

Perspective

Open Access



Enhancing big data for greentelligence across the production value chain

Reuben Seyram Komla Agbozo¹, Tao Peng^{1,2} , Huajun Cao³, Renzhong Tang^{1,2}

¹Institute of Industrial Engineering, School of Mechanical Engineering, Zhejiang University, Hangzhou 310027, Zhejiang, China.

²State Key Laboratory of Fluid Power & Mechatronic Systems, Zhejiang University, Hangzhou 310027, Zhejiang, China.

³State Key Laboratory of Mechanical Transmissions, Chongqing University, Chongqing 400044, China.

Correspondence to: Dr. Tao Peng, Institute of Industrial Engineering, School of Mechanical Engineering, Zhejiang University, 38 Zheda Road, Hangzhou 310027, Zhejiang, China. E-mail: tao_peng@zju.edu.cn

How to cite this article: Agbozo RSK, Peng T, Cao H, Tang R. Enhancing big data for greentelligence across the production value chain. *Green Manuf. Open* 2023;1:4. <https://dx.doi.org/10.20517/gmo.2022.02>

Received: 30 Jun 2022 **First Decision:** 2 Aug 2022 **Revised:** 25 Aug 2022 **Accepted:** 15 Sep 2022 **Published:** 23 Sep 2022

Academic Editor: Hongchao Zhang **Copy Editor:** Jia-Xin Zhang **Production Editor:** Jia-Xin Zhang

Abstract

The big data concept has been explosive, revealing, and transformative across manufacturing industries as it provides deeper insights into manufacturing operations for decision-making. However, big green data (BGA), a dedicated subset of big data, is not adequately structured for comprehensive sustainability analysis, particularly in smart factories. With our proposed green data balance (GDB), there will be accountability for each input and output composition in a production unit within the production value chain (PVC). Data will be exhaustively and accurately collected in each workshop to help uncover unknown issues in a production value chain while facilitating the development of sustainability metrics or index systems. Additionally, a structured big green data system will fuel "greentelligence", using intelligent systems and technologies to speed up digitalization toward sustainable manufacturing by measuring, tracking, and minimizing adverse environmental impacts. Lastly, with the support of the cognitive intelligence data analytic system (CIDAS), real-time and near real-time comprehensive sustainability analytics can be performed, leading to Self-X metacognitive adjustments and corrective actions.

Keywords: Big green data, data balance, Self-X, cognitive intelligence, smart factory

INTRODUCTION

Going green is the next global competition, the next revolution across industries. In manufacturing, going green means implementing green engineering practices into the production value chain to promote



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



sustainability. These practices include green engineering design, the use of pollution-reducing processes and products, holistic systems analysis, and the integration of environmental impact assessment methods and environmentally benign materials and energy inputs and outputs, as well as other eco-friendly and efficient principles^[1]. As manufacturing becomes smarter, industries progressively use sensors to capture data across a product's life cycle^[2]. This progression has profoundly impacted how products are manufactured, particularly with sensor-laden facilities continuously monitoring and gathering data across a product's life cycle, providing real-time and near real-time insights to understand their operations better. With the help of the Internet of Things (IoT) and sensor integration technologies, data on manufacturing operations such as pressure and temperature are continuously measured, monitored, and reported through manufacturing information systems^[3]. These datasets come with different attributes, i.e., the volume of data, the velocity of data streams, the variety of data types, their veracity, and the value of the data, thus making the data complex and challenging^[3,4]. The aggregation of these datasets (with variations in parameters) into more extensive and complex forms is described as big data. Besides the challenges, big data analysis has created opportunities for manufacturers of all types, from data-driven innovation to improving products, from enhancing real-time visibility to reducing costs, and from utilizing data to better manufacturing operations while providing actionable insights.

Presently, the big data concept has been explosive, revealing, and transformative across manufacturing industries as it contributes to revolutionizing manufacturing. Across a product's life cycle, big data analytics helps manufacturers streamline their objectives by optimizing resources, reducing carbon emissions, and emphasizing all critical sustainability factors^[5]. Fortunately, smart manufacturing operations generate big data to provide visibility and scientific and statistical evidence to help accelerate green manufacturing objectives. For instance, Bevilacqua *et al.*, in partnership with an Italian manufacturing company, developed an IoT-based energy management framework that uses big data analytics to improve energy efficiency in production systems^[6]. Their framework facilitates the discovery of machine energy trends with strategies to improve energy-efficient decision-making in a production system with a multi-layer model: data collecting, data management, and data analytics layers. Therefore, as production digitalization increases across manufacturing industries, so does big data, allowing for a more in-depth analysis of processes and operations inside a manufacturing system.

Research gaps in big green data for greentelligence

Big data powers "greentelligence". Greentelligence is the use of intelligent systems and technologies to accelerate digitalization towards sustainable manufacturing. For greentelligence to be purely effective and efficient, it must run on big green data - large datasets that hold environmentally sensitive information for monitoring, pattern identification, analysis, and optimization. Consequently, greentelligent manufacturing operations can deliver greener products by reconfiguring industrial processes and product designs, reducing waste and emissions, minimizing natural resource use, and improving energy and manufacturing efficiency^[7].

However, the nexus of green manufacturing and big data reveals some vulnerabilities. First, the big green data (BGD) characteristics in production are not explicit. Although some research has attempted to link big data and green manufacturing, there are limited standardized metrics or index systems for measuring big green data on production units, mainly for smart manufacturing. Metrics are relevant, especially in this age of data revolution. Thus, establishing standardized sustainability metrics contributes significantly to driving green-targeted goals in manufacturing. Standardized sustainability metrics (SSM) are benchmarks for measuring and tracking the performance of sustainability-related factors (such as energy, materials, greenhouse gases (GHG), and other forms of pollution) on the environment. SSM can assist in making sustainability-informed decisions such as fine-tuning operations towards greener goals, developing efficient

sustainability impact assessment methods, and even elevating the company's image and identity. Nevertheless, the absence of SSM limits exposure to identifying varying potential impacts of pollution on the environment, which could otherwise help streamline green-related operations. It is, therefore, prudent to gather and assess green data on the factory floor by considering the processes, production lines, methods, or techniques, as well as any activity that may have an adverse effect on the environment. In light of this, performing multi-perspective analytics on each operation within a defined area can deliver comprehensive information for innovative decision-making toward zero-waste manufacturing^[8]. That is, tracking manufacturing operations can identify waste spots and recommend effective and efficient solutions.

Second, manufacturing is gearing towards the next paradigm, the use of cognitive systems and technologies. For instance, IBM is exploring cognitive manufacturing by unifying millions of data points from systems to generate actionable insights across the whole value chain, from product design to operations and customer service^[9]. Besides, Zheng *et al.* presented a Self-X cognitive manufacturing network, which introduces a multi-agent reinforcement learning method based on industrial knowledge graphs and other related works^[10]. Generally, researchers are investigating the use of cognitive technologies across a product's life cycle in various industries. Hence, this paper proposes adopting and implementing cognitive intelligence-enabled systems to facilitate green manufacturing objectives. The recent advancements and publications on the conscious Turing machine^[11], cognitive computing^[12], cognitive manufacturing^[13], cognitive intelligence-enabled manufacturing systems^[14], and related fields provide the opportunity to research the technology's applicability in green manufacturing.

Indeed, big data is crucial and fundamental for cognitive systems to achieve higher performance. Thus, for sustainability operations to be effective in the production value chain, precise and structured BGD is required, where data collected via sensors and IoT technologies can contribute toward developing a zero-waste factory. However, simply collecting data is not enough. Instead, a structured mechanism for collecting, calculating, and analyzing data is required to provide better insights into achieving resource optimization while minimizing waste.

Research questions

Using a data-driven approach to implement sustainability in manufacturing offers industries a better perspective on their operations. After all, manufacturing and big data are symbiotic and pivotal in advancing industrial operations. Similarly, BGD, a dedicated subset of big data, plays a critical role in reconstructing manufacturing processes in ways that support sustainability. Therefore, as academics and manufacturing industries fight for a shared cause to achieve greening in manufacturing, we pose two research questions:

- How detailed should big green data be in the production value chain (PVC)?
- How can cognitive intelligent systems use big green data (BGD) to tackle sustainability issues?

In the following, we propose some techniques and methodologies for exploring sustainability issues in the PVC using BGD, where the leaner, the better. Data analytics using cognitive intelligence is also discussed, particularly on how cognitive intelligent-enabled systems can efficiently enhance green manufacturing.

BIG DATA FOR GREENTELLIGENCE

Big data offers industries tremendous opportunities, as it empowers them to adopt data-driven strategies to become more competitive^[3]. With innovative connective technologies, smart factories monitor and collect

data across a product's life cycle, including manufacturing operations. This helps industries gain broader insights into making better decisions. To adapt to an increasingly complex and dynamic environment, smart factories use data and technology to monitor and improve operations, making them flexible and efficient. In fact, smart factories are not perfect; they have blind spots. However, to gain the broadest possible insight into green manufacturing operations and identify waste, smart factories must use digital technologies to monitor and collect data across the production value chain to comprehensively analyze green-driven activities [Figure 1]. As a result, they can use those data to adjust the manufacturing process, improve the design, and produce sustainable products. For instance, in manufacturing operations, aggregating every measurable green data per production unit on the shop floor can provide extensive information on resource usage, waste generation, and GHG emissions through BGD analytics. Production units in this paper refer to a process, method, or technique used to manufacture a product (part or whole) within a defined area.

Generally, big data analytics influences green manufacturing strategies by extracting meaningful information from data, whereas green manufacturing provides complete mediation between big data analytics and environmental performance^[15]. As a result, areas that prove challenging or fall behind in meeting green goals can be overhauled for the better. Hence, with the help of intelligent technologies, methods, structured green data, and big data analytics, sustainability solutions can be effectively and efficiently implemented in the production value chain (PVC). Considering this, we suggest enhancing big data within the production value chain (PVC) as a fraction of the product life cycle. Given this, we describe the production value chain as the operations between raw material reception and end-user delivery. Data generated in this phase, PVC, are mostly process, energy, material, and/or equipment related. Thus, for the purpose of streamlining procedures, fostering innovation, and facilitating sustainability-driven decision-making, these data offer greater insights into manufacturing operations.

Developing a green data balance

Manufacturing sectors are achieving total material and energy balances in their operations. Generally, material and energy balances are forms of data representation used to identify and quantify losses (energy and material) and emissions in production units. For instance, an energy balance reveals how products are transformed into one another, highlights the various relationships among these products, and allows for the visualization of the energy consumption in a defined system^[16]. Comparably, green data balance (GDB) provides a holistic flow analysis in ratio forms, input to output. Therefore, extending this concept of balancing to green manufacturing advances sustainability goals by providing detailed data structuring, leading to an in-depth understanding of the direct and indirect impacts of manufacturing operations on the environment. In addition, it can unearth new information to aid in the development of short- and long-term sustainability roadmaps. Thus, based on this concept of balancing, merged with a detailed flow analysis of green manufacturing operations, we coin the term "green data balance".

A green data balance is a balancing technique for gathering sustainability-related information in a system by accounting for and balancing the input to output data at each node of a production unit using Equation (1) [Figure 2]. The green data balance is a sustainability data gathering framework for understanding and improving environmental assessment.

$$\text{Input} = \text{System consumptions} + \text{output} \quad (1)$$

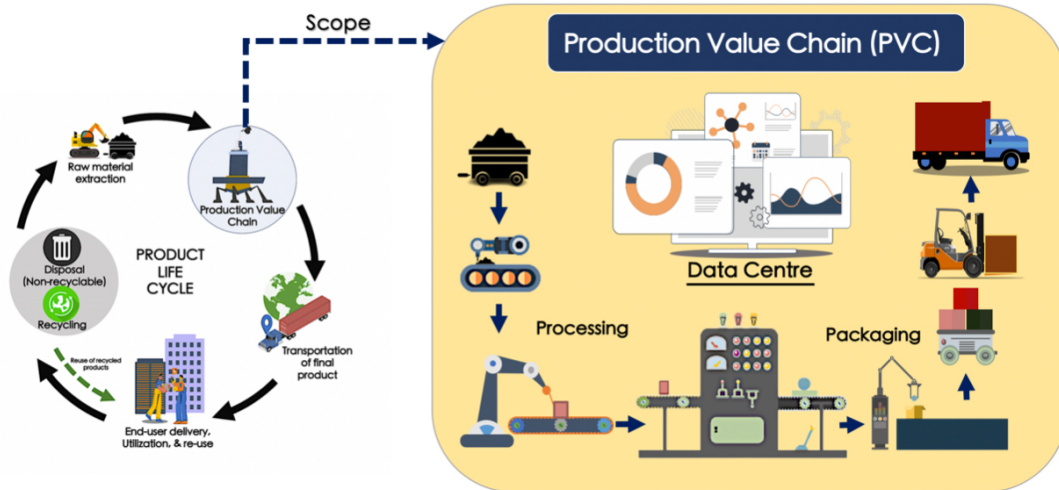


Figure 1. The scope: a production value chain, which is a fraction of the product life cycle.

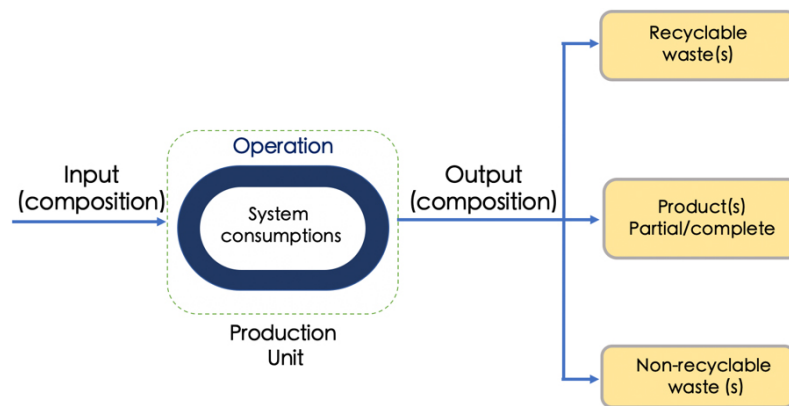


Figure 2. General view of a green data balance on a production unit.

It represents an exhaustive and accurate data flow (in forms such as material, energy, and other components) via the production unit at the input, system consumption (such as energy and material where applicable) and output levels (in the form of products and general wastes). In addition, it offers manufacturers the opportunity to gain new insights by extracting more detailed information in a production unit, revealing new green-related data where possible. As shown in Figure 2, any input in a production unit must be declared and documented on a component basis. This is followed by calculating the amount or rate of consumption in the system during operations and, finally, identifying and recording the output compositions. By obtaining and balancing green data on the unit, engineers can simultaneously perform both quantitative and qualitative studies, leading to greater knowledge acquisition at the input, process, and output levels. As a result, manufacturers will have statistical data to support their deep analytics, enabling them to gain a more comprehensive understanding of their operations and create environmentally friendly, productive solutions.

To explore this theory of GDB, we extended the working principles shown in Figures 2 and 3, based on a sample case study by Liu *et al.* on the critical life cycle inventory for the aluminum die casting workshop^[17]. Each process flow was examined in their case, along with data visualization on resources and emissions

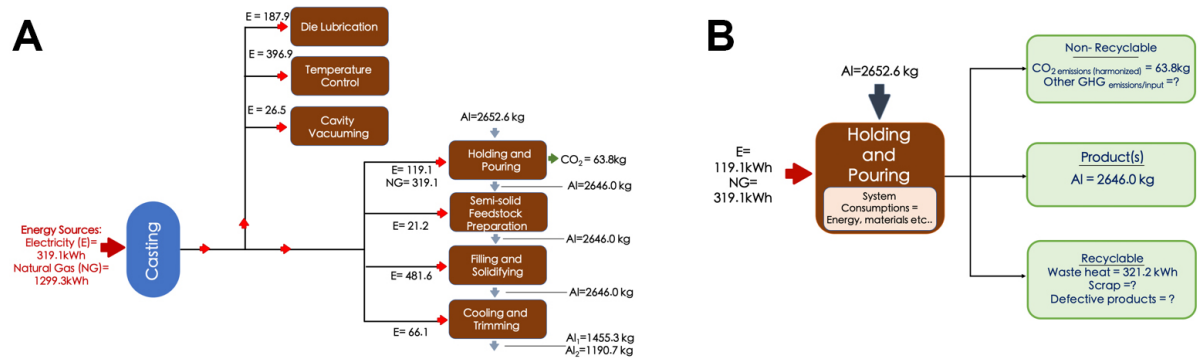


Figure 3. (A) Al casting process flow analysis. (B) Data balance approach for holding and pouring in Al casting. (Modified based on Liu et al. aluminum die casting case study^[17].)

based on a life cycle assessment approach. Although the data flow analysis was detailed and considered major green metrics, further analysis of the input-to-output ratio per unit in the PVC could yield a more comprehensive output. For instance, information on the types of equipment or technology used in their operation within a defined unit, and their corresponding energy consumption for a given timeframe could lead to determining what technologies and equipment are most suited for achieving a defined environmental target. To be more precise in our illustration, we focus on the holding and pouring unit in the casting processing stage in Figure 3 to show other variables that can be considered to balance the input variables. It is observed that each unit has information on energy input (electricity and natural gas), material inputs, and other components. However, the output data do not necessarily balance on each unit. For each energy source, we do not have data to account for all the respective emissions and the same applies to the material processing. Even though CO₂ emissions were recorded, we are unable to identify specific energy types for each category that generated the corresponding emissions, specifically other greenhouse gas emissions per production unit. With a GDB, this should be calculated and documented.

In fact, while CO₂ makes up a large fraction of greenhouse gases, the other GHG compositions are also a threat to our environment, and that is one of the benefits of adopting a green data balance, as it delves deeper into operations to help assess their direct and indirect impacts on the environment.

Furthermore, GDB can enhance life cycle assessment (LCA) and vice versa, as well as probably make LCA processes easier across the PVC. It can help resolve the issue of data uncertainty and create a readily available sustainability-related database for referencing and modeling, challenges stipulated by Curran^[18] in LCA. LCA is defined as “a tool to assess the environmental impacts and resources used throughout a product’s life cycle, i.e., from raw material acquisition, via production and use phases, to waste management”^[19]. It is a widely used methodology for assessing the environmental impacts associated with a product’s life cycle. This means sustainability-driven accountability and, in this case, a GDB for each unit at a compositional system level. Take, for example, the holding and pouring unit in Figure 3B; here, manufacturers can distinctly classify the output into product components (parts or whole) and waste (recyclable and non-recyclable and varying forms of emissions). By using the LCA approach, manufacturers can compute the greenhouse gas emissions per production unit by referencing the GHGs and materials factsheet^[20]. This includes the types of materials and the type of energy used in processing the product in each unit. During production, manufacturers can record and perform energy analysis to determine the energy consumed to produce a unit of a product, which serves as a reference for calculating the amount of GHGs per 1 kJ. In addition, for each waste stream that exits the system boundary, it should allocate the equivalent emissions. For instance, if 100 kg of material were fed into a system for processing with 90%

product output and 10% waste, we would need to allocate the emissions value for the waste on the said unit. To obtain a more holistic view of the system boundaries, manufacturers may perform simulations for different production units (using the data balance principles in [Figure 3](#)) to aid in identifying weak spots and facilitate decision-making strategies. This classification can be more thorough and balanced on the unit in terms of energy, material usage, waste, and all forms of emissions produced and other applicable treatments. For example, energy is a key resource in manufacturing; therefore, adopting this principle of green balance can push for a comprehensive energy performance analysis, considering the variation in energy consumption behavior in machine workshops^[21]. Hence, it can facilitate the energy usage tracking processes, determine the areas with higher energy consumption and which operations ignite higher consumption levels, and provide general usage statistics. In addition, it can boost innovative decision-making strategies by acquiring reliable data for energy modeling in terms of monitoring, consumption, performance, and optimization^[21,22].

Besides energy, the same principle can be applied to other green-sensitive elements. Primarily, adopting a green data balance can facilitate the development of sustainability metrics or index systems for measuring sustainability across the PVC since it can establish relationships between usage and impacts. In this context, usage refers to measuring the progress of key environmental metrics such as resource consumption against standardized indices. For instance, based on their consumption rate, where a higher consumption rate may infer more emissions depending on the type of energy source, manufacturers can perform sustainability assessments to track the impacts of these resources on the environment while developing mitigation strategies to keep manufacturing operations sustainable.

Cognitive intelligence-enabled data analytics for green manufacturing

Future computer system designers are beginning to prioritize a system's ability to self-repair, self-organize, adapt to changing workloads, and interact dynamically with other systems while anticipating user activities^[23]. The motivation is to provide computers with artificial cognitive abilities to make categorical judgments and decisions. Generally, humans can assess a situation, determine what is required to attain a goal, predict the outcome, and take the necessary actions while adapting^[24]. Using perception - the process of identifying, transforming, and interpreting sensory information - humans can create meaning for information received from their environment^[25]. These processes of sensation, perception, and identification lead to cognition, the act of perceiving and knowing^[26]. Likewise, cognitive systems require big data to learn and acquire knowledge and exhibit cognitive intelligence through Self-X features. In this paper, the term "Self-X" refers to the cognition of a system that makes it conscious of its surroundings and, as a result, causes it to make decisions to keep itself in check. Some Self-X attributes include self-monitoring, self-adaptation, self-diagnosis, self-configuration, and many others. An application of selected self-X attributes was tested by Zheng *et al.* by performing a visual reasoning-based approach for mutual-cognitive human-robot collaboration using advanced cognitive computing^[27]. Additionally, Park and Tran adopted cognitive abilities such as perception, reasoning, and cooperation in developing an intelligent agent-based manufacturing system on the manufacturing shop floor; where the system experiences disturbances, the system self-adjusts^[28]. Therefore, in a manufacturing environment, machines with embedded cognitive features can acquire green knowledge through green data acquisition with the help of machine perceptions. Machine perception is the ability of a computer system to understand data in a way that is similar to how humans interact with the world around them using their senses^[29]. By using machine senses (sensor networks including cameras, thermometers, microphones, electrodes, electronic noses, *etc.*), machines gather different formats of data. These large datasets form a knowledge layer that aids in tackling real-time complex problems.

Cognitive intelligence-enabled systems are knowledgeable systems that mimic human cognitive abilities to solve complex problems in a defined environment^[14]. Using Self-X cognitive abilities, cognitive systems optimize multi-modal data resources (numeric, texts, images, or video) to self-configure, -optimize, and -adapt^[24]. Supported by enabling technologies such as machine learning, deep learning, knowledge graphs, cognitive computing, reinforcement learning, and others, the cognitive system can self-learn and -unlearn. An example is AlphaGo Zero, a self-taught piece of computer software by Google's DeepMind group; it uses reinforcement learning to master the game of Go^[30]. By playing against itself, it learns the patterns and errors and becomes its own tutor. At a cognitive manufacturing network level, Zheng *et al.* explored an industrial knowledge graph-based multi-agent reinforcement learning approach that utilizes high-level Self-X capabilities, such as self-configuration, -optimization, and -adjustment^[10]. The study was supported by an illustrative example of a multi-robot reaching task to validate the proposed approach. In addition, Leng *et al.* proposed a loosely coupled deep reinforcement learning approach to precisely forecast the cost, makespan, and carbon footprint of orders for printed circuit boards and maximize material consumption across the production process^[31].

These enabling technologies can utilize algorithms that mimic metacognitive abilities such as human cognition. Metacognition is described as cognition about cognition: assessing and reflecting on one's thought processes to help fine-tune other cognitive processes, safeguarding against errors and confusion, and data management to increase efficiencies^[32]. Self-monitoring, -planning, and -evaluation are examples of metacognitive skills that can assist a system in reflecting on its performance. Therefore, the metacognitive task may involve explaining errors in the cognitive tasks or selecting between cognitive algorithms to carry out reasoning functions^[33]. Machines with these capabilities can detect problems early, self-adjust, and take corrective action^[14].

A cognition model is shown in Figure 4. The cognitive intelligence data analytics system (CIDAS) begins with operation monitoring, sensation, and real-time/near real-time data gathering (where applicable) using machine perceptive technologies - machine perception. In this paper, real-time data (RTD) refers to data that are collected instantly and sent to receiving systems on time and without delay, whereas near real-time data (NRTD) indicates data that are gathered before, during, and after operations and enter into the system's live database (requiring a live Internet connection) with a defined latency while maintaining data current online. Subsequently, data are accumulated and translated into machine processing formats, resulting in knowledge acquisition and storage within the cognitive data processor - knowledge acquisition and encoding. Then, the system performs a self-greening analysis, a form of semantic reasoning, by comparing real-time data to SSM standards (standard operational reference metrics) to derive meaning from the analytics. Based on the system response, as indicated in Equation (2), the system may notify engineers of the outcomes for authorization to self-adjust where necessary. Generally, a cognitive intelligent system is not necessarily responsible for making the final decisions; instead, it supplements information on the fly for engineers to make the necessary decisions^[14].

In the case of CIDAS, its main role is to perform cognitive analytics; thereafter, when and where necessary, it may self-adjust by exhibiting Self-X characteristics such as self-optimization or -configuration. Analytics from CIDAS may warrant diagnostics or prognostics to keep green manufacturing objectives in check by displaying statistical evidence in the form of data visualization or other data communicative formats. Acting as a machine expert, a cognitive system can use big data analytics to investigate the root causes of problems and give recommendations for decision-making and implementation by using historical and real-time data.

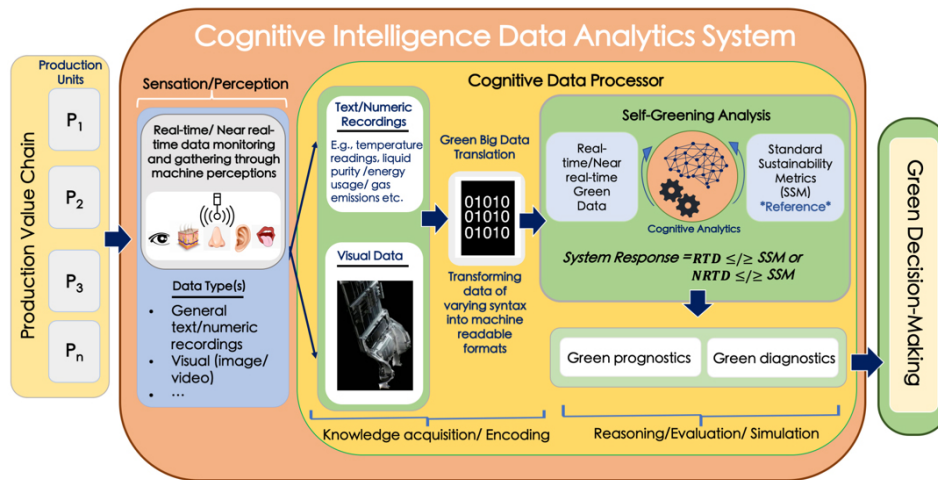


Figure 4. Cognitive intelligence data analytics system for big green data analytics. RTD: Real-time data; NRTD: near real-time data; SSM: standardized sustainability metrics.

On matters of sustainability, cognitive data analytics can be more detailed by adopting the green data balance principles and data analytics. By comparing input data to output data, CIDAS can perform comparative analytics and identify “green” weak spots. Green data balance helps in data structuring, thus contributing to analytics and reasoning for deeper analytics. As a result, manufacturers get foreground knowledge on resource usage (such as energy and material usage) and waste streams. It can also identify new forms of waste that were overlooked in previous processes. After all, the objectives of green manufacturing are to incorporate environmental performance indicators into PVC to prevent resource depletion or replace raw materials with eco-friendly alternatives, maximize energy efficiency, minimize the use of hazardous materials, and reduce emissions and other pollutants as well as the effects they have on the environment. With CIDAS, achieving these objectives is more feasible since data collection per production is very detailed. As a result, BGD analytics becomes very comprehensive and provides up-to-date information for green decision-making.

In addition, other relevant operations, such as waste classification, as depicted in Figure 3, can be performed. Hence, providing detailed analytics on the amount and types of waste generated and how they are generated can contribute to the development of sustainability metrics or index systems. Moreover, innovative recommendations on energy optimization and waste minimization can be discussed and implemented to boost manufacturing operations toward a zero-waste factory. Prognostics are another attribute, where anticipated deviations from the normal operating range that can negatively impact the environment are reported to the engineering team for a timely assessment. This can include issues such as emission prevention and reduction and safety problems. With their metacognitive features, the system is self-aware of deviations from standard green manufacturing processes and adjusts where necessary. In situations where deviations are perceived in advance based on pattern identification, engineers may perform prognostics to avoid any form of environmental degradation.

On the issue of ensuring green manufacturing practices, the system response in CIDAS serves as the prompt within the cognitive data processor (CDP) [Equation (2)]. Thus, in the event that real-time or near real-time operations data do not correspond to the anticipated standard sustainability metric (SSM), notifications will be issued to operators prior to self-adjustment. Another technology that can play a vital role in the effectiveness of CIDAS is the cognitive digital twin (CDT). To obtain a deeper understanding of

manufacturing operations and enhance decision-making, a cognitive digital twin can provide real-time simulation through a series of semantically connected digital models representing various stages of the physical system with specific cognitive abilities and assistance to carry out autonomous actions^[34,35]. Similar to conventional simulation software, cognitive digital twins can assist companies in comprehending the effects of decisions as well as the relationships between various factors, such as the effects of changing one variable on another. By relying on data from the physical manufacturing system, a CDT bridges the gap between digital and physical manufacturing worlds by creating digital replicas that can be used to simulate real-time behavior and respond to changes in a similar fashion to real-time manufacturing. Consequently, by gaining insights from the real system, manufacturers can understand and improve the performance of physical systems through effective decision-making processes. Other types of simulations can be carried out by testing various scenarios and using online data to create models that imitate the behavior of a current or proposed system. These models can then be used to support decision-making. The systems response equation can also be used for operation evaluation to identify deviations, hence enabling in-depth analysis and interpretation to be improved.

$$\text{System Response: } \text{RTD/NRTD} \leq \text{SSM} \text{ or } \text{RTD/NRTD} \geq \text{SSM} \quad (2)$$

PROOF OF CONCEPT, CIDAS—AN ILLUSTRATION

In a regular manufacturing operation, as shown in [Figure 5](#), machine perceptions use sensor networks to collect data across the PVC. Data gathered across the PVC makes up the big data - the perception stage (self-awareness). Big data at this stage includes different categories of data, specifically sustainability-related data (big green data), in the form of texts, images, videos, and sound, depending on the industry. With the help of enabling technologies, raw data formats that reach the cognitive processor are translated into machine-readable formats for analysis. By using real-time or near real-time data and standard sustainability metrics, CIDAS performs big green data analytics - the evaluation stage. Consequently, deviations in the real-time data are detected and reported for interactive assessment and decision-making at the decision-making stage.

Sustainability deviations detected in a production unit include temperature, greenhouse gases, energy, and material waste. Using the holding and pouring unit in [Figure 3B](#) as an illustrative case study, green-related data in and out of the production unit are monitored and recorded. With the help of machine perceptive technologies, big data is accumulated as described above for the perception stage. Data gathering is done in two categories (real-time and near real-time collection) where applicable. For instance, real-time data are collected by a network of sensors consisting of critical production data on the temperature within the holding and pouring unit (such as the equipment and/or machine or operations temperature data). With the help of IoT technologies, the temperature data gathered are continuously processed and analyzed. As a result, manufacturers gain insights about deviations or problems during cognitive analytics for rectification.

In the same format, near real-time data are gathered but at a delayed pace and enter into the local network and later into the online database. Near real-time data cover indicators where the real-time collection may be unlikely, such as energy consumption, GHG, and waste heat. For instance, IoT technologies can interact with system applications to track the energy consumption rate of machines or equipment and then record the data. Additionally, utility meters, such as power or electric meters, can be used to obtain data in close to real time. Data acquired can be transmitted via the local area network and updated over the Internet with minimal latency.

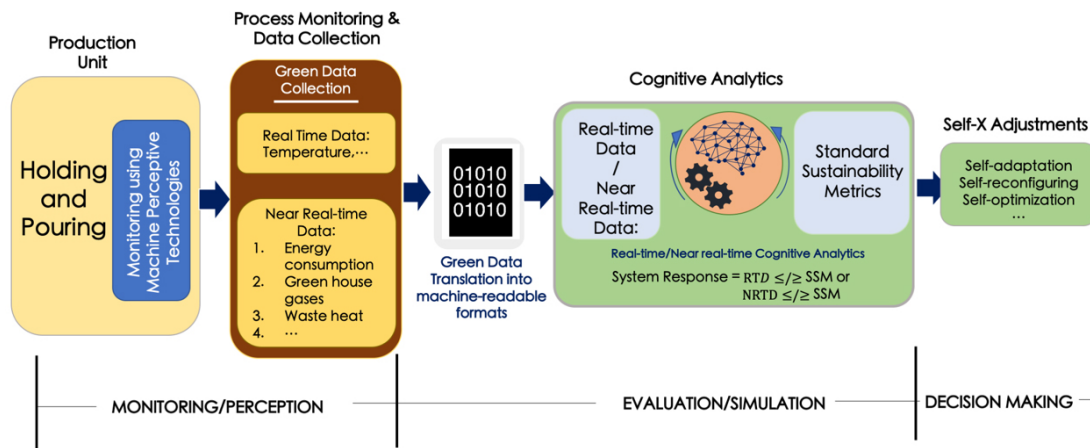


Figure 5. Cognitive intelligence analytics for the holding and pouring unit. RTD: Real-time data; NRTD: near real-time data; SSM: standardized sustainability metrics.

Regarding other green indicators such as GHG emissions, calculations based on referencing the GHGs and materials factsheet describe the emissions per production unit. By capitalizing on continuous monitoring of all connected systems and machinery within the production unit, manufacturers can obtain accurate, detailed, and up-to-the-minute data on energy consumption from every production unit. Subsequently, CIDAS performs cognitive analytics to illustrate how energy is consumed and when and where there are deviations; engineers can take evidence-based action to keep operations in check by identifying waste-generating spots and developing mechanisms to control them.

In addition, where necessary, the machine may self-adjust to maintain standard green operation. These indicators can have their data monitored and recorded using machine perceptive technologies. However, the data stream may be delayed and will have a defined near real-time latency. Regardless, given the available big green data from the holding and pouring, CIDAS performs cognitive analytics by relying on the system response functions. Based on the outcome of each analytics process, the system may perform Self-X adjustments (including self-configuration, -optimization, and -adaptation where necessary) or provide engineers the information for corrective measures, i.e., prognostics or diagnostics.

CONCLUSION

This paper explores two concepts: green data balance (GDB) and cognitive intelligence data analytic system (CIDAS) for analyzing big green data (BGD). With the digitalization of manufacturing comes the rise of big data, of which BGD is a subset. However, there are limited data structures for collecting sustainability-related data for analysis in detail per production unit, as described in Figures 2 and 3. Therefore, using methodologies such as data balance can contribute to facilitating the monitoring of input compositions on a production unit while accounting for the output on the same unit. This process will not only increase the level of accountability in each job shop but can also help unearth unknown issues in a production value chain while facilitating the development of sustainability metrics or index systems for measuring sustainability across the production value chain. Although green data balance may be challenging, it can contribute to developing clearer green data structures, data collection, and reporting and interpretation, which can lead to cleaner production. Most importantly, the principle of green data balance is applicable not only to the production value chain but to other fields that seek to achieve greener goals. In addition, the principle of green data balance can be added to educational curriculums to initiate green reasoning among future generations. The conversation on how detailed green data generation and analytics must happen

now, as our future depends on it.

CIDAS can enhance sustainability efforts by giving companies a competitive advantage over their industry peers. It can also contribute to the development of innovative solutions or dynamic roadmaps for future sustainability issues, such as facilitating the development of sustainability metrics or index systems. Cognitive intelligence-enabled systems are the future of computing, and thus employing such technologies in manufacturing can contribute to digital innovation solutions, strategic reasoning, and planning. Most importantly, it can contribute significantly to solving sustainability issues in manufacturing, as shown in [Figures 4](#) and [5](#). Hopefully, CIDAS can help incorporate environmental performance indicators into PVC to prevent resource depletion or replace raw materials with eco-friendly alternatives, maximize energy efficiency, minimize the use of hazardous materials, and reduce emissions and other pollutants, as well as the effects they have on the environment.

Going forward, we hope to experiment with data balance as a case study and explore other green-related fields, such as green graph modeling and graph-enabled reasoning.

DECLARATIONS

Authors' contributions

Conceptualization: Agbozo RSK, Peng T

Formal analysis: Agbozo RSK

Methodology: Peng T, Cao H

Validation: Peng T

Discussion: Cao H

Writing-original draft: Agbozo RSK, Peng T

Writing-review & editing: Agbozo RSK, Peng T, Cao H

Visualization: Agbozo RSK

Supervision: Peng T, Tang R

Resources: Tang R

Funding acquisition: Tang R

Availability of data and materials

Not applicable.

Financial support and sponsorship

None.

Conflicts of interest

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

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2022.

REFERENCES

1. US EPA. About green engineering. Available from: <https://www.epa.gov/green-engineering/about-green-engineering> [Last accessed on 20 Sep 2022].
2. Kusiak A. Smart manufacturing must embrace big data. *Nature* 2017;544:23-5. DOI PubMed
3. Tao F, Qi Q, Liu A, Kusiak A. Data-driven smart manufacturing. *J Man Syst* 2018;48:157-69. DOI
4. Chen M, Mao S, Liu Y. Big Data: a survey. *Mobile Netw Appl* 2014;19:171-209. DOI
5. Ren S, Zhang Y, Liu Y, Sakao T, Huisingh D, Almeida CM. A comprehensive review of big data analytics throughout product lifecycle to support sustainable smart manufacturing: a framework, challenges and future research directions. *J Cle Prod* 2019;210:1343-65. DOI
6. Bevilacqua M, Ciarapica FE, Diamantini C, et al. Big data analytics methodologies applied at energy management in industrial sector: a case study. *RFT* 2017;8:105-22. DOI
7. Li X, Wang B, Peng T, Xu X. Greentelligence: smart manufacturing for a greener future. *Chin J Mech Eng* 2021;34. DOI
8. Singh S, Ramakrishna S, Gupta MK. Towards zero waste manufacturing: a multidisciplinary review. *J Cle Prod* 2017;168:1230-43. DOI
9. Be ready for Industry 4.0 with cognitive manufacturing. Available from: <https://www.ibm.com/downloads/cas/JD71Q7RK> [Last accessed on 20 Sep 2022].
10. Zheng P, Xia L, Li C, Li X, Liu B. Towards Self-X cognitive manufacturing network: an industrial knowledge graph-based multi-agent reinforcement learning approach. *J Man Syst* 2021;61:16-26. DOI
11. Blum L, Blum M. A theory of consciousness from a theoretical computer science perspective: insights from the conscious turing machine. *Proc Natl Acad Sci USA* 2022;119:e2115934119. DOI PubMed PMC
12. Hu L, Miao Y, Wu G, Hassan MM, Humar I. iRobot-Factory: an intelligent robot factory based on cognitive manufacturing and edge computing. *Future Generation Computer Systems* 2019;90:569-77. DOI
13. IBM. Cognitive manufacturing: an overview and four applications that are transforming manufacturing today. Available from: <https://www.ibm.com/downloads/cas/VDNKMWM6> [Last accessed on 20 Sep 2022].
14. Agbozo R, Zheng P, Peng T, Tang R. Towards cognitive intelligence-enabled manufacturing. In: IFIP WG 5.7 APMS; 2022. Accepted DOI
15. Belhadi A, Kamble SS, Zkik K, Cherrafi A, Touriki FE. The integrated effect of big data analytics, lean six sigma and green manufacturing on the environmental performance of manufacturing companies: the case of North Africa. *J Cle Prod* 2020;252:119903. DOI
16. Understanding and using the Energy Balance - Analysis - IEA. Available from: <https://www.iea.org/commentaries/understanding-and-using-the-energy-balance> [Last accessed on 20 Sep 2022].
17. Liu W, Peng T, Kishita Y, et al. Critical life cycle inventory for aluminum die casting: a lightweight-vehicle manufacturing enabling technology. *Appl Energy* 2021;304:117814. DOI
18. Curran MA. Strengths and Limitations of Life Cycle Assessment. In: Klöpffer W, editor. Background and Future Prospects in Life Cycle Assessment. Dordrecht: Springer Netherlands; 2014. pp. 189-206. DOI
19. Finnveden G, Hauschild MZ, Ekvall T, et al. Recent developments in life cycle assessment. *J Environ Manage* 2009;91:1-21. DOI PubMed
20. Material Resources. Center for sustainable systems. Available from: <https://css.umich.edu/publications/factsheets/material-resources> [Last accessed on 20 Sep 2022].
21. Peng T, Xu X. An interoperable energy consumption analysis system for CNC machining. *J Cle Prod* 2017;140:1828-41. DOI
22. Peng T, Xu X, Wang L. A novel energy demand modelling approach for CNC machining based on function blocks. *J Man Syst* 2014;33:196-208. DOI
23. Morris R, Tarassenko L, Kenward M. Cognitive systems: information processing meets brain science. Available from: https://books.google.com.hk/books?hl=zh-CN&lr=&id=mpSdXu8LAtYC&oi=fnd&pg=PP1&dq=Morris+R,+Tarassenko+L,+Kenward+M.+Cognitive+Systems:+Information+Processing+Meets+Brain+Science.+Academic+Press%3B+2006.&ots=ZyhQpUB3m&sig=2PdvDIVAmgmVF005YO3gyk8JDnc&redir_esc=y#v=onepage&q&f=false [Last accessed on 20 Sep 2022].
24. Vernon D. Artificial cognitive systems: a primer. Available from: https://books.google.com.hk/books?hl=zh-CN&lr=&id=zzMqBQAAQBAJ&oi=fnd&pg=PR9&dq=Vernon+D.+Artificial+Cognitive+Systems:+A+primer.+MIT+Press%3B+2014.&ots=U29qp8a6u&sig=JB7ji0GgN8FoO-azSqFna0SbBng&redir_esc=y#v=onepage&q=Vernon%20D.%20Artificial%20Cognitive%20Systems%3A%20A%20primer.%20MIT%20Press%3B%202014.&f=false [Last accessed on 20 Sep 2022].
25. Levitin D. Foundations of Cognitive Psychology: Core Readings. Available from: <https://books.google.com.hk/books?hl=zh-CN&lr=&id=fDppm088AC&oi=fnd&pg=PP1> [Last accessed on 20 Sep 2022].
26. Stillings N. Cognitive science: an introduction. Available from: <https://books.google.com.hk/books?hl=zh-CN&lr=&id=U29qp8a6u&oi=fnd&pg=PP1>

- CN&lr=&id=wCRonP7EgDkC&oi=fnd&pg=PR11&dq=Stillings+N.+Cognitive+Science:+An+Introduction.+Cambridge,+Mass.:+MIT+Press%3B+1995.&ots=BLj9MmX71A&sig=CdBVhhHKNAYJ3Odf53iBhj6p7is&redir_esc=y#v=onepage&q=Stillings%20N.%20Cognitive%20Science%3A%20An%20Introduction.%20Cambridge%2C%20Mass.%3A%20MIT%20Press%3B%201995.&f=false [Last accessed on 20 Sep 2022].
27. Zheng P, Li S, Xia L, Wang L, Nassehi A. A visual reasoning-based approach for mutual-cognitive human-robot collaboration. *CIRP Annals* 2022;71:377-80. DOI
 28. Park H, Tran N. Development of an intelligent agent based manufacturing system. Available from: <https://pdfs.semanticscholar.org/4a31/058d3d90ccea0feb4a44ca01d7ae75b9bde5.pdf> [Last accessed on 20 Sep 2022].
 29. Florida Atlantic University. Machine perceptions. Available from: <http://www.ccs.fau.edu/~hahn/mpcr/> [Last accessed on 20 Sep 2022].
 30. Silver D, Schrittwieser J, Simonyan K, et al. Mastering the game of go without human knowledge. *Nature* 2017;550:354-9. DOI PubMed
 31. Leng J, Ruan G, Song Y, et al. A loosely-coupled deep reinforcement learning approach for order acceptance decision of mass-individualized printed circuit board manufacturing in industry 4.0. *J Cle Prod* 2021;280:124405. DOI
 32. Kralik JD, Lee JH, Rosenbloom PS, et al. Metacognition for a common model of cognition. *Proc Comp Sci* 2018;145:730-9. DOI
 33. Cox MT. Metacognition in computation: a selected research review. *Art Intell* 2005;169:104-41. DOI
 34. Faruque M, Muthirayan D, Yu S, Khargonekar P. Cognitive digital twin for manufacturing systems. *IEEE* ;2021:440-5. DOI
 35. Zheng X, Lu J, Kiritsis D. The emergence of cognitive digital twin: vision, challenges and opportunities. *Int J Prod Res* ;2021:1-23. DOI