MRS Communications

Check for updates

Toward artificial intelligence and machine learning-enabled frameworks for improved predictions of lifecycle environmental impacts of functional materials and devices

T. Ibn-Mohammed[iD]**,** WMG, The University of Warwick, Coventry CV4 7AL, UK
K. B. Mustapha, Department of Mechanical, Materials and Manufacturing Engineering, University of Nottingham (Malaysia Campus), 43500 Semenyih, Malaysia
M. Abdulkareem, William Harvey Research Institute, Queen Mary University, London EC1M 6BQ, UK
A. Ucles Fuensanta, WMG, The University of Warwick, Coventry CV4 7AL, UK
V. Pecunia, School of Sustainable Energy Engineering, Simon Fraser University, Surrey, BC V3T 0N1, Canada
C. E. J. Dancer, WMG, The University of Warwick, Coventry CV4 7AL, UK

Address all correspondence to T. Ibn-Mohammed at **t.ibn-mohammed@warwick.ac.uk**

Abstract
The application of functional materials and devices (FM&Ds) underpins numerous products and services, facilitating improved quality of life, but also constitutes a huge environmental burden on the natural ecosystem, prompting the need to quantify their value-chain impact using the bottom-up life cycle assessment (LCA) framework. As the volume of FM&Ds manufactured increases, the LCA calculation speed is constrained due to the time-consuming nature of data collection and processing. Moreover, the bottom-up LCA framework is limited in scope, being typically static or retrospective, and laced with data gap challenges, resulting in the use of proxy values, thus limiting the relevance, accuracy, and quality of results. In this prospective article, we explore how these challenges across all phases of the bottom-up LCA framework can be overcome by harnessing new insights garnered from computationally guided parameterized models enabled by artificial intelligence (AI) methods, such as machine learning (ML), applicable to all products in general and specifically to FM&Ds, for which adoption remains underexplored.

Introduction
The fundamental contributions of materials to humanity's progress are exemplified by historians defining the ages of civilization by the dominant materials that transformed and shaped society during those eras—the stone, iron, and bronze eras.[1,2] The Victorian era was defined by iron, steel, and cement, facilitating powerful engineering inspiration and designs, such as the creation of railways, suspension bridges, passenger liners, and steam engines.[3] Advancements in materials discovery and development have since led to the synthetic materials age, characterized by composites, and plastics, alongside well-designed, artificial engineering materials with better performance compared to traditional materials. These man-made engineering materials and structures are nonetheless considered passive, given that they are constrained by pre-processing and design to offer a restricted set of responses to external stimuli.[4] Facilitated by materials science breakthroughs that ushered in the silicon chip at the dawn of the 21st century, alongside the advancement of several emerging technologies, including nanotechnology, biotechnology, biomimetics, and information and communication technologies (ICTs), the smart materials age started most recently.[1,2] This era has opened up new frontiers that enable the energy transition to renewable, sustainable, and low-carbon technologies while enhancing the quality of life for billions of people across the globe.[5]

Smart materials constitute non-living stimuli-responsive material systems endowed with sensing, actuation, logic, and control functions to respond adaptively to the environment to which they are exposed, in a manner that is usually repetitive and beneficial.[6] They are part of the smart systems functional materials—the set of materials which are used for properties that are not structural (i.e., not used mainly for their mechanical, load-bearing capacity), but characterized by their physical–chemical properties responding to electrical, optical, magnetic, or chemical effects.[7] Functional materials and devices (FM&Ds) encompass conductive polymers, multiferroics, photovoltaics, optoelectronic materials, piezoelectric/ferroelectric materials, functional ceramics, functional alloys, semiconductors, and ionic conductors. In all, these materials are designed with determined functions, such as piezoelectric, magneto-caloric, thermoelectric, triboelectric, pyroelectric, dielectric, or electro-optic effects, among others.[4] Their applications underpin the supply chains of numerous products and services of modern life including ICTs, energy generation and storage, reliable and efficient transport systems, health care systems, smart space structures, intelligent buildings, and many more, all of which are made possible through the use of the materials in radio-frequency transmission and reception, information processing devices, light generation and detection, and sensing and actuation.[4,8] FM&Ds are therefore crucial to a sustainable future, given the high growth and development witnessed
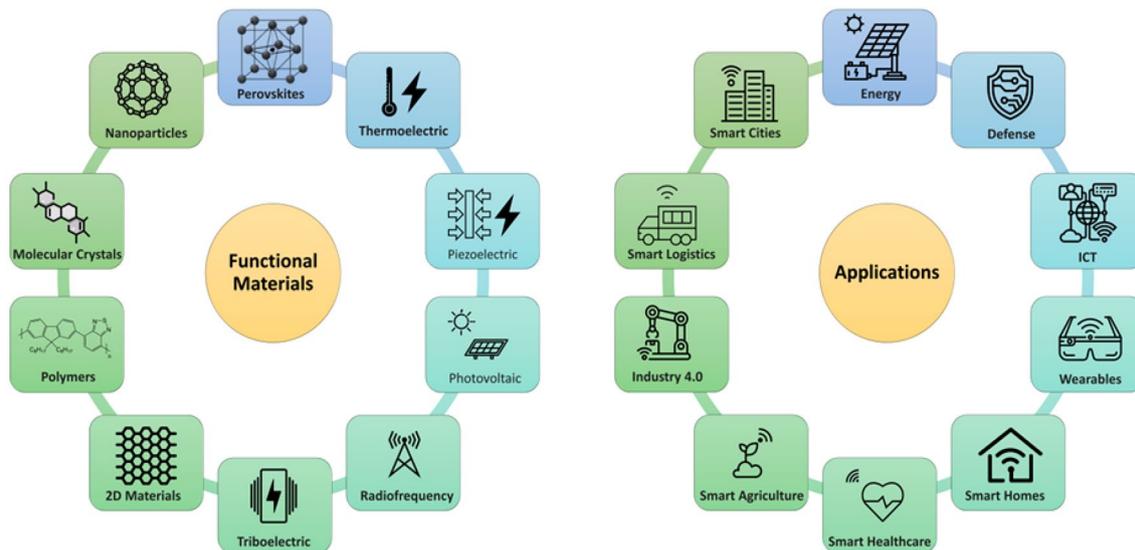
Figure 1. FM&Ds and areas of application. The figure combines materials by class and by chemistry/bonding (e.g., thermoelectrics for the first and polymers for the second). But we note that not all polymers are functional materials, and some will possess properties of the other classes. Fig. 1 icons are from the www.nounproject.com and are credited to (https://creativecommons.org/licenses/by/3.0/): "Photovoltaic" icon by Ian Rahmadi Kurniawan; "Thermoelectric" icon by Abdul Latif; "Radiofrequency" icon by Xinh Studio; "Piezoelectric" icon by ImageCatalog; "Triboelectric" icon by Iconz; "2D materials" icon by Loritas Medina; "Smart Homes" icon by Omar Cruz; "Smart Cities" icon by Justin Blake; "Smart Logistics" icon by Icon Market; "Smart Healthcare" icon by Shocho; "Industry 4.0" icon by Mutualism; and "Smart Agriculture" icon by Thossawat. The icons are from www.flaticon.com and are credited to "Molecular Crystals," "Nanoparticles," and "Defense" icons by Freepik; "ITC" icon by Eucalyp; "Wearables" icon by Uniconlabs; "Energy" icon by Good Ware.

through their discoveries and applications. Figure 1 depicts a selection of different types of FM&Ds, alongside their exploitable properties and areas of application.

Despite their functional use and cross-sector transformational benefits, their developments are not necessarily optimal for sustainability, given the environmental burden attributed to their widespread usage.[9] Therefore, there is an obligation to evaluate the impact of their mass production and their associated supply chain systems on the environment, so as to specify evidence-based mitigation strategies. Advances in the development of FM&Ds must therefore be integrated with lifecycle sustainability constructs such that resulting products and services are designed in a manner that establishes an optimal balance between the environmental burden they impose and improved quality of life.[5] Such evaluations, when conducted in a manner that anticipates foreseeable deleterious consequences while identifying opportunities for improvement and mitigation strategies, can aid the communication of key findings to materials developers.[9] One strategic tool that must therefore be embedded into the functional materials design and development decisions is bottom-up life cycle assessment (LCA), a computational technique for assessing the environmental impacts of products across their entire lifecycle, aiding new material/product design, particularly in terms of environmental performance.[10]

Although bottom-up LCA is a powerful tool, it has methodological limitations, especially regarding data quality and collection (e.g., the choice between average and marginal data or allocation problems), system boundary truncation, time

boundaries, and process modeling.[11] Its calculation speeds are constrained due to the time-consuming and costly nature of data collection and processing.[11] Moreover, the bottom-up LCA framework is limited in scope, typically static or retrospective, and laced with data gap challenges, resulting in the use of proxy values and limiting the relevance and quality of the LCA outputs.[12] A detailed survey of unresolved issues across all phases of bottom-up LCA is provided by Reap et al.[13,14]

In this context, this prospective article explores how the methodological limitations of the bottom-up LCA framework across its phases can be overcome by leveraging the proliferation of material databases and new insights garnered from computationally guided parameterized models enabled by artificial intelligence (AI) methods, such as machine learning (ML). The deployment of AI/ML strategies for environmental risk profiling is growing in different areas, such as the construction industry and the built environment, bioenergy systems, agriculture, and food production, building energy performance, environmental monitoring, ecotoxicological assessment, municipal solid waste management, and e-waste management. Currently, however, their applications remain underexplored for FM&Ds. Accordingly, an overview of how ML methods can be used to evaluate the environmental profile of FM&Ds is also presented.

The remainder of the paper is structured as follows. In Sect. "Life Cycle Assessment Overview," an overview of the bottom-up LCA framework detailing its features, applications, limitations, and the need to couple it with AI/ML capabilities is presented. Sect. "An overview of the applicability of AI/ML methods to LCA" provides an overview of different AI/

ML strategies and their applicability to resolving a wide range of bottom-up LCA-related challenges across a diversity of domains, particularly FM&Ds, covering all phases of LCA, leading to the conclusion in Sect. "Conclusion".

Life cycle assessment overview

Bottom-up LCA is a well-established computational technique used for evaluating the associated environmental impacts throughout the entire lifecycle of an activity, product, or process.[15] The life cycle includes various stages, such as raw material extraction, manufacturing, distribution, use, and different end-of-life scenarios, such as disposal, recycling, or reuse. This holistic perspective renders LCA uniquely suitable as a science-based methodology for assessing the environmental impacts of products, processes, or services. It considers all relevant inputs from the environment, such as raw materials, energy, water, and land use, as well as emissions into air, water, and soil, such as greenhouse gases and pollutants.[15] The primary objective of bottom-up LCA is to systematically identify and quantify the environmental impacts associated with a product, process, or activity, facilitating impact mitigation and promoting sustainability.[9] There are two types of bottom-up LCA approaches namely attributional LCA and consequential LCA. In attributional LCA, the goal is to assess the total environmental burden that can be attributed to a particular product. In consequential LCA, the aim is to assess the overall environmental impact of a product, typically in the context of a particular adoption scenario.[16] The LCA framework typically consists of four main phases: (i) goal and scope definition, (ii) inventory analysis, (iii) impact assessment, and (iv) interpretation,[15] with a wide range of applications as shown in Fig. 2.

The *first phase*, goal and scope definition, involves specifying the purpose of the LCA study, establishing the functional unit (i.e., a reference unit to which inventory data are normalized) and defining the boundaries and assumptions for the analysis.[14] The selection of an appropriate functional unit constitutes an integral component of this phase. This stage sets the foundation for the entire LCA study, clarifying the goals, objectives, and intended applications of the assessment. The *second phase*, inventory analysis, is a technical process that entails the identification and quantification of all the inputs to and outputs from the processes within the defined system boundary. The inputs are energy, water, and raw materials, and the outputs are emissions released to air, water, and land, as well as solid waste; products and co-products.[15] Essentially, inventory analysis is the energy-mass balance of the system and is crucial in establishing a comprehensive and accurate understanding of the environmental burdens associated with the product/process being assessed.[16] Due to the data collection processes involved, this phase is the most intensive and time-consuming.[17] The information required for constructing the lifecycle inventory can be obtained from direct measurements, commercial databases, manufacturers, and literature.

The *third phase*, life cycle impact assessment (LCIA), aims to understand and evaluate environmental impacts based on the inventory data. In this phase, the inventory inputs and outputs are evaluated and classified into different impact categories, such as global warming potential, ozone layer depletion, human health effects, acidification, and eutrophication.[15] Currently, there is no universal list of impact categories that exist but various categories may be used, depending on the specific goals and scope of the LCA study.[17] Essentially, impact assessment facilitates a profound understanding of the implications of the LCA results, and it often involves considering trade-offs and
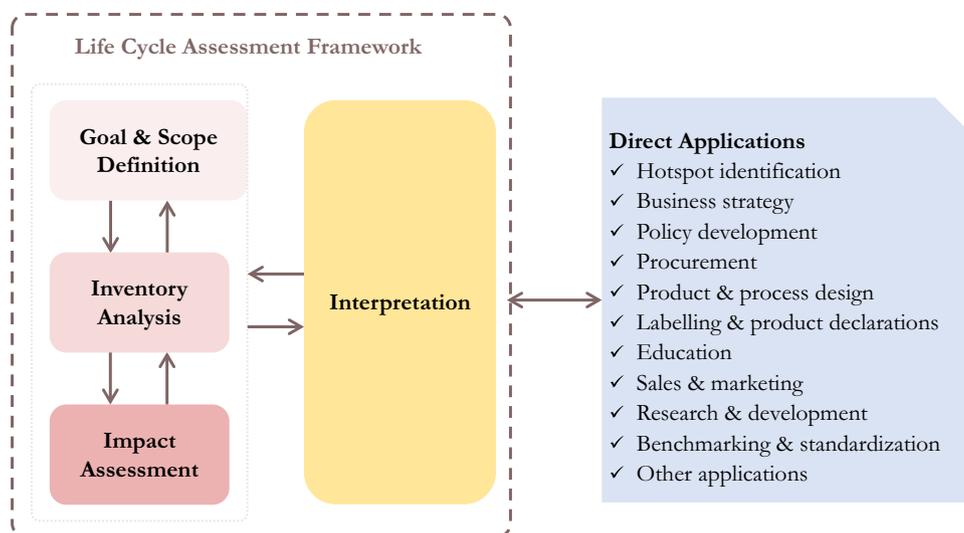

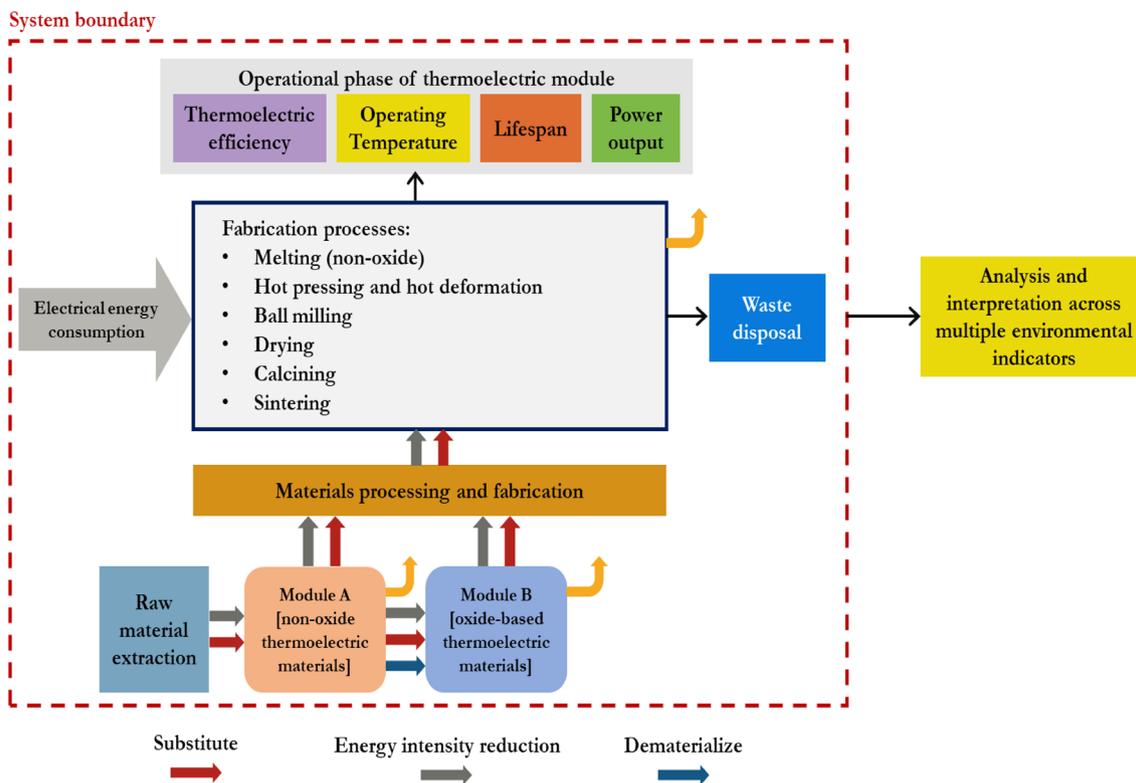**Figure 2.** LCA framework and applications.

Figure 3. A typical system boundary diagram for thermoelectric FM&Ds.[18]

uncertainties associated with the assessment.[13] The LCIA is a crucial and complex step in LCA, and it generally consists of four steps, namely classification, characterization, normalization, and valuation.[13] The *last phase* is the interpretation of the results obtained and entails drawing conclusions and recommendations for mitigation strategies to improve the environmental sustainability of the product or process.[15]

Application of LCA to FM&Ds
The LCA of FM&Ds follows similar steps as conventional ones, highlighted in Sect. "Life Cycle Assessment Overview." For emerging FM&Ds, for example, the LCA is informed by numerous steps including (i) gaining an understanding of the FM&Ds under consideration based on raw material requirements, and laboratory synthesis/manufacturing routes; (ii) system characterization (i.e., systems boundary setting, functional unit identification, modular components specification, material composition, operational efficiencies); (iii) LCI construction based on physical processes, material and energy flows, and upstream supply chain data; (iv) overall impact assessment and environmental profile evaluations across multiple environmental indicators; and (v) performance evaluation, analysis, and interpretation. As a representative example, Figure 3 shows a LCA system boundary diagram for thermoelectric FM&Ds.[18]

LCA has previously been applied to scrutinize the environmental profiles of different FM&Ds, including piezoelectric materials,[17] perovskite solar cells,[19] high volumetric

efficiency capacitors,[20] solid-state batteries,[21] lithium-ion batteries,[22] solid oxide fuel cells,[23] triboelectric nanogenerators,[24] thermoelectric materials,[18] and many more. A review of the LCA of selected FM&Ds is provided by Smith et al.[9] Figure 4 shows a typical LCA output of a laboratory-based *n-type* lanthanum-doped $SrTiO_3$ functional thermoelectric material across six environmental indicators.[18] This includes freshwater aquatic ecotoxicity (FAE), freshwater sedimentary ecotoxicity (FSE), marine aquatic ecotoxicity (MAE), marine sedimentary ecotoxicity (MSE), ionization radiation (IR), and malodorous air (MA). As shown, electrical energy consumption (EEC) during fabrication constitutes the most dominant hotspot across all environmental indicators considered.

Challenges and limitations of LCA
This section describes the challenges and limitations of the LCA framework both in general terms and more specifically to FM&Ds.

General challenges of bottom-up LCA framework
Although bottom-up LCA is a useful tool for assessing the environmental impacts of products/services, it suffers from several well-established limitations.[17] For instance, a principal task in the LCA of products is lifecycle inventory (LCI) modeling, as codified in the ISO LCA operational guideline.[15] Grounded in the need for resources, material, energy, and emission data
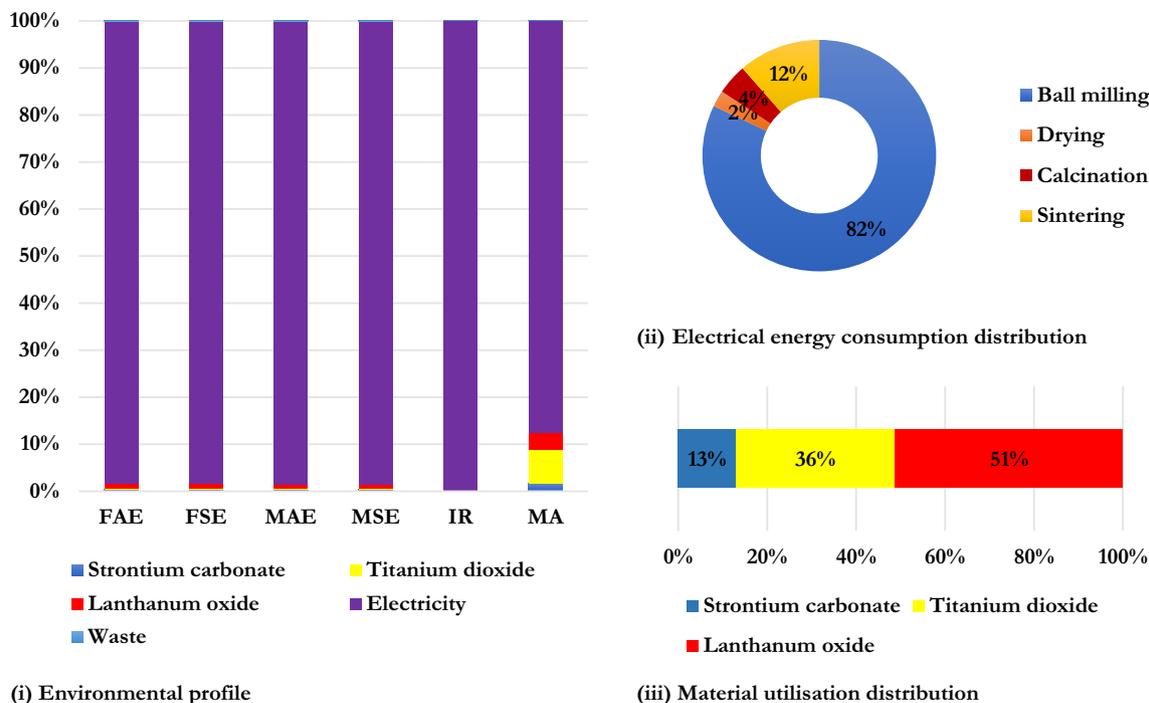
(ii) Electrical energy consumption distribution

(i) Environmental profile

(iii) Material utilisation distribution

Figure 4. Typical LCA output of a functional thermoelectric material. (i) Environmental profile of W/m² functional unit of laboratory-based *n*-type lanthanum-doped SrTiO₃ thermoelectric material, indicating the relative proportions of each of the impact categories based on the unit process exchanges. (ii) Distribution of electrical energy consumption of fabrication processes. (iii) Material utilization impact distribution.

compilation at various lifecycle stages, LCI is time-consuming, and the processes involved in data collection can be expensive, thus constraining LCA calculation speed.[11] Other constraints range from subjectivity or non-representativeness of available data (e.g., where an element of choice of the analyst is required to analyze average and marginal data or where allocations are subjective), hard-to-quantify spatiotemporal variability of available data, low data quality, and poor knowledge of uncertainty surrounding the myriad of LCA parameters in the input space, system boundary truncation, time boundaries, and process modeling issues, to, in some cases, complete lack of data.[11]

The bottom-up LCA framework suffers from limitations in scope by virtue of their static or retrospectively deployed applications. They can also be restricted by data gap challenges, resulting in the use of proxy values and limiting the relevance and quality of the LCA outputs.[12] In most LCA studies, environmental emissions are considered independent of the place and time of occurrence. This poses a challenge at the LCIA phase when translating burdens into environmental impacts due to the requirement of connecting the right burdens with the right impacts at the appropriate time and place.[13] Other challenges at the LCIA phase include the difficulty in selecting the appropriate impact category and methodology due to lack of standardization; spatial variations; local environmental uniqueness; dynamics of the environment; and time horizons.[13] The interpretation of impact assessment results requires careful and integrated consideration of uncertainties, trade-offs, and

limitations of the chosen methods, all of which can be complex depending on the assessment objective. Reap et al.[13,14] provided a detailed survey of unresolved issues across all LCA phases. Overall, all of the challenges associated with bottom-up LCA can lead to misguided emission reduction initiatives, wasted resource allocation, exposure to greenwashing accusations from environmental sustainability blind spots, and failure to pass audits and meet regulation standards.[13,14]

Specific LCA challenges of emerging technologies and FM&Ds

Although the inherent methodological challenges and limitations of the bottom-up LCA framework are highlighted in Sect. "General challenges of bottom-up LCA framework," there are specific challenges pertaining to LCA of emerging FM&Ds, discussed in this section. Generally, the comparative LCA of emerging and mature technologies is predicated upon the technology maturity and stage of development.[7] This prompted Gavankar et al.[25] to conclude that the interpretation of LCA results should be exclusively specified based on universally recognized classification schemes, like the technology readiness level (TRL)[26] and manufacturing readiness level (MRL),[27] both of which describe the technology or manufacturing development from the lowest (i.e., the conceptual fundamentals: TRL/MRL 1) to the highest levels (i.e., the proven applicable technology: TRL 9 or the full rate manufacturing: MRL 10).

Documented studies of LCA about many of the advanced technology sectors have predominantly geared toward retrospective impact assessments of matured technologies, generally referred to as conventional or *ex post* LCA.[28] In other words, *ex post* LCA assesses mature technologies at a current development stage using real-world data,[7] with the primary rationale of leveraging the outcome to prove compliance with environmental regulations or to acquire green certifications.[29] Currently, the majority of studies evaluating the environmental profile of emerging FM&Ds (identified in "Application of LCA to FM&Ds" section) have also employed the conventional LCA framework. However, this poses significant challenges due to the differences in data requirements, availability, and access compared to matured technologies. For instance, conventional LCA of emerging materials or technologies focus mainly on upstream emissions of laboratory fabrication processes[7] and are based on inventory data estimated from such processes using engineering heuristics, stoichiometric relationships, and relevant data from within the literature.[17]

Other key challenges comprise the (i) difficulty in harnessing and testing the uniqueness of different materials to obtain physical optimum, (ii) data gaps resulting in the use of proxy values, and (iii) the use of laboratory-scale processes as a representation for industrial-scale processes.[7,30] Furthermore, EEC during fabrication constitutes the most dominant hotspot (see Fig. 4) for emerging FM&Ds due to inefficient manufacturing with energy-intensive laboratory equipment compared with mature technologies fabricated commercially. This is particularly the case for capacitors,[20] lead-based piezoelectric ceramics,[17] fuel cells,[23] perovskite solar cells,[19] and other emerging materials including biochemicals,[31] nanomaterials,[32] and emerging technologies in general.[33] At the laboratory level, new approaches for lowering EEC have been demonstrated using fabrication routes, such as microwave-assisted sintering, hot extrusion and melt spinning, spark plasma sintering, rapid laser melting, and solidification, alongside the use of sintering aids and low-temperature processing technology, such as cold sintering.[34] Nonetheless, in an industrial setting, these materials will be processed on a large scale and EEC will be minimized by leveraging the capacity of energy-efficient machinery and batch manufacturing processes with a greater throughput.[18] As such, LCA, as designed for matured technologies, requires a thorough interpretation for application to emerging technologies.

Consequently, for the development of advanced FM&Ds, the strategic use of LCA is likely to be the ex-ante application (also known as prospective LCA).[28] As noted by van der Giesen, et al.[35], the *ex ante* LCA centers around conducting rigorous environmental LCA of "*a new technology before it is commercially implemented to guide R&D decisions to make this new technology environmentally competitive as compared to the incumbent technology mix.*" Arvidsson et al.[36] also noted: "*an LCA is prospective when the (emerging) technology being evaluated is in an early phase of development, but is modeled

at a future, more-developed phase.*" In other words, prospective LCA can be used to project how emerging technology such as FM&Ds that are currently available at a lower TRL may look and function at a higher TRL by using different upscaling methods.[37] Studies on upscaling methods (e.g., expert interviews, scenario modeling and analysis, process simulation, molecular structure models, manual calculations, or proxy) for projecting future process performances and the modeling of life cycle inventory data have been reported.[7,37]

Despite the potential of prospective LCA, challenges abound due to the highly complex and nonlinear interactions among key variables in manufacturing processes,[30] thus inhibiting the prediction of the future environmental impacts of emerging FM&Ds. Considering this gap between the theoretical and practical implementation of upscaling scenarios in prospective LCA, in the next section, we explore how coupling AI/ML with LCA can be leveraged to tackle some of the general methodological limitations of bottom-up LCA, alongside the specific challenges of the LCA of emerging FM&Ds.

An overview of the applicability of AI/ML methods to LCA

Recently, Zargar, et al.[38] and Venkatraj and Dixit[39] identified two categories of methods that are currently in use to alleviate the impact of some of the bottlenecks, most notably data gaps challenges, highlighted in "Challenges and limitations of LCA," section namely (i) *mechanistic or deterministic or dynamics-based or mathematical approach* and (ii) *data-driven or empirical techniques*. Under the classical mechanistic approach, knowledge of the physio-mechanical/physio-chemical relationships within a product/process is leveraged for inventory data modeling. Essentially, the approach is predicated upon the ability to develop mathematical expressions for all the dynamical and physical processes and discretize them for numerical evaluations. However, not all products and physical processes lend themselves easily to explicit mathematical relationships. Consequently, an evolving methodology for LCI modeling is the data-driven approach, which allows data to be represented statistically, thus enabling the recognition of reasonable patterns to inform accurate predictions. Driven by the ascendant wave of digitalization in various sectors of the economy, the data-driven approach is endowed with various suites of toolkits to tackle wider scenarios beyond the capability of the mechanistic approaches. The data-driven approach encompasses methods that rely on the latest advances in algorithmic soft computing paradigms, and it has become the preferred approach due to the complexity of systems for which LCA is conducted. Multiple studies have adopted data-driven approaches to overcome data gap challenges in LCA, as will be described in Sect. "Potential role of AI/ML in each phase of LCA" below.

Broadly clustered under the umbrella term of AI, these soft computing techniques have enabled the rapid development
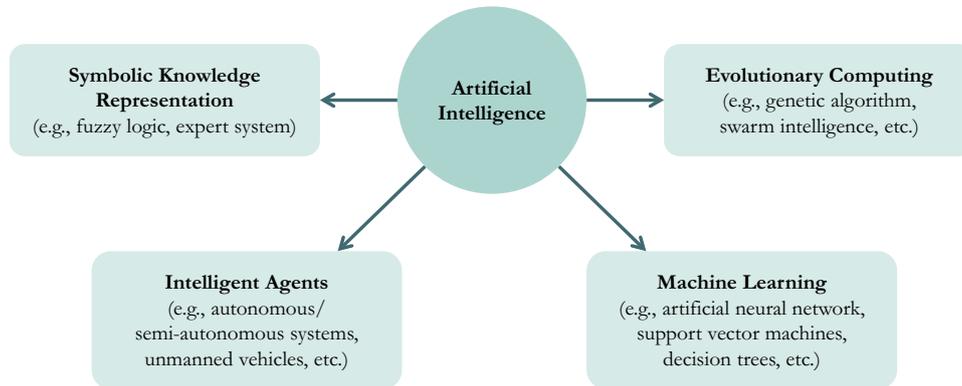
Figure 5. A highlight of contemporary AI techniques.

of perceptual, cognitive, and decision-making intelligence systems.[40] Various sub-categories of the AI framework (see Table A1 in Appendix for a brief description of AI terms and techniques), some of which are highlighted in Fig. 5, have garnered critical interest as evidenced by the growing number of studies in this area.[41,42] The versatility of AI techniques, based on their ability to learn from diverse and voluminous datasets[43] renders them well suited for addressing a wide range of challenges related to bottom-up LCA. Indeed, the integration of AI techniques in LCA can support data collection, modeling, analysis, monitoring, and presentation in the various stages of the product life cycle.

From a survey of the literature, the branch of AI that has seen the most rapid uptake with the convergence of LCA and AI is the ML technique.[44] The utilization of ML methods for improving the prediction of LCA outputs is a direct outcome of digitization, where data are available in formats that can be readily processed by computers. Unlocking the full potential of the vast scope and scale of such data necessitates moving beyond the constraints of traditional statistical methods.[41,42] Accordingly, advanced AI techniques offer a promising avenue for harnessing the true capabilities of ML by leveraging large volume and high multidimensionality data across all stages of LCA.

Figure 6 depicts some of the core methods under the ML schemes. Generally, the robustness of these ML methods has seen them deployed for a wide range of LCA-related tasks. These include inventory optimization, data augmentation for
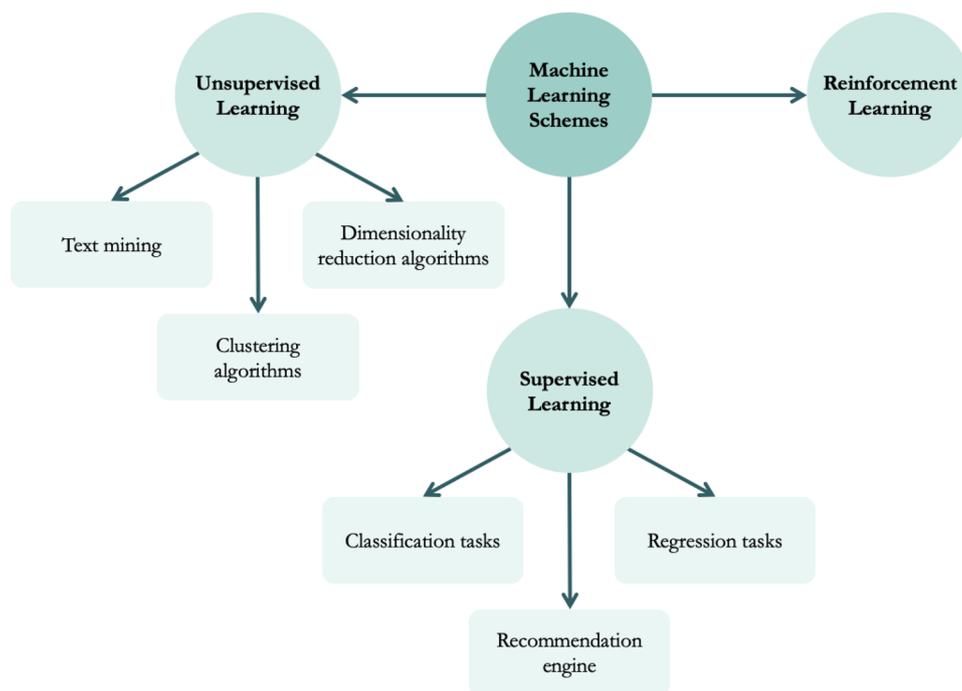

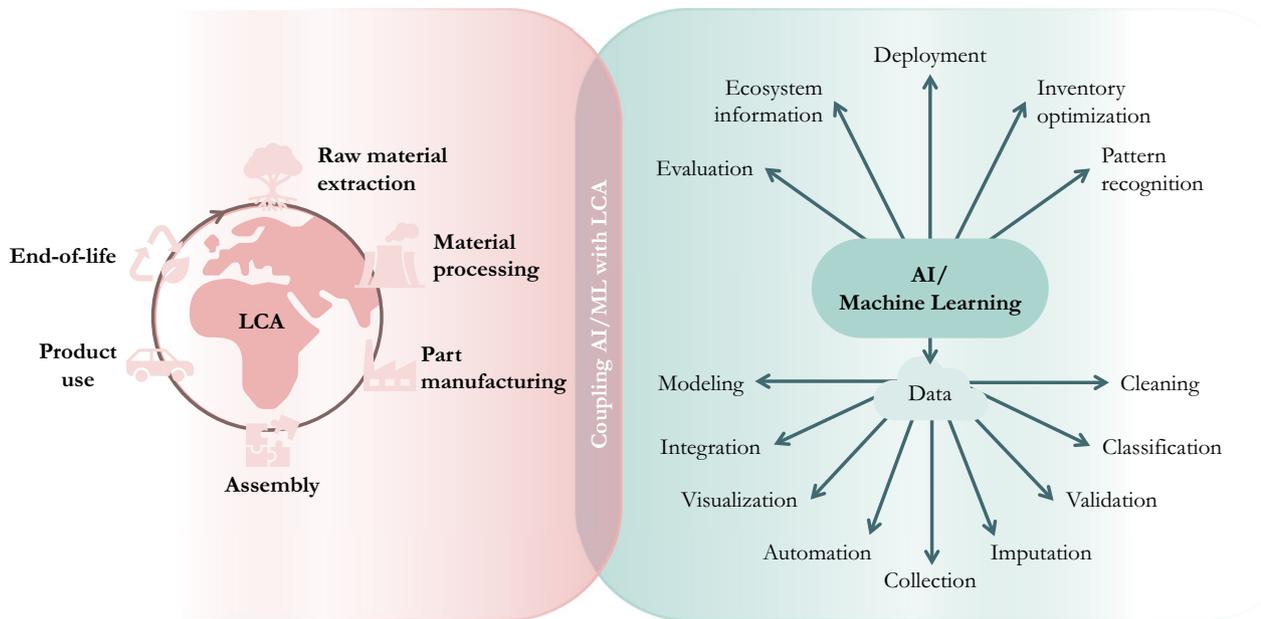**Figure 6.** Principal types of machine learning algorithms.

Figure 7. Coupling of ML and LCA for improved prediction of environmental impact.

LCI, resource utilization, construction of resource forecasting models, prediction of energy/emission hotspots, ecosystem informatics,[45,46] and many more.

Figure 7 provides a schematic representation of the coupling of AI/ML and LCA for improved prediction of environmental impact. In the next section, an overview of how these powerful data-driven techniques can be used to overcome some of the LCA challenges across the four phases, alongside examples of previous studies that have applied them is presented.

Potential role of AI/ML in each phase of LCA

Figure 8 provides a schematic representation of the potential role of AI/ML in each phase of the LCA processes, covering inventory analysis, characterization, normalization, impact assessment, and interpretation. A description of each block of the LCA impact assessment pipeline and the corresponding AI application is provided in each of the subsections that follow. In practice, the implementation of AI/ML strategies in LCA is informed by the building blocks shown in Fig. 9, consisting of four main blocks: (a) data collection; (b) data pre-processing (consisting of data pre-processing, data cleaning and feature engineering sub-blocks); (c) model training (consisting of dataset preparation, model training and model evaluation, & performance analysis sub-blocks); and (d) interpretation.

Potential role of AI in inventory analysis, characterization, and normalization

Unsupervised ML algorithms such as neural net clustering and K-Nearest Neighbors (KNN) can be used for intelligent data collection and extraction from diverse sources, such as databases, sensor networks, and online repositories, thus

streamlining the data collection process, minimizing manual effort, and enhancing data accuracy and comprehensiveness.[47,48] Also, ML algorithms including artificial neural networks (ANN), support vector machines (SVM), and deep neural networks (DNN) can be employed to categorize or cluster data based on predefined criteria, metrics, or learned patterns, facilitating the organization and structuring of inventory data into a form that is more amenable to analysis and interpretation.[49] These algorithms can also be used to validate LCI data by identifying, isolating, and rectifying errors, inconsistencies, and outliers, thus improving the reliability and accuracy of the inventory data, leading to more robust LCA results.[50] ANN, for instance, can handle large and complex datasets, capture nonlinearity in the data, and provide robust predictions. Also, SVMs are particularly effective when dealing with small datasets or datasets with complex patterns.[47,48]

ML algorithms such as Naive Bayes Classifier and Decision Trees[51] can be used for data classification and categorization purposes (e.g., categorizing inventory data into predefined categories, such as different types of materials, energy sources, or processes). The decision trees ML method works by recursively splitting the data into subsets based on the values of the input features and then assigning a category or class label to each subset based on the majority class. This approach can therefore be used to create rules or criteria for categorizing inventory data based on specific attributes or characteristics.[51]

By leveraging models such as DNN and generative adversarial network (GAN),[38,52] LCA data imputation and extrapolation can be achieved, facilitating the completeness and representativeness of inventory data.[53] ML algorithms can integrate and harmonize data from different sources, formats, and units to ensure consistency and comparability in inventory data.[54]
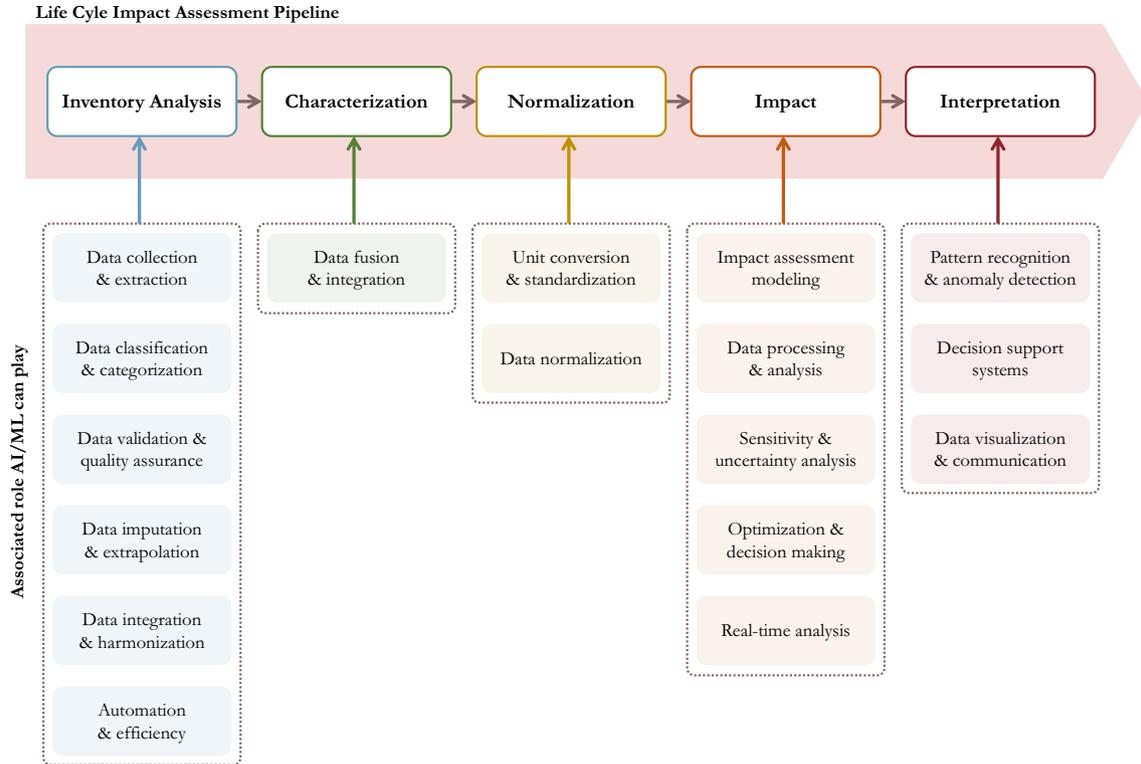
Life Cyle Impact Assessment Pipeline

Figure 8. Illustration of the processes in the LCA analysis pipeline and the associated role of AI/ML.

Moreover, ML methods can be used to integrate LCA results with other relevant information or data, such as economic, social, or other sustainability-related factors. They can also be employed to automatically convert inventory data to a common unit, standardized data formats and align data to a reference database.[55] ML algorithms can be trained to normalize data by accounting for different units and timeframes, ensuring consistency and accuracy in inventory analysis.[56] They can also be used to automatically classify inventory data, reducing the need for manual data entry and processing, thus increasing the efficiency of the inventory analysis stage of LCA.

Natural Language Processing (NLP) techniques can be used to automatically extract relevant data from text-based sources, such as scientific literature, reports, or websites.[38] NLP web scraping algorithms can be utilized to perform tasks such as text mining, entity recognition, and sentiment analysis to extract data related to the inventory of a product or process, including raw material inputs, energy consumption,
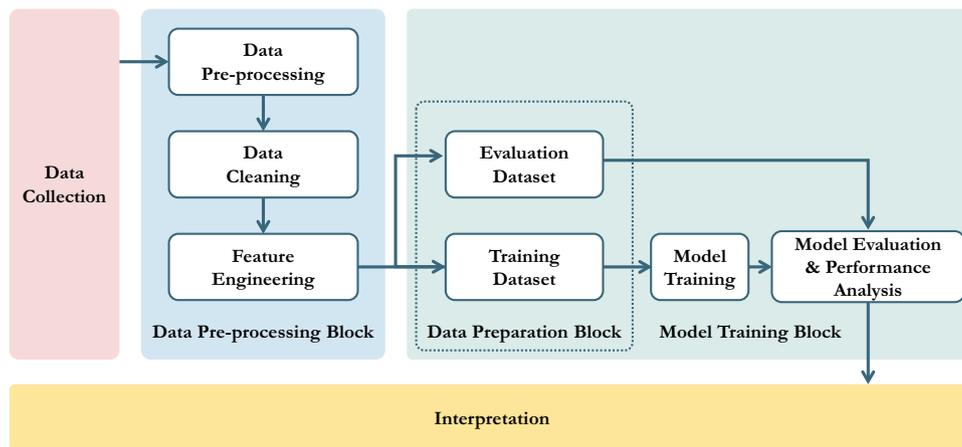

Figure 9. A holistic AI/ML-enabled LCA framework for improved prediction of environmental impact.

emissions estimates, and waste generation.[38] Similarly, deep learning (DL) approaches like convolutional neural networks (CNN) can be used to automatically develop a bill of quantity of materials from images, diagrams, or maps to serve as input data for LCA.[52,57] Clustering is an ML technique that groups similar data points based on similar features or attributes. ML methods such as k-means, hierarchical clustering, or Density-Based Spatial Clustering of Applications with Noise (DBSCAN)[41,42] can be used to group data with similar characteristics. This helps to identify patterns or trends in the data, which can be useful for quality improvement by identifying areas that may require further investigation or improvement.

As shown in Fig. 9, LCA requires the collection and processing of large amounts of data which are not always available in formats that are immediately useable. As such, data pre-processing tools are required for the transformation of raw data into suitable computational formats. These pre-processing tools can range from simple data manipulation techniques to advanced computational methods, such as Principal Component Analysis (PCA), SVM, and ANN.[47,48] PCA is a dimensionality reduction technique that identifies the most important patterns or features in datasets and projects the data onto a reduced dimensional space, thereby transforming it into a lower-dimensional representation while retaining important information. PCA can therefore be used to identify patterns and trends in complex datasets, thus facilitating data fusion from different sources. Similarly, SVM can be used to develop models that can classify and integrate data from different sources based on criteria, such as environmental impact categories or system boundary conditions.[49] Likewise, ANN can learn complex patterns and relationships from large datasets, enabling data integration from different sources by training the network to recognize patterns and make predictions based on the input data.[56,58]

Indeed, not all data collected for LCA are completely accurate, hence the need to perform data cleaning (Fig. 9) using methods such as KNN which is a simple but effective method for data imputation and extrapolation in LCA. For instance, KNN imputes missing values or extrapolates data points by finding the K-nearest neighbors based on similarity metrics and using their values to estimate the missing or extrapolated data. Data cleaning also entails data matching and reconciliation (i.e., aligning and harmonizing data from different sources), which can be achieved by leveraging ML methods, such as clustering, classification, or similarity-based methods. Specifically, these methods can identify similar data points, reconcile conflicting data, and match data from different sources to create a consistent and harmonized dataset for further analysis. Also, similarity-based methods (e.g., cosine similarity or Jaccard similarity) can be used for data matching and reconciliation by calculating the similarity or dissimilarity between data points based on their attributes or other relevant parameters. Time-series analysis methods (such as autoregressive integrated moving averages (ARIMA)) can be used for data extrapolation in LCA when considering temporal factors. Specifically, time-series analysis methods can model the temporal patterns and trends in a given dataset and forecast future values based on historical data, which can be useful for extrapolating data in LCA studies involving time-dependent variables, such as energy consumption or emissions. Generally, similarity-based and time-series methods can be used in conjunction with ML methods to identify and reconcile data points.

Potential role of AI in environmental impact evaluations and interpretation

ML algorithms can be employed to develop advanced models for impact assessment, considering the complex relationships between environmental factors, emissions, and impact categories, including global warming potential (GWP), acidification potential, eutrophication potential, human toxicity, and many others. They can also be used to analyze emissions or energy consumption data to identify patterns, trends, and relationships that can inform the evaluation of environmental impacts as recently demonstrated by Ross, et al.[59] ML algorithms can be utilized to conduct sensitivity analysis and uncertainty assessment in LCA calculations, thus helping to assess the robustness of results to changes in input parameters or assumptions. Besides, ML algorithms can also be employed to estimate uncertainty intervals for ill-defined parameters, improving the accuracy and reliability of LCA calculations[59] They can also be used to optimize the allocation of environmental impacts among different products or processes in a life cycle or identify optimal process configurations that minimize environmental impacts.

At a high level, ML methods that can be used for impact assessment modeling include (i) regression techniques, such as linear regression, multiple regression, or nonlinear regression, used to model the relationship between input variables and environmental impacts; (ii) ANN-based models consist of interconnected nodes organized in layers, and the parameters of these models can be trained to capture complex patterns and relationships between the input and output data variables; (iii) decision trees for decision-making processes or rules formulation to determine the environmental impacts associated with different life cycle stages; (iv) random forests for building an ensemble of decision trees that collectively predict impact values; (v) SVM models for training historical data to predict impact values based on input variables; (vi) ensemble methods such as bagging or boosting to combine multiple ML models such as regression, decision trees, or SVMs to improve prediction accuracy and robust model performance.[60] Overall, the choice of ML method(s) for impact assessment modeling in LCA depends on the characteristics of the data, the level of accuracy/reliability required, and the availability of labeled or historical data.

Furthermore, model evaluation should be carried out to ensure the accuracy and reliability of the impact assessment

models.[60] Datasets are often divided into training and evaluation datasets (e.g., 70% of the dataset can be used for training and the remaining 30% for evaluation); training datasets are used for model training while the evaluation datasets (i.e., those not used during the training process) are used to evaluate the model after training. The evaluation assesses the generalizability of the model on independent datasets. A model has good performance when its accuracy on an evaluation dataset degrades negligibly, meaning that the model has learnt the underlying relationship that produces the data rather than overfitting.

Uncertainties related to data, models, and assumptions are inherent in LCA. Conducting uncertainty and sensitivity analyses to assess the robustness and reliability of the model prediction results is therefore pertinent. Monte Carlo simulation and Bayesian statistics are examples of techniques that are commonly used to randomly sample input parameters from their respective probability distributions and run the LCA models multiple times to obtain a distribution of the results.

Some ways in which ML can be applied at the interpretation phase include pattern recognition and anomaly detection; decision support systems that aid the interpretation of LCA results, facilitating effective decision-making; data visualization; and communication.[41] ML-based methods for supporting data visualization and interpretation include (i) decision tree models, which generate decision rules or thresholds that can be visualized in the form of decision trees, heatmaps, or radar charts to understand the factors driving environmental impacts and (ii) explainable AI (XAI) techniques, which provides interpretable explanations for the predictions or results obtained from ML models. Other XAI techniques, such as feature importance analysis, Local Interpretable Model-agnostic Explanations (LIME),[61] or SHAP (SHapley Additive exPlanations),[62] can be used to understand and interpret the outputs of ML models. This can help improve the transparency of the LCA by providing insights into the underlying mechanisms of the models.

Past studies on the growing applications of AI/ML methods in LCA

Interestingly, the diversity of domains where the quantification of environmental impacts of products/services/processes is imperative has translated into the diversity of AI/ML-enabled LCA across different sub-fields.[63] For instance, Ghoroghi, et al.[64] Barros and Ruschel,[65] Hong et al.[66] and Koyamparambath et al.[41] chronicled the growing collection of technical articles that have embraced ML-assisted LCA within the construction industry and the built environment. Successful applications of ML-enabled environmental impact assessments have also been documented in areas of agriculture[67–70] and other applications. Table I provides a short description of ML algorithms described so far and examples of previous LCA that have adopted them.

Application of AI/ML to LCA of emerging FM&Ds

Beyond the aspects identified above, an important area where the use of ML-assisted LCA holds potential is the development of new chemicals or materials in general and FM&Ds in particular. Despite a noticeable pace of progress in this area, challenges remain. Key obstacles include the uniqueness of different materials in terms of their chemical properties and complexities (e.g., SPIRO-OMeTAD molecule for perovskite solar cell applications[19]), as well as the non-trivial hurdle in the scaling of synthesis methods from the laboratory-scale to industrial production.[7,30] While the coupling of ML for the LCA of various novel materials confers notable universal advantages, this approach has not been implemented evenly across the different categories of materials. Nonetheless, some areas where promising results have already been shown include the (i) prediction of missing data[50,56,77]; (ii) prediction of ecotoxicity characterization factors and impacts for green/functional materials[87,88]; (iii) prediction of lifecycle impact of chemical materials across different categories based on their molecular structure information[51,89]; and (iv) reduction of uncertainty of existing chemical fate model by improving the accuracy of fate factor, which is a function of a chemical's persistence in an environment.[90]

Weyand et al.[7] developed a robust scheme for generating upscaling scenarios of emerging FM&Ds. As such, ML models such as DNN and GAN[38,52] can be used for extrapolation and scaling up of laboratory synthesis methods for FM&Ds to industrial production, based on production site-specific data, thus facilitating the accurate prediction of energy consumption beyond the laboratory. Song et al.[51] developed ANNs to estimate the characterized results of chemical materials using their molecular structural information as inputs, across six environmental impact categories including global warming potential, cumulative energy demand, acidification potential, human health, ecosystem quality, and eco-indicator 99. The application domain of the model was also estimated for each impact category where higher reliability was exhibited, indicating that ANN models can be deployed as an initial screening tool for chemical material lifecycle impacts estimations, even when more reliable information does not exist.

In chemical materials impact assessment, the effect factor which is the overall ecotoxicity impact of a material on the ecosystem is derived from the toxicity to numerous species via Species Sensitivity Distribution (SSDs)—a key parameter for understanding the potential ecotoxicity impacts of chemicals. By leveraging ANN model to process > 2000 experimental toxicity data collected for eight aquatic species across twenty sources, Song[90] estimated the chemical toxicities (i.e., Lethal Concentration (LC50)) to numerous aquatic species, using these to build SSD and to estimate the effect factor of organic chemicals. Using the bootstrapping method, the ANN model output was used to fit SSDs and subsequently

Table I. A short description of popular ML algorithms for LCA and example studies.

ML Techniques	Tasks	Brief Description	Related studies
Logistic regression	Classification	A probability-based classification algorithm premised on the logistic/Sigmoid function.	71
Gaussian Process Regression	Regression	An advanced and versatile tool for nonlinear and non-parametric regression for interpolating between data points that are distributed within a multi-dimensional input space.	72,73
Artificial neural networks (ANN)	Classification and regression	A biologically inspired, layer-based algorithm where the input and output data are connected by edges and training can be achieved by backpropagation.	51,56,58,74
Extreme learning machines	Classification and regression	A feedforward neural network with a single hidden layer.	75
Deep learning	Classification, regression, object detection, natural language processing	A subset of neural networks with more hidden layers.	76,77
Generative adversarial network (GAN)	Classification, data augmentation, image synthesis, etc.	An unsupervised form of deep neural network.	38,52
Adaptive neuro-fuzzy inference system (ANFIS)	Classification, regression, automatic control, time series modeling, etc.	A fuzzy inference system with an architecture constructed on the backbone of an adaptive ANN.	58,78
Decision trees (DT)	Classification and regression	A non-parametric algorithm characterized by the recursive splitting of data into tree-like nodes and branches.	50,79
Gradient-boosted decision trees	Classification, regression, anomaly detection, etc.	An ensemble DT where trees (called learners) are built one at a time using gradient descent and decision from the learners are additively combined in a forward stage-wise manner.	80,81
Random forests	Classification and regression	An ensemble DT with multiple trees that are simultaneously built on random bootstrap samples and decisions from the various trees aggregated at the end.	82,83
Bayesian belief networks	Prediction, hypothesis testing, decision-making under uncertainty, etc.	A probabilistic graphical computing model based on a directed acyclic graph.	84
Support vector machines	Classification, outlier detection, regression, natural language processing, text categorization, etc.	Premised on the use of kernel function and hyperplanes (decision boundaries) for dealing with highly nonlinear, high-dimensional feature space.	85
K-nearest neighbors (KNN)	Classification, semantic segmentation	A simple classification scheme for multi-modal classes.	49,86
Naïve Bayes classifiers	Classification	A family of probabilistic classifiers based on Bayes' theorem.	59,86

used to generate SSDs for more than 8000 chemicals in the ToX21 database.

Traditionally, due to the high resources and time costs of laboratory tests, the evaluation of chemical toxicity has generally been backed by computational toxicology techniques, such as Quantitative Structure–Toxicity Relationships (QSTRs).[91] However, the reliance on linear relationships between chemical structures and biological activities often hinders the effectiveness of QSTR and its variants (such as Ecological Structure Activity Relationships (ECOSAR)). In addition, the generalization of such methods tends to be limited due to a lack of input data. To overcome these limitations, ML methods such as ANN, SVM, and RF have recently been found to show higher accuracy in toxicity estimation than the aforementioned traditional methods.[51,89] The attraction of these ML methods is partly due to their inherent ability to capture important nonlinear relationships that are prevalent in the determination of the aquatic/terrestrial ecotoxicity of environmental pollutants.[92,93]

Despite this, as recently highlighted by Miller et al.[94] the deployment of AI/ML for environmental risk profiling of FM&Ds remains underexplored but some of the ML methods described can be adopted to accomplish this task. Consequent to the notable sparsity of data, the integration of AI/ML with LCA is recognized to hold a plethora of opportunities for *ex ante* LCA in the early-stage development and technology readiness evaluations of new FM&Ds.[29] For these scenarios, the joint use of AI/ML with LCA enables the prediction of energy consumption of a prospective new technology related to FM&Ds, estimation of chemical safety assessment, normalization of LCA outcomes, screening of functional materials for substitution in devices, prioritization of materials with low-carbon footprint, estimation of missing empirical data, and prediction of ecotoxicity of legacy and prospective materials, among others.[44,74,95–97]

While the extension of ML methods coupled with LCA to these tasks will play a transformative role, challenges remain before this can occur, largely due to (i) the unavoidable requirement of large amounts of data for the training of robust ML models[98]; (ii) lack of interpretability due to the "black box" attributes of ML methods, such as ANN, thus rendering the outcome of model development not immediately interpretable to humans[90]; (iii) lack of proper model validation, as almost every ML model excels at interpolation but is constrained at extrapolation, as such model performance output may not reflect the actual performance of the model[90]; and (iv) difficulty in measuring the model uncertainty induced by constrained external validation due to limited experimental data, for example.[98]

Conclusion
FM&Ds are continually being embedded into numerous applications as they can operate in diverse conditions while meeting the wide-ranging needs of consumers. As a result, modern society has witnessed high growth and development through the discovery and applications of these materials. Concerns about the "health" of our planet, therefore, necessitate an evaluation of the environmental profile of these materials at the design or pilot stage before expensive investments and resources are committed. Such evaluations, which are carried out using bottom-up LCA, when conducted in a manner that anticipates foreseeable deleterious consequences while identifying opportunities for improvement and mitigation strategies, can aid the communication of key findings to various stakeholders. However, there are several notable methodological limitations of bottom-up LCA as it is currently being utilized in the literature for environmental impact predictions of products.

Some of the methodological limitations of the bottom-up LCA framework across its phases, with a focus on FM&Ds, can be overcome by harnessing the power of AI/ML techniques. By coupling AI/ML strategies with bottom-up LCA, lifecycle environmental impacts can be predicted with a high degree of accuracy. Despite the capabilities of AI/ML techniques, their potential in the context of LCA can only be fully realized when applied to large and multi-dimensional datasets. Interestingly, with the advent of digitization, the Internet-of-Things (IoT) and advancement in ML methods, large multi-dimensional datasets that can be used for real-time LCA data capture and analysis will increase, bringing unprecedented opportunities. To tap the full potential of these enormous opportunities for LCA through improvements in robustness, standardization, and accuracy of environment impact predictions, multi-disciplinary efforts involving innovation in data collection and collaborations among several stakeholders including sustainability professionals, AI and computer scientists, and experts from different disciplines will be required.

Funding
No funding was received to conduct this work.

Declarations
Conflict of interest
The authors declare that they have no known conflict of interests or competing interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix
See Table II.

Table II. A brief description of AI terms and techniques.

Term	Description
Artificial intelligence (AI)	AI refers to the ability of machines or computer programs to perform specific tasks that typically require human intelligence, such as visual perception and speech recognition, and are developed to learn from experience or data.
Machine learning (ML)	A subset of AI that involves training models to 'learn' patterns and make predictions or decisions based on some input variables. Learning involves using computational algorithms to improve the performance of the models on a specific task over time by adjusting model parameters to optimize a specific objective function.
Supervised learning	A set of ML algorithms for developing models that are trained on a labeled dataset consisting of input data and corresponding output data. The algorithm learns the relationship between the input data and the output data by analyzing and finding patterns and structures in the training data and then uses this knowledge to predict the output for new, unseen input data.
Unsupervised learning	A set of ML algorithms for developing models using unlabelled dataset. The algorithms try to identify patterns or relationships in the data by clustering or grouping similar data points together which can be used in applications such as anomaly detection or feature extraction.
Reinforcement learning	A set of ML algorithms in which models or agents learn to make decisions based on trial-and-error interaction with their environments. During the learning process, an agent receives feedback in the form of rewards or penalties based on its actions in the environment and uses this feedback to adjust its behavior over time.
Artificial neural networks (ANN)	ANN is a powerful ML technique that is composed of interconnected nodes called artificial neurons or nodes that are organized into layers; these nodes communicate with each other through weighted connections. The weights are optimized during the learning process to minimize the difference between the predicted output and the actual output for a given input data. ANN are trained using large datasets.
Deep learning (DL)	A set of ANN algorithms involving a large number of layers. The term 'deep' emphasizes the fact that there are many layers, often referred to as hidden layers, between the input and the output layers.
Convolutional neural network (CNN)	A set of DL algorithms that is commonly used in applications involving images or videos. It consists of multiple layers that perform convolution operations, allowing the network to learn spatial features and patterns in data.
Natural language processing (NLP)	A set of AI algorithms involving the development of models to understand, interpret, and generate human language. Recently, ML algorithms are being used to develop NLP models (as against using statistical methods and algorithms) and applications include data or information extraction from texts.
Principal component analysis (PCA)	PCA is a dimensionality reduction technique in ML that is used to identify patterns and reduce the number of features in a dataset while retaining as much information as possible. The technique works by identifying the principal components of a dataset, which are linear combinations of the original features that account for the most variability in the data making PCA suitable for removing noise and redundancy in the data.
Regression problems	In ML, regression problems involve predicting a continuous numerical value, such as predicting the price of a house or the amount of rainfall in a particular region. The goal is to find a function that maps the input variables the output variable based on a given set of training data.
Classification problems	In ML, classification problems involve assigning a label or category to a given input data point based on its features or attributes. They are supervised learning problem that involves learning to predict a discrete or categorical target variable (class label) based on one or more predictor variables (features). Some ML problems can have both discrete and continuous aspects. For example, predicting the price of a house could involve both classifying it as a certain type of house (classification task) and then predicting its price based on its features (regression task).
Regression algorithm	In ML, regression is a type of supervised learning algorithm used for predicting a continuous outcome variable based on one or more input variables. The goal of regression analysis is to estimate the relationship between the input variables and the outcome variable and use this relationship to make predictions on new data. The output of a regression model is a continuous numerical value, such as a price, a quantity, or a probability.
Support vector machines (SVM)	SVM is a supervised ML method that can be used for classification or regression tasks. The algorithm tries to find a hyperplane in a high-dimensional space that can best separate the data into different classes.
Decision trees (DT)	DT is a supervised ML method that builds a classification or regression model in the form of a tree structure by recursively splitting the dataset into subsets based on the values of the input features and assigning a class or value to each subset based on the majority class or mean value of the target variable.

Table II. (continued)

Term	Description
Random forests (RF)	RF is a supervised ML algorithm that is commonly used for classification and regression tasks. It works by creating multiple decision trees and aggregating the outputs of all decision trees (e.g., by taking the maximum prediction in the case of classification or the mean of the predictions in the case of regression) to make a final prediction. In a random forest, each decision tree is trained on a random subset of the data and a random subset of the features.
Naive Bayes (NB)	NB is a probabilistic ML method that is based on Bayes' theorem and the assumption of conditional independence between features. The algorithm calculates the probability of each class label given the input features and selects the class with the highest probability as the output. Naive Bayes algorithm is a fast, simple, and can handle high-dimensional data without requiring a large dataset.
K-nearest neighbors (KNN)	KNN is a supervised ML method used for classification and regression tasks. For a given new data point, the algorithm identifies its k-nearest neighbors based on some distance metric (using, e.g., Euclidean distance) from the training dataset. The output for the new data point is then based on the class or average value of its k-nearest neighbors. KNN is a simple and versatile algorithm, but it may be sensitive to the choice of distance metric and the value of k.
K-Means	K-means is a clustering algorithm, making it an unsupervised ML algorithm. It groups a set of data points into K number of clusters, where K is a user-defined parameter, by iteratively assigning each data point to its nearest centroid and then recalculating the centroids – repeating the process until the centroids of the clusters no longer change or a maximum number of iterations is reached.

Open Access

References

1. B. Zhang, On typical materials acting as the dividing standard of the development stages of human substance civilization. Interdiscip. Descr. Complex Syst.: INDECS **10**(2), 114–126 (2012)
2. L.A. Dobrzański, Significance of materials science for the future development of societies. J. Mater. Process. Technol. **175**(1–3), 133–148 (2006)
3. The Guardian. Why the story of materials is really the story of civilisation. The Guardian Newspaper. https://www.theguardian.com/science/2014/sep/14/story-of-materials-human-civilisation-mark-miodownik. Accessed 29 April 2023
4. Department of Trade and Industry. Functional Materials Report. Materials Innovation and Growth Team. http://www.matuk.co.uk/docs/Functioanmat.pdf. Accessed 22 March 2023
5. R.E. Kirchain Jr., J.R. Gregory, E.A. Olivetti, Environmental life-cycle assessment. Nat. Mater. **16**(7), 693 (2017)
6. H. B. Strock, Emerging smart materials systems, American Ceramic Society Bulletin, vol. 75, no. 4, 1996.
7. S. Weyand, K. Kawajiri, C. Mortan, L. Schebek, Scheme for generating upscaling scenarios of emerging functional materials based energy technologies in prospective LCA (UpFunMatLCA). J. Ind Ecol **27**(3), 676–692 (2023)
8. V. Pecunia et al., Roadmap on energy harvesting materials. J. Phys.: Mater. (2023). https://doi.org/10.1088/2515-7639/acc550
9. L. Smith, T. Ibn-Mohammed, S.L. Koh, I.M. Reaney, Life cycle assessment of functional materials and devices: opportunities, challenges and current and future trends. J. Am. Ceram. Soc. **102**(12), 28 (2019). https://doi.org/10.1111/jace.16712
10. S. Hellweg, L. Milà i Canals, Emerging approaches, challenges and opportunities in life cycle assessment. Science **344**(6188), 1109–1113 (2014)
11. J.B. Guinée, Handbook on life cycle assessment operational guide to the ISO standards. Int. J. LCA **7**(5), 311–313 (2002)
12. H. Ward, L. Wenz, J.C. Steckel, J.C. Minx, Truncation error estimates in process life cycle assessment using input-output analysis. J. Ind. Ecol. **22**(5), 1080–1091 (2018)
13. J. Reap, F. Roman, S. Duncan, B. Bras, A survey of unresolved problems in life cycle assessment: part 2: impact assessment and interpretation. Int. J. LCA **13**, 374–388 (2008)
14. J. Reap, F. Roman, S. Duncan, B. Bras, A survey of unresolved problems in life cycle assessment: part 1: goal and scope and inventory analysis. Int. J. LCA **13**, 290–300 (2008)
15. Environmental management: life cycle assessment; requirements and guidelines, ISO, 2006.
16. T. Schaubroeck, S. Schaubroeck, R. Heijungs, A. Zamagni, M. Brandão, E. Benetto, Attributional & consequential life cycle assessment: Definitions, conceptual characteristics and modelling restrictions. Sustainability **13**(13), 7386 (2021)
17. T. Ibn-Mohammed et al., Integrated hybrid life cycle assessment and supply chain environmental profile evaluations of lead-based (lead zirconate titanate) versus lead-free (potassium sodium niobate) piezoelectric ceramics. Energy Environ. Sci. **9**(11), 3495–3520 (2016)
18. T. Ibn-Mohammed et al., Techno-environmental analysis of material substitution in thermoelectric modules: non-oxide (bismuth telluride alloys) vs. oxide-based (lanthanum-doped strontium titanate and calcium cobaltite) materials. Energy Convers. Manag.: X **19**, 100395 (2023)
19. T. Ibn-Mohammed et al., Perovskite solar cells: an integrated hybrid lifecycle assessment and review in comparison with other photovoltaic technologies. Renew. Sustain. Energy Rev. **80**, 1321–1344 (2017)
20. L. Smith, T. Ibn-Mohammed, S.L. Koh, I.M. Reaney, Life cycle assessment and environmental profile evaluations of high volumetric efficiency capacitors. Appl. Energy **220**, 496–513 (2018)
21. L. Smith, T. Ibn-Mohammed, D. Astudillo, S. Brown, I.M. Reaney, S.L. Koh, The Role of Cycle Life on the Environmental Impact of Li6. 4La3Zr1.

4Ta0. 6012 based Solid-State Batteries. Adv. Sustain. Syst. **5**(2), 2000241 (2021)

22. X. Sun, X. Luo, Z. Zhang, F. Meng, J. Yang, Life cycle assessment of lithium nickel cobalt manganese oxide (NCM) batteries for electric passenger vehicles. J. Clean. Prod. **273**, 123006 (2020)

23. L. Smith, T. Ibn-Mohammed, F. Yang, I.M. Reaney, D.C. Sinclair, S.L. Koh, Comparative environmental profile assessments of commercial and novel material structures for solid oxide fuel cells. Appl. Energy **235**, 1300–1313 (2019)

24. A. Ahmed et al., Environmental life cycle assessment and techno-economic analysis of triboelectric nanogenerators. Energy Environ. Sci. **10**(3), 653–671 (2017)

25. S. Gavankar, S. Suh, A.A. Keller, The role of scale and technology maturity in life cycle assessment of emerging technologies: a case study on carbon nanotubes. J. Ind. Ecol. **19**(1), 51–60 (2015)

26. S. R. Hirshorn, L. D. Voss, and L. K. Bromley, Nasa systems engineering handbook, 2017.

27. US DoD, Manufacturing readiness level (MRL) deskbook version 2.0, United States Department of Defence Manufacturing Technology Program, 2011.

28. S. Cucurachi, C. van der Giesen, J. Guinée, Ex-ante LCA of emerging technologies. Procedia CIRP **69**, 463–468 (2018)

29. P. Karka, S. Papadokonstantakis, A. Kokossis, Digitizing sustainable process development: From ex-post to ex-ante LCA using machine-learning to evaluate bio-based process technologies ahead of detailed design. Chem. Eng. Sci. **250**, 117339 (2022). https://doi.org/10.1016/j.ces.2021.117339

30. F. Piccinno, R. Hischier, S. Seeger, C. Som, From laboratory to industrial scale: a scale-up framework for chemical processes in life cycle assessment studies. J. Clean. Prod. **135**, 1085–1097 (2016)

31. Ó. Ögmundarson, M.J. Herrgård, J. Forster, M.Z. Hauschild, P. Fantke, Addressing environmental sustainability of biochemicals. Nat. Sustain. **3**(3), 167–174 (2020)

32. A.C. Hetherington, A.L. Borrion, O.G. Griffiths, M.C. McManus, Use of LCA as a development tool within early research: challenges and issues across different sectors. Int J LCA **19**, 130–143 (2014)

33. N. Thonemann, A. Schulte, D. Maga, How to conduct prospective life cycle assessment for emerging technologies? A systematic review and methodological guidance. Sustainability **12**(3), 1192 (2020)

34. T. Ibn-Mohammed et al., Decarbonising ceramic manufacturing: A techno-economic analysis of energy efficient sintering technologies in the functional materials sector, Journal of the European Ceramic Society, 2019.

35. C. van der Giesen, S. Cucurachi, J. Guinée, G.J. Kramer, A. Tukker, A critical view on the current application of LCA for new technologies and recommendations for improved practice. J. Clean. Prod. **259**, 120904 (2020)

36. R. Arvidsson et al., Environmental assessment of emerging technologies: recommendations for prospective LCA. J. Ind. Ecol. **22**(6), 1286–1294 (2018)

37. N. Tsoy, B. Steubing, C. van der Giesen, J. Guinée, Upscaling methods used in ex ante life cycle assessment of emerging technologies: a review. Int. J. LCA **25**, 1680–1692 (2020)

38. S. Zargar, Y. Yao, Q. Tu, A review of inventory modeling methods for missing data in life cycle assessment. J. Ind. Ecol. **26**(5), 1676–1689 (2022)

39. V. Venkatraj, M.K. Dixit, Challenges in implementing data-driven approaches for building life cycle energy assessment: a review. Renew. Sustain. Energy Rev. **160**, 112327 (2022). https://doi.org/10.1016/j.rser.2022.112327

40. Y. Xu et al., Artificial intelligence: a powerful paradigm for scientific research. Innovation **2**(4), 100179 (2021). https://doi.org/10.1016/j.xinn.2021.100179

41. A. Koyamparambath, N. Adibi, C. Szablewski, S.A. Adibi, G. Sonnemann, Implementing artificial intelligence techniques to predict environmental impacts: case of construction products. Sustainability **14**(6), 3699 (2022)

42. P. Hou, O. Jolliet, J. Zhu, M. Xu, Estimate ecotoxicity characterization factors for chemicals in life cycle assessment using machine learning models. Environ. Int. **135**, 105393 (2020). https://doi.org/10.1016/j.envint.2019.105393

43. D. Bzdok, N. Altman, M. Krzywinski, Points of significance. Stat. Versus Machine Learn. **15**(4), 233–234 (2018)

44. M. Akhshik, A. Bilton, J. Tjong, C.V. Singh, O. Faruk, M. Sain, Prediction of greenhouse gas emissions reductions via machine learning algorithms: toward an artificial intelligence-based life cycle assessment for automotive lightweighting. Sustain. Mater. Technol. **3**, e00370 (2022). https://doi.org/10.1016/j.susmat.2021.e00370

45. H. Abbasimehr, R. Paki, A. Bahrini, A novel XGBoost-based featurization approach to forecast renewable energy consumption with deep learning models,". Sustain. Comput.: Inform. Syst. **38**, 100863 (2023). https://doi.org/10.1016/j.suscom.2023.100863

46. M. Algren, W. Fisher, A.E. Landis, Chapter 8 - Machine learning in life cycle assessment, in *Data Science Applied to Sustainability Analysis*. ed. by J. Dunn, P. Balaprakash (Elsevier, Amsterdam, 2021)

47. B. Köck, A. Friedl, S. Serna Loaiza, W. Wukovits, B. Mihalyi-Schneider, Automation of life cycle assessment—A critical review of developments in the field of life cycle inventory analysis. Sustainability **15**(6), 5531 (2023)

48. D. Wehner, T. Prenzel, T. Betten, A.-K. Briem, S. H. Hong, and R. Ilg, The Sustainability Data Science Life Cycle for automating multi-purpose LCA workflows for the analysis of large product portfolios, In E3S Web of Conferences 2022. EDP Sciences, 349, 11003

49. H. Yılmaz, M. Şahin, Solar panel energy production forecasting by machine learning methods and contribution of lifespan to sustainability. Int. J. Environ. Sci. Technol. (2023). https://doi.org/10.1007/s13762-023-05110-5

50. B. Zhao, C. Shuai, P. Hou, S. Qu, M. Xu, Estimation of unit process data for life cycle assessment using a decision tree-based approach. Environ. Sci. Technol. **55**(12), 8439–8446 (2021)

51. R. Song, A.A. Keller, S. Suh, Rapid life-cycle impact screening using artificial neural networks. Environ. Sci. Technol. **51**(18), 10777–10785 (2017)

52. A.J. Green et al., Leveraging high-throughput screening data, deep neural networks, and conditional generative adversarial networks to advance predictive toxicology. PLoS Comput. Biol. **17**(7), e1009135 (2021)

53. X. Chen, C. Shuai, B. Zhao, Y. Zhang, K. Li, Imputing environmental impact missing data of the industrial sector for Chinese cities: a machine learning approach. Environ. Impact Assess. Rev. **100**, 107050 (2023)

54. F.N. Elouariaghli, S.M. Kozderka, T.G. Quaranta, F.D. Pena, F.B. Rose, S.Y. Hoarau, Eco-design and life cycle management: consequential life cycle assessment, artificial intelligence and green IT,". IFAC-PapersOnLine **55**(5), 49–53 (2022). https://doi.org/10.1016/j.ifacol.2022.07.638

55. S. Sala, D. Marinov, D. Pennington, Spatial differentiation of chemical removal rates from air in life cycle impact assessment. Int. J. Life Cycle Assess. **16**, 748–760 (2011)

56. Y. Sun, X. Wang, N. Ren, Y. Liu, S. You, Improved machine learning models by data processing for predicting life-cycle environmental impacts of chemicals. Environ. Sci. Technol. **57**(8), 3434–3444 (2022). https://doi.org/10.1021/acs.est.2c04945

57. Y. Xu, J. Pei, L. Lai, Deep learning based regression and multiclass models for acute oral toxicity prediction with automatic chemical feature extraction. J. Chem. Inf. Model. **57**(11), 2672–2685 (2017)

58. A. Nabavi-Pelesaraei, S. Rafiee, S.S. Mohtasebi, H. Hosseinzadeh-Bandbafha, K.-W. Chau, Comprehensive model of energy, environmental impacts and economic in rice milling factories by coupling adaptive neuro-fuzzy inference system and life cycle assessment. J. Clean. Prod. **217**, 742–756 (2019)

59. N.S. Ross et al., Carbon emissions and overall sustainability assessment in eco-friendly machining of Monel-400 alloy. Sustain. Mater. Technol. **37**, e00675 (2023). https://doi.org/10.1016/j.susmat.2023.e00675

60. R. Yedida, S. Saha, Beginning with machine learning: a comprehensive primer. Eur. Phys. J. Spec. Topics **230**(10), 2363–2444 (2021). https://doi.org/10.1140/epjs/s11734-021-00209-7

61. M. T. Ribeiro, S. Singh, and C. Guestrin, Why should i trust you?" Explaining the predictions of any classifier," In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 2016, 1135–1144.

62. S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems,* vol. 30, 2017.

63. J. O. de Jesus, K. Oliveira-Esquerre, and D. L. Medeiros, 2021 Integration of Artificial Intelligence and Life Cycle Assessment Methods, In IOP Conference Series: Materials Science and Engineering 2021, IOP Publishing, 1196, 012028.

64. A. Ghoroghi, Y. Rezgui, I. Petri, T. Beach, Advances in application of machine learning to life cycle assessment: a literature review. Int. J. Life Cycle Assess. **27**(3), 433–456 (2022). https://doi.org/10.1007/s11367-022-02030-3

65. N. N. Barros and R. C. Ruschel, Machine learning for whole-building life cycle assessment: A systematic literature review, In Proceedings of the 18th International Conference on Computing in Civil and Building Engineering: ICCCBE 2020, 2021: Springer, 109–122

66. T. Hong, Z. Wang, X. Luo, W. Zhang, State-of-the-art on research and applications of machine learning in the building life cycle. Energy and Buildings **212**, 109831 (2020)

67. A. Kaab, M. Sharifi, H. Mobli, A. Nabavi-Pelesaraei, K.-W. Chau, Combined life cycle assessment and artificial intelligence for prediction of output energy and environmental impacts of sugarcane production. Sci. Total Environ. **664**, 1005–1019 (2019)

68. B. Khoshnevisan, S. Rafiee, M. Omid, H. Mousazadeh, S. Clark, Environmental impact assessment of tomato and cucumber cultivation in greenhouses using life cycle assessment and adaptive neuro-fuzzy inference system. J. Clean. Prod. **73**, 183–192 (2014)

69. B. Khoshnevisan et al., Evaluation of traditional and consolidated rice farms in Guilan Province, Iran, using life cycle assessment and fuzzy modeling. Sci. Total Environ. **481**, 242–251 (2014)

70. A. Nabavi-Pelesaraei, S. Rafiee, S.S. Mohtasebi, H. Hosseinzadeh-Bandbafha, K.-W. Chau, Integration of artificial intelligence methods and life cycle assessment to predict energy output and environmental impacts of paddy production. Sci. Total Environ. **631**, 1279–1294 (2018)

71. S. Gounder, A. Hasan, A. Shrestha, A. Elmualim, Barriers to the use of sustainable materials in Australian building projects. Eng. Constr. Archit. Manag. **30**(1), 189–209 (2023)

72. J. Kleinekorte, J. Kleppich, L. Fleitmann, V. Beckert, L. Blodau, A. Bardow, APPROPRIATE life cycle assessment: a PRO cess-specific, PR edictive I mpact A ssessmen T method for emerging chemical processes. ACS Sustain. Chem. Eng. **11**(25), 9309–9319 (2023)

73. V.L. Deringer, A.P. Bartók, N. Bernstein, D.M. Wilkins, M. Ceriotti, G. Csányi, Gaussian process regression for materials and molecules. Chem. Rev. **121**(16), 10073–10141 (2021)

74. J. Kleinekorte, L. Kröger, K. Leonhard, A. Bardow, A Neural Network-Based Framework to Predict Process-Specific Environmental Impacts, in *Computer Aided Chemical Engineering*. ed. by A.A. Kiss, E. Zondervan, R. Lakerveld, L. Özkan (Elsevier, Amsterdam, 2019)

75. Y. Ma, L. Wu, Y. Guan, Z. Peng, The capacity estimation and cycle life prediction of lithium-ion batteries using a new broad extreme learning machine approach. J. Power Sources **476**, 228581 (2020)

76. B. Zhao, C. Shuai, S. Qu, M. Xu, Using deep learning to fill data gaps in environmental footprint accounting. Environ. Sci. Technol. **56**(16), 11897–11906 (2022)

77. X. Zhu, C.-H. Ho, X. Wang, Application of life cycle assessment and machine learning for high-throughput screening of green chemical substitutes. ACS Sustain. Chem. Eng. **8**(30), 11141–11151 (2020)

78. S.H. Mousavi-Avval, S. Rafiee, M. Sharifi, S. Hosseinpour, A. Shah, Combined application of life cycle assessment and adaptive neuro-fuzzy inference system for modeling energy and environmental emissions of oilseed production. Renew. Sustain. Energy Rev. **78**, 807–820 (2017)

79. A. Karim, A. Mishra, M.H. Newton, A. Sattar, Efficient toxicity prediction via simple features using shallow neural networks and decision trees. ACS Omega **4**(1), 1874–1888 (2019)

80. X. X. Romeiko, Z. Guo, and Y. Pang, Comparison of support vector machine and gradient boosting regression tree for predicting spatially explicit life cycle global warming and eutrophication impacts: A case study in corn production, In 2019 IEEE International Conference on Big Data (Big Data), 2019. IEEE, 3277–3284

81. E.K. Lee et al., Projecting life-cycle environmental impacts of corn production in the U.S. Midwest under future climate scenarios using a machine learning approach. Sci. Total Environ. **714**, 136697 (2020). https://doi.org/10.1016/j.scitotenv.2020.136697

82. A. Marvuglia, M. Leuenberger, M. Kanevski, E. Benetto, Random Forest for toxicity of chemical emissions: features selection and uncertainty quantification. J. Environ. Account. Manag. **3**(3), 229–241 (2015)

83. M. Budihardjo, I. Priyambada, A. Chegenizadeh, S. Al Qadar, and A. Puspita, "Environmental impact technology for life cycle assessment in municipal solid waste management," Global Journal of Environmental Science and Management, 2023.

84. L. Xu, F. Wu, R. Chen, L. Li, Data-driven-aided strategies in battery lifecycle management: prediction, monitoring, and optimization. Energy Storage Mater. (2023). https://doi.org/10.1016/j.ensm.2023.102785

85. E. Bolandnazar, A. Rohani, M. Taki, Energy consumption forecasting in agriculture by artificial intelligence and mathematical models. Energy Sources Part A: Recovery, Utilization, and Environmental Effects **42**(32), 1618–1632 (2020). https://doi.org/10.1080/15567036.2019.1604872

86. D. Yılmaz, A.M. Tanyer, İD. Toker, A data-driven energy performance gap prediction model using machine learning,". Renew. Sustain. Energy Rev. **181**, 113318 (2023)

87. I.I. Baskin, V.A. Palyulin, N.S. Zefirov, Neural networks in building QSAR models, in *Artificial Neural Networks: Methods and Applications*. ed. by D.J. Livingstone (Humana Press, Totowa, 2009)

88. A. Marvuglia, M. Kanevski, E. Benetto, Machine learning for toxicity characterization of organic chemical emissions using USEtox database: Learning the structure of the input space. Environ. Int. **83**, 72–85 (2015)

89. G. Wernet, S. Hellweg, U. Fischer, S. Papadokonstantakis, K. Hungerbühler, Molecular-structure-based models of chemical inventories using neural networks. Environ. Sci. Technol. **42**(17), 6717–6722 (2008)

90. R. Song, *Machine learning for addressing data deficiencies in life cycle assessment* (University of California, Santa Barbara, 2019)

91. C. Isarankura-Na-Ayudhya, T. Naenna, C. Nantasenamat, and V. Prachayasittikul, "A practical overview of quantitative structure-activity relationship," 2009.

92. B.-A. Tuulaikhuu, H. Guasch, E. García-Berthou, Examining predictors of chemical toxicity in freshwater fish using the random forest technique. Environ. Sci. Pollut. Res. **24**(11), 10172–10181 (2017). https://doi.org/10.1007/s11356-017-8667-4

93. R. Shoji, 2005 The potential performance of artificial neural networks in QSTRs for predicting ecotoxicity of environmental pollutants. Curr. Comput.-Aided Drug Design **1**(1), 65–72 (2005)

94. T.H. Miller et al., *Machine learning for environmental toxicology: a call for integration and innovation* (ACS Publications, Washington, D.C., 2018)

95. M. Slapnik, D. Istenič, M. Pintar, A. Udovč, Extending life cycle assessment normalization factors and use of machine learning–a Slovenian case study. Ecol. Ind. **50**, 161–172 (2015)

96. P. Ahmadi, M. Raeesi, S. Changizian, A. Teimouri, A. Khoshnevisan, Life-cycle assessment of diesel, diesel-electric and hydrogen fuel cell transit buses with fuel cell degradation and battery aging using machine learning techniques. Energy **259**, 125003 (2022)

97. F. Meng, C. LaFleur, A. Wijesinghe, J. Colvin, Data-driven approach to fill in data gaps for life cycle inventory of dual fuel technology. Fuel **246**, 187–195 (2019)

98. X. Hu, A.K. An, S.S. Chopra, Life Cycle Assessment of the Polyvinylidene Fluoride Polymer with Applications in Various Emerging Technologies. ACS Sustain. Chem. Eng. **10**(18), 5708–5718 (2022)