[26].

Four drivers were selected for the morphological scheme. The first is *technology maturity*, ranging from laboratory prototypes to piloted solutions and validated applications at scale. The second is *regulation and standards*, ranging from nascent guidance to sector-specific rules and harmonized standards. The third is *market adoption*, ranging from niche use to broader industrial scaling and mainstream adoption. The fourth is *education and skills*, ranging from ad-hoc upskilling to elective modules and core curricula with practice-oriented studios. These drivers were derived from the bibliometric evidence, the literature on process systems and hybrid modeling, and the discussion of Peru's productive structure [10]–[13].

The combinations of these drivers were used to construct three internally coherent horizons. Horizon 1 reflects near-term use of AI to improve existing operations. Horizon 2 describes a transition toward hybrid modeling, data infrastructure, and more formal governance. Horizon 3 represents a longer-term configuration in which autonomous and sustainability-constrained operations become more feasible from both technical and institutional perspectives. The scenarios are exploratory and policy-dependent; therefore, they should be read as strategic learning tools rather than predictions.

B. Three Horizons

Table V summarizes the three foresight horizons for data-driven chemical engineering. Horizon 1 focuses on automation and optimization of existing process units through soft sensors, PAT, anomaly detection, and model-based control. This horizon is closest to current industrial practice because it builds on established process systems engineering and process analytics capabilities [10], [11]. Horizon 2 emphasizes hybrid modeling, digital infrastructure, model lifecycle management, and data governance. This stage is especially relevant for applications that require the combination of mechanistic understanding and data-driven components [12], [19]. Horizon 3 considers a longer-term transition toward autonomous, resilient, and sustainability-constrained plants, where closed-loop optimization, human-in-the-loop oversight, and lifecycle indicators are embedded in supervisory decision-making.

Representative applications were selected from the literature and aligned with Peru's potential sectoral opportunities. These include spectroscopy-enabled mining analytics, AI-assisted molecular design, hyperspectral quality assessment in agro-industry, and data-driven materials discovery [16]–[18], [20]. The proposed horizons therefore connect global research trajectories with possible pathways for capability building in resource-based economies.

The scenarios are aligned with the capability gaps and sectoral opportunities discussed for Peru, particularly in mining, agro-export, pharmaceutical, and bioprocess activities [13]–[15].

TABLE V: Three Horizons foresight scenarios for data-driven chemical engineering.

Horizon	Focus	Dominant drivers	Illustrative applications / enablers
H1 (0–3 yrs)	Automation and optimization of existing units; soft sensors; quality by design	Technology maturity: piloted; regulation: developing; market adoption: early majority; education: targeted upskilling	PAT with chemometrics for spectral monitoring; anomaly detection and MPC in continuous or batch lines; mining feedstock characterization to support blending and energy-use decisions [10], [11], [16]
H2 (3–7 yrs)	Hybrid modeling; digital infrastructure; model lifecycle management; data governance	Technology maturity: validated at scale; regulation: sector guidance; market adoption: scaling; education: elective and practice-based modules	Integrated pipelines for computational design and process intensification; agri-food imaging deployed in production environments; PBPK-coupled multi-objective molecular design platforms [12], [17], [18]
H3 (7–12 yrs)	Autonomous and sustainable plants; closed-loop optimization; resilience and safety assurance	Technology maturity: widespread; regulation: harmonized standards; market adoption: mainstream; education: core curricula with sector studios	Data-driven materials discovery coupled with process synthesis; self-tuning operations with human-in-the-loop oversight; integration of emissions, water use, and lifecycle indicators into supervisory control [2], [19], [20]

Notes: PAT = Process Analytical Technology; MPC = model predictive control; PBPK = physiologically based pharmacokinetics; MLOps = lifecycle management for models and data. Timeframes are indicative and depend on policy, investment, infrastructure, regulation, and human capital formation.

VII. Conclusions and Future Work

A. Key Findings

This study examined recent developments at the intersection of data-driven methods and chemical engineering through bibliometric mapping and foresight analysis. The results show sustained growth in publications during the 2015–2024 period (Fig. 3) and a thematic structure organized around five clusters: adsorption and material characterization, process optimization and prediction, computational chemistry, neural networks and chemometrics, and process automation and analytics (Fig. 4, Table III). These clusters show that AI applications in chemical engineering are distributed across process systems, analytical methods, materials research, and digitally enabled operations, rather than being concentrated in a single area.

The temporal analysis indicates a gradual transition from early uses of neural networks, rule-based systems, and optimization models toward more integrated applications involving spectroscopy, chemometrics, hybrid modeling, molecular design, materials discovery, and agri-food quality assessment (Fig. 5) [10], [17], [18], [20], [30]. In this context, AI contributes mainly by connecting experimental data, process knowledge, and computational models to support monitoring, prediction, optimization, and decision-making in complex chemical engineering systems [11], [12], [19].

B. Strategic Implications for Peru

For Peru, the findings are relevant because the country remains highly dependent on primary commodities and shows limited productive diversification [13], [14]. The results do not suggest that AI adoption alone can overcome these structural limitations. However, they indicate that selected AI-enabled chemical engineering applications may support technological upgrading in sectors where Peru already has productive capabilities. Mining analytics, agro-export quality systems, and pharmaceutical or bioprocess development appear as near- and medium-term opportunity areas.

In mining, spectroscopy-assisted feedstock characterization may improve process monitoring and operational decisions [16]. In agro-industry, hyperspectral imaging and machine vision can support quality control, moisture prediction, sorting, and traceability [17]. In pharmaceutical and bioprocess activities, computational design and hybrid modeling may contribute to candidate prioritization, process understanding, and product development [18], [19]. These opportunities are consistent with the broader need to strengthen knowledge-intensive activities and increase the technological content of Peru's productive structure [15].

C. Educational Implications and Roadmap

The results also suggest the need to update chemical engineering education in a gradual and discipline-based manner. Data-driven methods should be incorporated without displacing core foundations such as thermodynamics, transport phenomena, reaction engineering, separations, and process design. The proposed program framework (Fig. 7) connects quantitative foundations with process systems engineering, PAT and control, sustainability, and sectoral application studios. This structure would allow students to develop computational and analytical skills while maintaining a clear connection with chemical engineering problems.

The roadmap derived from the Three Horizons exercise (Table V) organizes this transition into three stages. Horizon 1 emphasizes automation, soft sensors, process monitoring, and quality by design. Horizon 2 focuses on hybrid modeling, data governance, digital infrastructure, and model lifecycle management. Horizon 3 anticipates more advanced configurations involving autonomous, resilient, and sustainability-constrained operations. This sequence suggests that curriculum reform, research agendas, and industry collaboration should evolve together, supported by technology surveillance and periodic review of emerging capabilities [24], [25], [31].

D. Limitations

The study has several limitations. First, the bibliometric analysis is based on Scopus records from 2015 to 2024 and only includes English-language documents; therefore, regional publications, technical reports, Spanish-language outputs, Web of Science records, and patent documents may be underrepresented. Second, the study reports the search parameters and filters applied in Scopus, but does not include a PRISMA-style screening flow or detailed record-count traceability. Third, keyword co-occurrence maps describe associations among terms, but they do not establish causal relationships or directly measure industrial adoption. Fourth, the case applications are representative examples from the literature rather than empirical case studies conducted in Peru. Finally, the foresight component is exploratory and depends on the interpretation of bibliometric trends, literature-based evidence, and contextual assumptions about Peru's productive structure. For this reason, the proposed horizons should be understood as strategic scenarios rather than forecasts.

E. Future Work

Future research should extend the evidence base in four directions. First, the bibliometric corpus should be expanded to include Web of Science and patent databases in order to improve coverage and capture technology-transfer dynamics. Second, patent landscaping and assignee-network analysis could be used to identify firms, universities, and technology domains with greater potential for industrial application. Third, sectoral case studies with operational metrics should be developed in mining, agro-industry, pharmaceuticals, and bioprocessing to assess the benefits, costs, constraints, and implementation risks of AI-enabled chemical engineering applications. Finally, the foresight scenarios should be updated and validated through expert consultation, such as Delphi rounds or structured interviews, to strengthen the roadmap and support decisions on research priorities, infrastructure, curriculum design, and industry partnerships [21], [23], [26].

References

[1] A. Sailaja, "Emerging technologies in chemical engineering: Advancements in process optimization, sustainable practices, and future innovations," *International Journal for Multidimensional Research Perspectives*, vol. 2, no. 5, pp. 59–71, May 2024, doi: 10.61877/ijmrp.v2i5.148.

[2] G. Gwehenberger and M. Narodoslawsky, "Sustainable processes—The challenge of the 21st century for chemical engineering," *Process Safety and Environmental Protection*, vol. 86, no. 5, pp. 321–327, Sep. 2008, doi: 10.1016/j.psep.2008.03.004.

[3] P. R. Westmoreland, "Opportunities and challenges for a golden age of chemical engineering," 2014, doi: 10.1007/s11705-014-1416-z.

[4] J. Li, "Perspectives on chemical engineering in the 21st century," *Huagong Xuebao/Journal of Chemical Industry and Engineering (China)*, vol. 59, no. 8, pp. 1879–1883, 2008.

[5] Y.-H. Pao, "Engineering artificial intelligence," *Engineering Applications of Artificial Intelligence*, vol. 1, no. 1, pp. 5–10, Mar. 1988, doi: 10.1016/0952-1976(88)90062-0.

[6] D. G. Goldstein, "Artificial intelligence in concurrent engineering," *Concurrent Engineering*, vol. 2, no. 2, pp. 83–83, Jun. 1994, doi: 10.1177/1063293X9400200201.

[7] S. C. Shapiro, "Artificial intelligence (AI)," in *Springer Encyclopedia of Cognitive Science*. Berlin/Heidelberg: Springer-Verlag, 2003, pp. 89–93, doi: 10.1007/SpringerReference_6389.

[8] S. D. Mohaghegh, *Artificial Intelligence for Science and Engineering Applications*. Boca Raton, FL, USA: CRC Press, 2024, doi: 10.1201/9781003369356.

[9] M. Khaleel, A. A. Ahmed, and A. Alsharif, "Artificial intelligence in engineering," *Brilliance: Research of Artificial Intelligence*, vol. 3, no. 1, pp. 32–42, Mar. 2023, doi: 10.47709/brilliance.v3i1.2170.

[10] S. J. Qin and L. H. Chiang, "Advances and opportunities in machine learning for process data analytics," *Computers & Chemical Engineering*, vol. 126, pp. 465–473, 2019, doi: 10.1016/j.compchemeng.2019.04.003.

[11] P. Daoutidis *et al.*, "Machine learning in process systems engineering: Challenges and opportunities," *Computers & Chemical Engineering*, 2024, doi: 10.1016/j.compchemeng.2023.108523.

[12] A. M. Schweidtmann, D. Zhang, and M. von Stosch, "A review and perspective on hybrid modeling methodologies," *Digital Chemical Engineering*, vol. 10, Art. no. 100136, 2024, doi: 10.1016/j.dche.2023.100136.

[13] Harvard Growth Lab, "Peru. Atlas of Economic Complexity," 2025. [Online]. Available: https://atlas.hks.harvard.edu/countries/604. Accessed: Jul. 23, 2025.

[14] Instituto Nacional de Estadística e Informática (INEI), *Perú: Cuentas Nacionales 1950–2023: Cuentas de Bienes y Servicios y Cuentas por Sectores Institucionales*. Lima, Peru: INEI, 2024.

[15] M. E. Porter and J. Ramírez Vallejo, "Towards a national value proposition. A strategy for competitiveness," *ReVista: Harvard Review of Latin America*, vol. XIV, no. 1, 2023. [Online]. Available: https://revista.drclas.harvard.edu/towards-a-national-value-proposition-revista. Accessed: Aug. 19, 2025.

[16] A. Bernicky, B. Davis, J. Barnes, and H.-P. Loock, "Spectroscopic characterisation of feedstock for copper smelters by machine-learning," *Canadian Metallurgical Quarterly*, vol. 63, no. 2, pp. 576–585, Apr. 2024, doi: 10.1080/00084433.2023.2215013.

[17] M. B. Farrar *et al.*, "Hyperspectral imaging predicts macadamia nut-in-shell and kernel moisture using machine vision and learning tools," *Computers and Electronics in Agriculture*, vol. 224, Art. no. 109209, Sep. 2024, doi: 10.1016/j.compag.2024.109209.

[18] J. Jones, R. D. Clark, M. S. Lawless, D. W. Miller, and M. Waldman, "The AI-driven drug design (AIDD) platform: An interactive multi-parameter optimization system integrating molecular evolution with physiologically based pharmacokinetic simulations," *Journal of Computer-Aided Molecular Design*, vol. 38, no. 1, Art. no. 14, Dec. 2024, doi: 10.1007/s10822-024-00552-6.

[19] C. L. Gargalo *et al.*, "The role of hybrid modelling in (bio)chemical engineering's digital transition," *Frontiers in Chemical Engineering*, vol. 6, 2024, doi: 10.3389/fceng.2024.1494244.

[20] K. Zheng *et al.*, "Deep alloys: Metal materials empowered by deep learning," *Materials Science in Semiconductor Processing*, vol. 179, Art. no. 108514, Aug. 2024, doi: 10.1016/j.mssp.2024.108514.

[21] N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, "How to conduct a bibliometric analysis: An overview and guidelines," *Journal of Business Research*, vol. 133, pp. 285–296, 2021, doi: 10.1016/j.jbusres.2021.04.070.

[22] I. Zupic and T. Čater, "Bibliometric methods in management and organization," *Organizational Research Methods*, vol. 18, no. 3, pp. 429–472, 2015, doi: 10.1177/1094428114562629.

[23] I. Passas, "Bibliometric analysis: The main steps," *Encyclopedia*, vol. 4, no. 2, pp. 1014–1025, 2024, doi: 10.3390/encyclopedia4020065.

[24] R. Popper, "How are foresight methods selected?," *Foresight*, vol. 10, no. 6, pp. 62–89, 2008, doi: 10.1108/14636680810918586.

[25] A. L. Porter, "Technology foresight: Types and methods," *International Journal of Foresight and Innovation Policy*, vol. 6, no. 1–3, pp. 36–45, 2010.

[26] A. Curry and A. Hodgson, "Seeing in multiple horizons: Connecting futures to strategy," *Journal of Futures Studies*, vol. 13, no. 1, pp. 1–20, 2008.

[27] Elsevier, "Scopus database," 2024. [Online]. Available: https://www.scopus.com. Accessed: Aug. 16, 2025.

[28] N. J. van Eck and L. Waltman, "VOSviewer: A computer program for bibliometric mapping," 2023. [Online]. Available: https://www.vosviewer.com/. Accessed: Aug. 28, 2025.

[29] M. Aria and C. Cuccurullo, "bibliometrix: An R-tool for comprehensive science mapping analysis," *Journal of Informetrics*, vol. 11, no. 4, pp. 959–975, 2017, doi: 10.1016/j.joi.2017.08.007.

[30] A. Baum, R. Moiseyenko, S. Glanville, and T. M. Jørgensen, "Image-based characterization of flocculation processes through PLS inspired representation learning in convolutional neural networks," *Journal of Chemometrics*, vol. 38, no. 6, Jun. 2024, doi: 10.1002/cem.3534.

[31] K. E. Pilario, "Teaching classical machine learning as a graduate-level course in chemical engineering: An algorithmic approach," *Digital Chemical Engineering*, vol. 11, Art. no. 100163, 2024, doi: 10.1016/j.dche.2024.100163.