Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 120 (2023) 1498–1503

www.elsevier.com/locate/procedia

56th CIRP Conference on Manufacturing Systems, CIRP CMS '23, South Africa

Development of a product return process in the context of the Circular Economy with the help of Machine Learning

Selina Walter[a,b,*], Anja Braun[a], Louis Louw[b]

[a,*]ESB Business School, Alteburgstraße 150, 72762 Reutlingen, Germany
[b]University of Stellenbosch, Stellenbosch Central, 7602 Stellenbosch, South Africa

* Corresponding author. Tel.: +491728636507. *E-mail address:* selina.walter@student.reutlingen-university.de

Abstract

The Circular Economy aims to reintroduce the value of products back into the economic cycle at the same value chain level. While the activities of the Circular Economy are already well-defined, there exists a gap in how returned products are treated by the industry. This study aims to examine how a process should be designed to handle returned products in the context of the Circular Economy. To achieve this, a machine learning-based algorithm is used to classify data and extract relevant information throughout the product life cycle. The focus of this research is limited to land transportation systems within the Sharing Economy sector.

Keywords: Circular Economy; Machine Learning; Sharing Economy; Data Analysis; Classification

1. Introduction

According to the Circular Economy (CE) Report of the Ellen MacArthur Foundation, only 8.6 % of the global economy implemented circularity for their products and materials [1]. One reason for this small number is the embedded nature of the linear system in the economy [1]. The problem of the Linear Economy is traced back to the high level of material consumption [2], which ends up in substantial environmental impacts, such as significant energy-related emissions, air pollution, climate change, and lots more [3]. Supporting the implementation of the CE is an integral approach to reducing these unfavourable impacts on the environment since this solution aims to maintain the products, equipment, and infrastructure to improve resource productivity [4].

A vital enabler of the CE is digitisation. Digital solutions support the CE digital platforms with technologies such as Machine Learning (ML) [5]. The use of ML for implementing the CE has become increasingly popular in the last few years. ML enables an automatic recommendation of what to do with a product at the end of its life based on the CE's capabilities. However, the ML approach includes lifetime data and data-sharing challenges. Therefore, this research paper develops a design for a return process based on the CE recommendations of a ML algorithm. The ML algorithm is created for a sharing economy (SE) that tracks and publishes lifetime data to address the challenges of lifetime data and data sharing.

2. Theoretical background

2.1. Digital Technologies in the CE

Based on the results of Antikainen et al. [5], digitalisation is a massive enabler for the realisation of CE. The different technologies offer opportunities for successful implementation [5].

According to Pagoropoulos et al. [6], three different application fields of Digital Technologies exist in the CE. The

10.1016/j.procir.2023.09.201

first is "Data collection", which includes vital technologies such as Radio Frequency Identification and the Internet of Things. This builds the basis for the later use of Digital Technologies. The second field is "Data integration", with Relational Databased Management Systems and Product Lifecycle Management. The last is about "Data analysis" and includes ML, Artificial Intelligence (AI) and Big Data analytics [6].

2.2. Machine Learning in the Circular Economy

Following Pagoropoulos et al. [6], the use of AI for solving specific problems in the environmental sector gets increasingly important. AI is already used in the CE to get information about the correct way to recycle waste based on its material. Sensors and AI are reliably used to separate waste into their raw materials [7, 8]. In other research, deep learning techniques distinguish between products and allocate them to different material groups [9]. In the listed use cases, the focus is on recycling waste.

Another AI scope of application is to accomplish reverse logistics. In this context, AI solves complex location and routing problems. However, there are many application possibilities in ML logistics [10]. The focus of this paper is not on one specific activity of the CE, but it is more about the general efficiency and sustainability in logistics.

There already exist applications of analytical AI to predict the number of returned items. AI provides an optimal monetary incentive to increase the willingness of consumers to return the product after the usage phase instead of throwing it away [10].

In conclusion, many approaches already exist for specific problems within the CE that use ML and AI. The following research step focuses on what can be done to avoid waste generation and how AI can be helpful.

Besides this selective assistance of AI for implementing the CE, AI has the potential for a systematic transformation in the CE [11]. Specifically meant is the decision-making of AI regarding the different activities within the CE - share, reuse, maintain, repair, repurpose, remanufacture and recycle.

The paper does not focus on the problem of what should be done with the waste, but it focuses on generating less waste. It presents a solution for maintaining the highest value of products on the market. Thus extending the life cycle of the product and significantly reducing waste.

2.3. Sustainable reverse logistics

In the SE, it is part of the concept that the product is returned after its usage phase. That does not include further procedures with the product when it is broken or reaches the end of its life. This procedure is described in sustainable reverse logistics, which consists of the activities included in the CE.

Reverse logistics is the complete process of efficient and effective planning, implementation and control of the flow of raw materials [12].

Despite the enormous cost-saving opportunities, companies rarely consider reverse logistics of returned products as their most important "value-added" activity [13, 14]. There are already several ways to design a reverse logistics process sustainably. One option is that the returned product is collected, selected, re-processed, shredded and redistributed as recycled or remanufactured material [12]. Another reverse logistics process focuses more on the principles of the CE. Therefore, products that can be reused get a second chance in the secondary market. Refurbished products are sold again by the distributor. The assembler takes back remanufactured products, and recycled products are returned by the fabricator as raw material. Everything else ends up in the landfill [15]. The reverse logistics framework of Dev et al. [16] considers a closed-loop environment to develop its framework. In this case, the returned products are collected by a supplier responsible for recycled materials. The quantity of the returned products is calculated with an innovation theory (Bass) model. The supplier performs a typical recycling process which includes disassembly, recycling, remanufacturing and reassembly [16].

However, building on the already developed options and adjusting them is advantageous. In addition, implementing such a sustainable return process in the economy is expandable. There are two reasons why companies need to realise functions like this. The first is that companies need to see the resulting value of the process [17]. The second is about the effort of implementation.

2.4. Sharing Economy

Sharing is the innermost circle in the CE and describes the collaborative use of one product. There are various ways to do this, such as online platforms where a product can be offered or booked or companies that make the products available for short-term use. Thus, a SE is a perfect example of the successful implementation of a CE. However, the problem is not the sharing process but what happens afterwards. [18–20]

Most companies dispose of the product without using a CE activity, so sustainability is nullified. The reason for the minimal implementation of reverse logistics in this sector, as mentioned above, is that companies need to see the added value or an efficient way to do it. [21]

3. Development of the Product Return Process

This study comprises quantitative data analysis in developing a ML algorithm. A product return process is created based on the results of the data analysis. For one SE provider, a published dataset of e-bike rides is used for the ML algorithm. The company publishes the data set citibike on its homepage citibikenyc.com/system-data.

The dataset comprises 2.5 billion trips made with 16.5 thousand e-bikes in 2020. For each trip, the information, which is listed in Table 1, is stored in the dataset. First, steps were taken to cleanse the data to obtain only relevant and correct data. In this context, the following measures were carried out. The columns that refer to the user and the location data are not relevant in this context and are, therefore, not considered further. Furthermore, some dataset data types were adjusted to continue working with them. In the process, date values stored as text were converted to datetime and values saved as text were changed to floats. The column bike ID is sorted in ascending order, while the column start time is sorted to show a history per bike ID. The dataset is checked for duplicates and missing values. The duplicates were deleted, and the average values of the entire column were added for the missing values. During cleaning, it turned out that January, February and March contained erroneous, unrealistic data that doesn't fit the remaining months. This can be seen by comparing the average trip amount of the months. Between April and December, the average is around 28 thousand trips per month, compared to an average of over 1 million trips from January to March. As the cause of the significant difference in values cannot be determined and filtering out incorrect values has not yielded successful results, the decision was made to delete the three-month period entirely. This was done to prevent any potential falsification of the data evaluation in subsequent analyses.

Table 1: Information stored in the dataset

Information stored in the dataset	
Bike ID	End station name
Trip duration	Start station GPS coordinates
Start time and date.	End station GPS coordinates
Stop time and date	User type
Start station id	Gender
End station id	Year of birth
Start station name	

Additional information is generated, including the duration of standing time, trip distance, and trip amount for each e-bike. A supervised classification algorithm is then developed to predict different activity classes, including sharing, reusing, maintaining, repairing, remanufacturing, and recycling. These classes and corresponding CE recommendations are added to the dataset as a new column. As this study represents a preliminary approach, the classes are artificially assigned to the values. The values are divided into six classes based on the average time it takes for a product to break down. The data is divided into a training and test set, with the training set comprising 80% of the total dataset.

The classification models Decision Tree Classifier, the KNeighbor Classifier, SVC Classifier, the Gaussian NM Classifier, and the MLP Classifier are used. All classification models are tested with the accuracy score, where the Decision Tree obtained the best result of 1.0, which means that the data is overfitted. The MLP Classifier obtained an accuracy score of 0.99, the KNeigbors Classifier obtained 0.96, the SVC Classifier obtained 0.95, and the GaussianNB Classifier 0.88. Therefore, the Multilayer perceptron (MLP) Classifier is used for further procedure due to its best accuracy score. When running the MLP algorithm on the presented dataset, an execution time of 172.144 seconds, a memory consumption of 1324.61 MB and a CPU consumption of 80.5 % is required. To evaluate the performance of the MLP algorithm, cross-validation is performed. The dataset is divided into ten subsets, one of these sets serves as the validation set and the model is trained on the remaining sets. After multiple runs, an average accuracy of 98.8% was obtained, confirming the algorithm's performance. Based on the six CE classes, a primary process to get a CE recommendation after each trip is developed with the help of the MLP classifier. This provides real-time monitoring and, therefore, real-time advice on what should be done with the product from a CE perspective.

The following procedure describes the execution process for obtaining the results. For each bike ID, a loop is performed. By sorting the trips chronologically, it is possible to determine how many trips have already been made with the bicycle, which is added to the dataset. All the other information about this trip is stored in a new data frame. Based on this information, the previous MLP algorithm assigns this trip to the suitable class. In the second step, the given class is checked because other considerations depend on the class. Nothing special must be considered if the class share, reuse, maintain, or repair appears. If the class recycle or remanufacture appears, the bike ID will be deleted after this trip because it is no longer part of the ecosystem. Figure 1 visualises the return process based on the CE. Since the process can be added to existing SE with already-used transportation objects, the products may be in every class. Figure 1 also shows from the literature the adapted sustainable return process used in this concept.

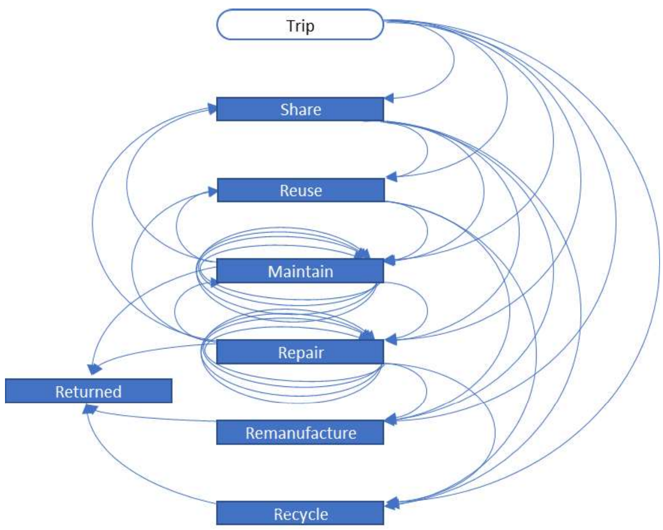

Figure 1: Product return process

In the third step, the loop looks for the second shortest start time, trip number two. The trip distance is cumulated with the trip distance of the previous trip and the standing time with the standing time of the previous trip. The trip's start time is also stored in the new dataset. The MLP algorithm is applied to the suitable class based on this information. This procedure is repeated several times until no trip of the same bike ID is left. In this case, it is moved to the following bike ID and the whole process is repeated until all trips are included in the new data frame with the information. For each class, a different process for further product usage is developed. The products' return gets more predictable based on the classes' information. The result section shows how the return process is adapted to the information that resulted through the ML Classification.

4. Results

The product return process can be supplemented with the resulting information of the ML algorithm. Therefore, the product return process is more reliable and predictable, and the enterprise's business can be adjusted, resulting in higher efficiency and a higher probability of sustainable implementation of CE.

Based on the results of the ML algorithm, the different bike IDs are added to the six classes, with their characteristics, after a year of usage. Figure 2 shows the distribution of the six activity classes after applying the ML algorithm.

The diagram shows that the significant part is repair with 45 %, followed by share with 35 % and reuse with 19 %. The classes maintain, remanufacture and recycle are significantly smaller. This is because the product is not returned to the sharing economy after the remanufacturing and recycling activities have been carried out.

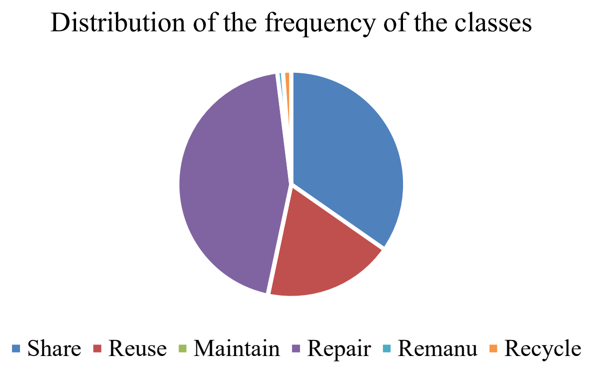

Figure 2: Distribution of the frequency of the classes

A closer look at the individual products is necessary to learn more about the cycle a means of transport goes through in a year. The following part shows how often a product passes through the respective CE activity in a year. As is already clear from the distribution, the sharing part is the second largest. Since the reuse area in connection with the SE can be classified in the same class as share, it will be included in the sharing part in the following. Therefore in 53 % of the trips in one year, sharing is the algorithm's recommendation. On average, each product gets this recommendation 47 times. Based on the Return Process, the following CE activities are intriguing: maintain, repair, remanufacture and recycle. Because with the recommendations, maintain, repair, remanufacture and recycle, the product is physically withdrawn from the SE so that the SE provider can carry out the recommendation. However, after the remanufacture and reuse activities, the product is no longer returned to the market, as already explained. Therefore, a product can receive a maximum of one remanufacturing or recycling recommendation per product. These two activities still increase the lifetime of the components of the product rather than the lifetime of the product itself. That is, the lifetime is increased based on a lower value.

The activities maintain and repair increase the product's lifetime, as the same product is fed into the economic cycle several times at the same value creation level. Accordingly, these two CE activities can occur any number of times. The evaluation of the ML Algorithm showed that between 1 and 2 maintenance is performed on a product per year. Repairs occur much more frequently, averaging 39 per product per year. The reason for this difference is that a repair is necessary because the means of transport are broken. In contrast, maintenance only improves performance and extends the life cycle, but it is not mandatory because there is no explicit fault.

Since the frequency and purposes of returning a product are known, it is significant when and for what reason it is returned. In the following, it is first shown how the usage behaviour of the means of transport changes seasonally over the entire year.

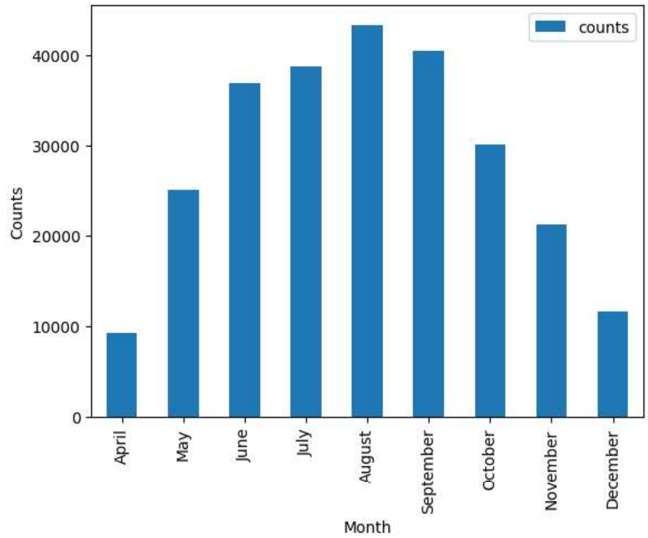

Figure 3: Annual distribution of the trips

Consequently, it can be observed in Figure 3 that the use increases in the summer months and decreases during the winter. Accordingly, it can be deduced that the returns also increase over the summer and decrease in the winter. The following graphs check whether this hypothesis is true. The charts in Figure 4 show that the already suspected number of

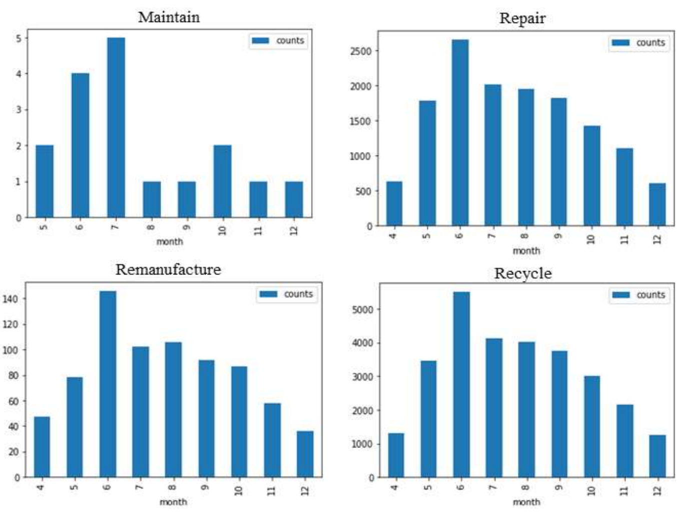

Figure 4: Annual distribution of the trips for each class

returns during summer is higher than in winter. However, it must be mentioned that in the Maintain area, the values in August dropped more significantly than in the other three activities. This, in turn, can also be traced back to the necessity of the activity, which is not guaranteed in maintenance. Based on the distribution throughout the year, the following recommendation can increase the value of the CE activities. In summer, the concentration lies on the necessary CE activities, such as repair, remanufacture, and recycle. During the winter, when there are fewer returns due to broken products, the focus is on the activity maintain. This makes the currently underappreciated action maintain more efficient. To show the second aspect, the return reason, the input values and the algorithm's results must be analysed in more detail. Because, based on the introduced values, the algorithm can only create predictions. Based on the values of standing time, driven meters and the number of trips made, the algorithm generates its results, which include the activities. Table 2 shows the Data Quality Report values of the features per activity. The values don't offer a significant difference between each activity class. Therefore, a more detailed analysis of the classes of the activities is necessary. Accordingly, the differences and the lower quartile significantly impact the classification in this algorithm.

Table 2: Data Quality Report of the activity classes

	Maintain			Repair		
Activity	Trip distance [m]	Standing time [s]	Trip amount	Trip distance [m]	Standing time [s]	Trip amount
Mean	1.660.211	14.903.310	244	1.517.545	15.355.810	240
min	141.824	4.360.598	32	888	2.895	1
25%	566.936	10.911.060	98	655.925	11.019.810	111
50%	1.107.944	13.637.040	200	1.260.942	14.752.630	220
75%	2.945.912	18.550.680	365	2.177.709	19.775.900	343
max	4.373.440	26.393.650	607	9.942.044	18.737.240	772

	Remanufacture			Recycle		
Activity	Trip distance [m]	Standing time [s]	Trip amount	Trip distance [m]	Standing time [s]	Trip amount
Mean	1585166	15436640	247	1526111	15404390	242
min	6068	559000	3	448	106	1
25%	66264	11319620	110	66650	11089720	112
50%	1293662	15049320	231	1294310	14828150	223
75%	2219879	19878610	361	2198596	19785230	349
max	9605836	28179400	702	9933284	28947660	771

To visualise this data, boxplots were used. In Figure 5, it is clear how the activities are distributed in the case of the trip distance feature. The same applies to the standing time feature, shown in Figure 6, and the trip amount in Figure 7.

A summary of why a product is sorted out lies in the context of the three different features. In this case, a more detailed

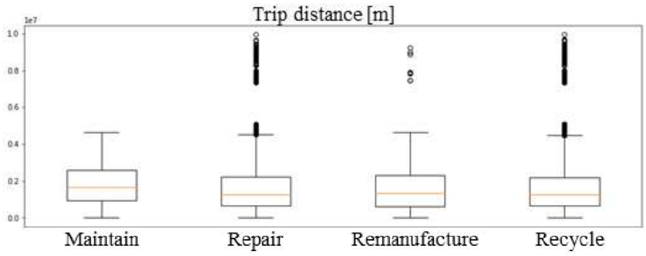

Figure 5: Visualisation of the range of the activity classes of the feature Trip Distance

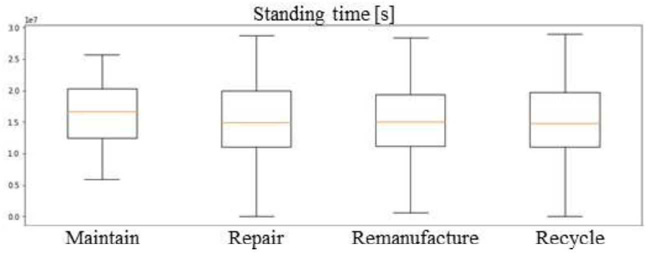

Figure 6: Visualisation of the range of the activity classes of the feature Standing time

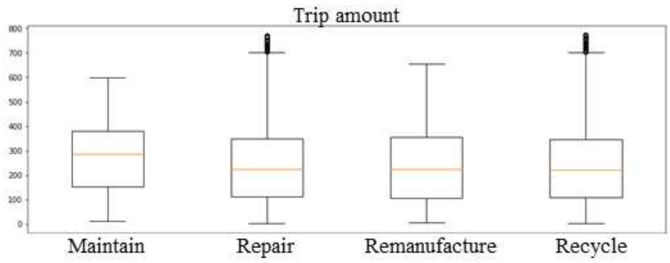

Figure 7: Visualisation of the range of the activity classes of the feature Trip amount

explanation cannot be provided, as two features, trip amount and trip distance, refer to the trips themselves, while the standing time feature refers to the time between trips. This also explains the classes that are close to each other. While a product is in use, the number of trips and the distance travelled increases, but the standing time cannot be increased. The service life is a significant factor that can also influence the product's condition due to service damage. An unused product can cause the same failures as a product used always. The difference is in the cause of the loss. In conclusion, the following points can be summarised to create a return process according to the best possible implementation of the CE. In this case, each product is returned for maintenance one to two times per year, 39 times for repair, one time for remanufacturing and one time for recycling. 0.05 % of all recommendations in one year are to do maintenance, almost half recommend repair, 0.84 % recommend remanufacture, and 1.15 % recycle. The rest is assigned to sharing. The annual distribution shows that more product returns occur over the summer months, the only exception lies in the activity maintain.

5. Conclusion and discussion

In conclusion, this is a first practice-oriented creation of a Product Return Process which aims to implement the CE in the

SE in the transportation sector. Nevertheless, it is helpful to implement the CE in the SE to get actual values of when and why the product was returned. This makes the recommendations that the ML algorithm creates more reliable. In addition, it must be noted that the product's service life increases when a repair or maintain activity is performed. This still needs to be taken into account in this development process. Another valuable addition to optimising the ML algorithm would be the implementation of whether the product still works in the input data. These results are related to the presented problem but can also be generalised under certain conditions and thus applied to diverse data sets. In the following, the prerequisites must be fulfilled to successfully transfer machine learning applications to create a return process based on the CE activities.

First, it must be a company that provides SE for transportation and records the data of its products. A suitable classification method with the most accurate results must be selected for this data set. Furthermore, this approach can be transferred to various other products and businesses. The classification process is generally valid because the classes are created based on the individual dataset. In addition, this is supervised learning, which means that the classes for learning must already be available in the dataset. Since the result of this paper is a return process, the procedure can be easily applied to companies that want to implement CE in their organisation.

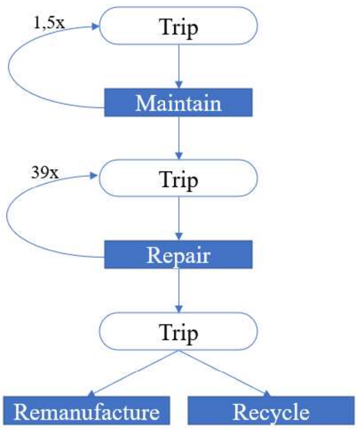

Figure 8: Developed Return Process based on ML

The following adjustments must be made to apply the procedure to companies not offering SE. Outside of an SE, it must first be precisely defined who is responsible for the CE activities as the ownership rights are transferred to the buyer. Furthermore, the recommendations Share and Reuse must be differentiated. In the case of share, the product remains with the same owner and is lent out. The product is passed on with reuse, but the owner is changed.

This raises challenges outside the SE because the company no longer owns the product. Hence, the responsibility of the product needs to be determined and recording the data is more complicated. Suppose the SE is for a product that is not used in transportation. This solution can be transferred, and it is only necessary to note that the data set contains the character traits responsible for the product's wear.

References

[1] van Veldhoven S, Brende B, Holdorf D, van Houten F, Raworth K, Vreeze D de, Watts M, Potocnik J, Sijbesma, Feike, Kortenhorst, Jules, Mcginty D. The circularity gap report; 2022.
[2] OECD. Environment at a glance 2020. Paris: OECD Publishing; 2020.
[3] OECD. The OECD RE-CIRCLE project: The economics of the transition to a more resource-efficient, circular economy; 2022.
[4] Zeng H, Chen X, Xiao X, Zhou Z. Institutional pressures, sustainable supply chain management, and circular economy capability: Empirical evidence from Chinese eco-industrial park firms. Journal of Cleaner Production 2017;155:54–65.
[5] Antikainen M, Uusitalo T, Kivikytö-Reponen P. Digitalisation as an Enabler of Circular Economy. Procedia CIRP 2018;73:45–9.
[6] Pagoropoulos A, Pigosso DC, McAloone TC. The Emergent Role of Digital Technologies in the Circular Economy: A Review. Procedia CIRP 2017;64:19–24.
[7] Chidepatil A, Bindra P, Kulkarni D, Qazi M, Kshirsagar M, Sankaran K. From Trash to Cash: How Blockchain and Multi-Sensor-Driven Artificial Intelligence Can Transform Circular Economy of Plastic Waste? Administrative Sciences 2020;10(2):23.
[8] Kumar NM, Mohammed MA, Abdulkareem KH, Damasevicius R, Mostafa SA, Maashi MS, Chopra SS. Artificial intelligence-based solution for sorting COVID related medical waste streams and supporting data-driven decisions for smart circular economy practice. Process Safety and Environmental Protection 2021;152:482–94.
[9] Nañez Alonso SL, Reier Forradellas RF, Pi Morell O, Jorge-Vazquez J. Digitalization, Circular Economy and Environmental Sustainability: The Application of Artificial Intelligence in the Efficient Self-Management of Waste. Sustainability 2021;13(4):2092.
[10] Wilson M, Paschen J, Pitt L. The circular economy meets artificial intelligence (AI): understanding the opportunities of AI for reverse logistics. MEQ 2022;33(1):9–25.
[11] Noman AA, Akter UH, Pranto TH, Haque AB. Machine Learning and Artificial Intelligence in Circular Economy: A Bibliometric Analysis and Systematic Literature Review. AETiC 2022;6(2):13–40.
[12] Nikolaou IE, Evangelinos KI, Allan S. A reverse logistics social responsibility evaluation framework based on the triple bottom line approach. Journal of Cleaner Production 2013;56:173–84.
[13] Casper R, Sundin E. Reverse Logistic Transportation and Packaging Concepts in Automotive Remanufacturing. Procedia Manufacturing 2018;25:154–60.
[14] Zaarour N, Melachrinoudis E, Solomon MM, Min H. The optimal determination of the collection period for returned products in the sustainable supply chain. International Journal of Logistics Research and Applications 2014;17(1):35–45.
[15] Kumar S, Putnam V. Cradle to cradle: Reverse logistics strategies and opportunities across three industry sectors. International Journal of Production Economics 2008;115(2):305–15.
[16] Dev NK, Shankar R, Qaiser FH. Industry 4.0 and circular economy: Operational excellence for sustainable reverse supply chain performance. Resources, Conservation and Recycling 2020;153:104583.
[17] Zhang X, Li Z, Wang Y. A Review of the Criteria and Methods of Reverse Logistics Supplier Selection. Processes 2020;8(6):705.
[18] Belezas F, Daniel AD. Innovation in the sharing economy: A systematic literature review and research framework. Technovation 2022:102509.
[19] Eckhardt GM, Houston MB, Jiang B, Lamberton C, Rindfleisch A, Zervas G. Marketing in the Sharing Economy. Journal of Marketing 2019;83(5):5–27.
[20] Geissinger A, Laurell C, Öberg C, Sandström C. How sustainable is the sharing economy? On the sustainability connotations of sharing economy platforms. Journal of Cleaner Production 2019;206:419–29.
[21] Mallick PK, Salling KB, Pigosso DCA, McAloone TC. Closing the loop: Establishing reverse logistics for a circular economy, a systematic review. J Environ Manage 2023;328:117017.