Communication in Design-for-Circularity: Requirements to a Knowledge Graph

Jan Martin Keil[1,*], Tobias Köhler[1] and Sirko Schindler[1]

[1]German Aerospace Center (DLR), Institute of Data Science, Mälzerstraße 3-5, 07745 Jena, Germany

Abstract
In times of potential resource shortages and urgent need of a reduction in CO_2 emission, reclaiming resources from end-of-life products becomes more and more relevant. This importance has already resulted in a growing number of legal regulations against the waste of resources and a strong push towards a circular economy. The design of a product has a strong influence on its environmental impact and its recyclability. In the project Methods and Technologies for an intelligent Circularity of Materials (MaTiC-M), we aim to develop a Design-for-Circularity methodology which accounts for product recyclability right from the design phase. However, lack of information for both product developers and recyclers about the entire life cycle of a product impedes recycling-friendly designs and efficient recycling processes. Overcoming this obstacle requires the integration of a wide range of data from different sources. Our goal is to develop a knowledge graph to facilitate the exchange of information between product designers and recyclers. In a first step, we analysed the requirements of various stakeholders on such a knowledge graph. In this paper, we present our procedure for the requirements analysis, as well as the resulting catalogue of competency questions.

Keywords
Circular Economy, Competency Questions, Knowledge Graph Engineering

1. Introduction

In the manufacturing industry, we can currently observe two major trends: On the one hand, there is a strong push towards net-zero emissions as witnessed, e.g., by the European Climate Law, which aims at a climate-neutral European economy and society by 2050 [1]. On the other hand, global supply chains are more and more under pressure due to political tensions, natural disasters but also a growing demand for various rare resources. Both trends cause a shift from relying solely on primary raw materials towards an increased use of recycled ones, e.g., from end-of-life products. Besides reducing the dependence from suppliers, this also lowers CO_2 emissions caused by production, processing, and transport of those materials. The vision is a circular economy in which fewer and fewer resources are disposed of as waste but instead are reintroduced back into the economic cycle.

The product design and development is of particular importance in this context as it influences the environmental impact along the entire life cycle of a product: It affects emissions, resource consumption, and waste generation during manufacturing, the extend of its use (including, e.g., robustness and durability but also the ability to repair it or replace broken components), and the degree to which its components can be recycled at its end-of-life. At the moment, product designers often lack proper information to make informed decisions on many of these sustainability aspects. In the same way, recyclers currently have little information available about products often causing suboptimal recycling results with respects to both ecological as well as economic regards as, e.g., rare and valuable materials cannot be specifically preserved.

In the project Methods and Technologies for an intelligent Circularity of Materials (MaTiC-M) [2], we develop a Design-for-Circularity methodology that already accounts for the recyclability of a product during its design phase. This allows product designers to make better decisions early on, resulting in

The 2nd International Workshop on Knowledge Graphs for Sustainability (KG4S2024) – Colocated with the 21st Extended Semantic Web Conference (ESWC2024), May 27th, 2024, Hersonissos, Greece.
*Corresponding author.
jan_martin.keil@dlr.de (J. M. Keil); tobias.koehler@dlr.de (T. Köhler); sirko.schindler@dlr.de (S. Schindler)
0000-0002-7733-0193 (J. M. Keil); 0000-0003-0567-6934 (T. Köhler); 0000-0002-0964-4457 (S. Schindler)

higher recycling rates and thus a lower environmental footprint at the end-of-life or end-of-use of a product. The methodology is initially developed for a range of demonstrators from the fields of aviation, energy, and transport. This broad range of domains and the early involvement of various industry partners will ensure a generic approach and a possible transfer into a wide range of application areas.

A crucial aspect of this methodology is the information exchange between the various involved stakeholders, including the aforementioned product designers and recyclers but also, e.g., life cycle analysts, to rate different design options based on their environmental impact. In this paper, we report on the development of the knowledge graph that is at the heart of this information exchange. Here, we collect and interconnect all relevant information ranging from material characteristics, over life cycle assessments, to recycling strategies. The goal is to provide a decision support system for both product designers and recyclers that helps to make more efficient use of available resources, reduce the reliance on primary raw materials, and thus contribute to meeting the net-zero goals set out by politics.

This paper is structured as follows: In Section 2, we provide background information on circular economy and the Linked Open Terms (LOT) methodology. In Section 3, we introduce the method we applied to determine the requirements for the MaTiC-M knowledge graph. In Section 4, we present the results of the requirements analysis. In Section 5, we reflect on lessons learned with the requirements analysis in this projects. In Section 6, we provide a conclusion and give an outlook for the further goals of the project.

2. Background

In this section we provide background information on key concepts for this paper.

2.1. Circular Economy

A circular economy [3] is a economy model that aims to reduce waste to a minimum. It involves the maximisation of the use of materials and products by sharing, leasing, reusing, repairing, refurbishing, and recycling them. The circular economy model is in contrast to the linear economy model, which relies on cheap, easily accessible materials and energy and that encourages one-time-use of materials and generation of waste.

2.2. Linked Open Terms Methodology

The Linked Open Terms (LOT) methodology [4] is an approach to ontology engineering. It describes a development process that consist of four main activities: (1) the ontology requirements specification activity, (2) the ontology implementation activity, (3) the ontology publication activity, and (4) the ontology maintenance activity. Each activity consists of several sub-activities. For the ontology requirements specification it defines seven sub-activities: (1.1) the use case specification activity, (1.2) the data exchange identification activity, (1.3) the purpose and scope identification activity, (1.4) the functional ontology requirements proposal, (1.5) the functional requirements completion activity, (1.6) the ontology requirements specification document formalisation, and (1.7) the optional functional ontology requirements formalisation activity.

3. Method

An important obstacle to implement a circular economy is the lack of information of all stakeholders along the product life cycle. In particular, product developers often can not assess the implications of certain design decisions for the sustainability and recyclability of the product. Vice versa, recyclers lack information on the individual parts and materials used in a product to select the optimal recycling strategy. In MaTiC-M we aim to address this obstacle with an RDF knowledge graph that supports the exchange of the required information. RDF supports us in the integration processes of the tool landscapes of the involved stakeholders, e.g., via a rather seamless transition in communication protocols using JSON-LD [5]. Further, RDF allows for an easy integration with and reuse of large knowledge bases like ChEBI [6]. Finally by using RDF as the base for information storage and exchange, we encourage adherence to the FAIR principles [7] with prospects of, e.g., FAIR Digital Twins [8] as a later use case.

For the development of the MaTiC-M knowledge graph, we apply the Linked Open Terms (LOT) methodology [4], a current and yet widespread methodology for ontology engineering. However, we needed to slightly adapt the process to account for the conditions of the project framework and we skipped the optional functional ontology requirements formalisation activity:

3.1. Use Case Specification

Based on an internal preliminary study, the use cases have been outlined in the MaTiC-M project plan:

1. provide information on recycling capabilities to enable an assessment of the recyclability of a product during its design as part of a Design-for-Circularity methodology,

2. provide information about designs and materials to enable the selection of an optimal decomposition and recycling at the end of life of a product.

However, as the Design-for-Circularity methodology itself is developed during the project, we expect further refinement of the use case specifications over the course of the project.

3.2. Data Exchange Identification

Interviews were conducted with domain experts, to identify the file formats currently used for storing or exchanging data between stakeholders along the product life cycle. These interviews revealed a range of relevant data formats, including tabular data stored in CSV and Excel files. Furthermore, 3D models of the product designs must be analysed by several stakeholders. However, these analyses require different computer-aided design (CAD) data formats. The standard for the exchange of product model data (STEP) format [9] enables to derive 3D data in all other required formats. Therefore, we selected the STEP format as the format for the exchange of 3D data. Unfortunately, some software tools do not export all relevant information into STEP, although it is the broadly accepted baseline format for CAD data. This makes some workarounds necessary to ensure comprehensive data integration.

3.3. Purpose And Scope Identification

The purpose and scope of the knowledge graph is already outlined in the project proposal. However, the specific methods the knowledge graph needs to support are still under active development. As a consequence, the purpose and scope definition could not be detailed so far. This is a unavoidable deviation from the LOT methodology. Therefore, the following knowledge graph development is based on the outlined assumptions, which might change to a certain degree over the course of the project, though. Nevertheless, some decisions have already been made. For example, the primary language of the knowledge graph was decided to be English, supplemented by German as secondary language.

3.4. Functional Ontology Requirements Proposal

We specified the functional ontology requirements in the form of competency questions (CQs). In a first step, all involved domain experts were asked to provide CQs that are relevant for the use cases in their domain and area of expertise. Overall, 14 domain experts from seven institutes of different domains within the German Aerospace Center (DLR) contributed to the collection of CQs. To prepare the domain experts for the task, example questions were provided along with an introduction of the purpose of the CQs. This resulted in several independent list of CQs (one per stakeholder or group of stakeholders involved in the project; in total seven independent lists). However, the level of granularity and compliance with the question format differed. In particular, we experienced a clear difference between domain experts with and without prior experience in developing or using knowledge graphs. Initially, the CQs were collected and edited in German.

In a second step, we consolidated the collected CQs into a single list. During this consolidation and in close collaboration with the domain experts, we also clarified and aligned the CQs. This included the following activities:

Ensure Question Format
Some initial CQs were only single keywords instead of full questions or the questions were consecutive questions that are understandable only in the context of the previous ones. We rephrased these CQs to comply with the intended question format and to be unambiguous and independent of one another.

Split Complex Questions
Some initial CQs were rather long or complex and involved a multitude of independent aspects. We split these CQs into multiple CQs to obtain a more even level of granularity.

Highlight Important Terms
Each CQ contains important terms. We highlighted these important terms to identify relevant concepts. Further, this helps to keep an overview during the question consolidation.

Group Questions
We noticed that there is couple of central topics among the CQs. Therefore, we grouped the CQs based on their main topic of concern. This activity is merely for organisational purposes and does not imply any priority or rating among the CQs. It helps to keep an overview during the question consolidation and we expect it to help during the ontology implementation, too.

Merge Duplicated Questions
Some equivalent or very similar questions were part of or derived from multiple initial CQ lists. We merged these CQs to avoid redundancy.

Align Terminology
We identified terminology gaps between different domains and contributors and clarified the meaning of used terms. For example, we decided that we consider (a) the term "material" as the chemical composition and structure of things like in "materials science" and therefore rejected the interpretation as things that are used to produce another thing like in a "bill of material" [10], (b) the term "part" as a thing that is not assembled of other things and therefore rejected the interpretation as a thing that is (meant to be) assembled into an assembly, (c) the term "component" instead of the term "part" as a thing that is (meant to be) assembled into an assembly and thereby has a special role in this context, (d) the term "product" as a part or assembly that is (meant to be) sold and thereby has a special role in this context and therefore rejecting the interpretation as a thing that can be sold to a consumer. These linguistic ambiguities apply to German and English in a similar way.

During these activities, we maintained the links between initial CQs and their according consolidated CQs to track the provenance of the latter. Due to splitting and merging during the consolidation, there is a many-to-many relation between initial and consolidated CQs. The provenance of each CQ helped to ensure that all initial CQs have been considered entirely in the consolidated CQs and that the consolidated CQs do not extend the scope of the initial CQs. Further, it eased the validation of consolidated CQs in the next step. These activities resulted in a CQ catalogue draft.

3.5. Functional Requirements Completion

To complete the CQ catalogue, we conducted a verification of the catalogue draft. All involved domain experts were asked to check whether the consolidated CQs cover all their information needs and that

Table 1
Example competency question (CQ) for each topic.

ID	Topic	Example CQ
108	material	What **properties** (e.g., **tensile strength**, **thermal conductivity**, **heat capacity**, **density**) does a **material** have?
202	individual part / assembly / product	Which **parts** and **assemblies** does an **assembly** / **product** consist of?
301	connection / joining	Which **connections** of which **connection type** / which **joining method** exist between two **components**?
401	material source	In what **amount** is a **material** available from a primary / secondary **material source**?
503	manufacturing / production	How much **electrical energy** does a **production process** / **manufacturing process** require per **workpiece** or **amount of material**?
601	scrap	What **material** does an amount of **scrap** consist of?
704	collection	What **personnel** is required for the **collection** of the scrap of a **product** or **assembly**?
804	disassembly	Which **disassembly steps** are to be carried out for a **product**?
902	recycling	How much **electrical energy** does a **recycling step** require per **product** or **component** or **amount of material**?
1002	appliance / machine	What is the **lifespan** of a **device**?
1101	personnel requirements	How many **person hours** with which **qualifications** does a **personnel requirement** comprise?

none of their intents was misinterpreted. This revealed a few gaps in the consolidated CQs that have subsequently been addressed. In one case, one of several alternative concepts in a question was missing. In other cases, the links between the initial CQs and the consolidated CQs were incorrect.

3.6. Ontology Requirements Specification Document Formalisation

To produce the ontology requirements specification document (ORSD), we decided to create a table containing the consolidated CQs, stored as a CSV file. As the non-functional requirements can not be comprehensively specified yet due to the project's setup, this is sufficient to store the known requirements. Each CQ is stored both in German as well as in English. We used the translation service DeepL Pro[1] to generate a draft translation from German into English, which was manually reviewed and corrected afterwards. Further, we assigned a numeric identifier to each CQ.

4. Results

At this stage of development, we identified a total number of 97 CQs that are grouped regarding their central concepts into 11 topic groups: *Material* contains questions around characteristics of individual materials employed. *Individual part / assembly / product* centres around which components make a final product whereas *connection / joining* includes question of how they are assembled. *Material source* and *manufacturing / production* provide necessary information for the life cycle assessment [11] during the creation of a product. *Scrap* and *collection* describe the collection of products at their end-of-life and end-of-use phase. Finally, *disassembly* and *recycling* hold CQs for recycling processing while *appliance / machine* and *personnel requirements* do so for the resources required for processes along the whole product life cycle. Table 1 shows one example CQ per topic. The CQs are recorded in tabular form including translations in both English and German. The topics list illustrates the broad range of domains that need to be involved in the Design-for-Circularity methodology.

[1] https://www.deepl.com

Unfortunately, we can not publish all consolidated CQs for now to protect intellectual property rights of our project partners. Therefore, we omitted or generalised some of the question in the consolidated CQ table for publishing. The resulting table contains 65 CQs. It is publicly available [12] and licensed under a permissive CC BY 4.0 license[2].

5. Lessons Learned

As mentioned in Section 3, we had to adapt the LOT methodology according to the framework of the project. Even though very helpful, it did not fit perfectly in our situation, where the application environment of the knowledge graph is not yet largely defined. In such case, a less linear but more agile methodology might be a better option.

Maintaining the provenance of CQs turned out to be very helpful for later activities during the requirements specification. But as we used a wiki platform to collect and consolidate the CQs, maintaining the links between initial CQs and consolidated CQs was a rather cumbersome manual task.

During the collection of the initial CQs we experienced several misconceptions regarding their purpose and form. In particular, there was a misconception that the CQs are about information needed within the project as opposed to information needed to apply the Design-for-Circularity methodology. We also noticed a clear difference between domain experts with and without prior experience in developing or using knowledge graphs. As a take-away, we need to put more emphasis on the introduction of the requirements specification activities to the domain experts in future projects.

6. Conclusion

The implementation of a circular economy requires stakeholders along a product life cycle to exchange information about the product. For instance, product designers need more information about the processes at the end-of-life/end-of-use of a product, to ease the product recycling and recyclers have to know the structure and involved materials of a product to decide for an optimal recycling strategy. We investigated the requirements for a knowledge graph that supports product designers and recyclers to increase the recyclability and the actual recycling rate of products on the path to a circular economy. Based on the identified requirements, we will develop the MaTiC-M knowledge graph to support the Design-for-Circularity methodology currently under development. We intend to make extensive use of existing ontologies and knowledge graphs to (a) reduce the development effort and (b) ensure compatibility with other efforts in the domain.

Acknowledgments

Many thanks to Mohamad Abdallah (DLR Institute of Vehicle Concepts), Elmar Beeh (DLR Institute of Vehicle Concepts), Katrin Bugelnig (DLR Institute of Materials Research), Jan Haubrich (DLR Institute of Materials Research), Sree Sameer Kumtamukkula (DLR Institute of Vehicle Concepts), Tom Lorenz (DLR Institute of Low-Carbon Industrial Processes), Miriam Löbbecke (DLR Institute of Materials Research), Nicole Carina Neumann (DLR Institute of Future Fuels), Korbinian Nottensteiner (DLR Institute of Robotics and Mechatronics), Diana Peters (DLR Institute of Data Science), Robert Reiser (DLR Institute of System Dynamics and Control), Ismael Valentin Rodriguez Brena (DLR Institute of Robotics and Mechatronics), Ralf Sturm (DLR Institute of Vehicle Concepts), Bernhard Thiele (DLR Institute of System Dynamics and Control), and Sven Torstrick-von der Lieth (DLR Institute of Lightweight Systems) for their valuable contributions to the requirements analysis.

[2] https://creativecommons.org/licenses/by/4.0/

References

[1] Regulation (EU) 2021/1119 of the European Parliament and of the Council of 30 June 2021 establishing the framework for achieving climate neutrality and amending Regulations (EC) No 401/2009 and (EU) 2018/1999 ('European Climate Law'), in: Official Journal, L 243, 2021, pp. 1–17. URL: http://data.europa.eu/eli/reg/2021/1119/oj.

[2] J. M. Keil, D. Peters, T. Lorenz, S. Schindler, Digital knowledge exchange for circularity of materials, in: DLRK 2023, 2023, p. 1. URL: https://elib.dlr.de/197942/.

[3] European Parliament, Circular economy: definition, importance and benefits, 2023. URL: https://www.europarl.europa.eu/topics/en/article/20151201STO05603/circular-economy-definition-importance-and-benefits, Accessed: 2024-02-28.

[4] M. Poveda-Villalón, A. Fernández-Izquierdo, M. Fernández-López, R. García-Castro, LOT: An industrial oriented ontology engineering framework, Engineering Applications of Artificial Intelligence 111 (2022) 104755. doi:10.1016/j.engappai.2022.104755.

[5] G. Kellogg, P.-A. Champin, D. Longley, JSON-LD 1.1 – A JSON-based Serialization for Linked Data, W3C Recommendation, 2020. URL: https://www.w3.org/TR/2020/REC-json-ld11-20200716/.

[6] K. Degtyarenko, P. de Matos, M. Ennis, J. Hastings, M. Zbinden, A. McNaught, R. Alcantara, M. Darsow, M. Guedj, M. Ashburner, Chebi: a database and ontology for chemical entities of biological interest, Nucleic Acids Research 36 (2007) D344–D350. doi:10.1093/nar/gkm791.

[7] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. 't Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. Mons, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S.-A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, B. Mons, The fair guiding principles for scientific data management and stewardship, Scientific Data 3 (2016). doi:10.1038/sdata.2016.18.

[8] D. Peters, S. Schindler, Fair for digital twins, CEAS Space Journal 16 (2023) 367–374. doi:10.1007/s12567-023-00506-y.

[9] ISO 10303-1:2024, Industrial automation systems and integration – Product dara representation and exchange – Part 1: Overview and fundamental principles, Standard, International Organization for Standardization, 2024. URL: https://www.iso.org/standard/83105.html.

[10] H. Li, M. Abd Nikooie Pour, Y. Li, M. Lindecrantz, E. Blomqvist, P. Lambrix, A survey of general ontologies for the cross-industry domain of circular economy, in: Companion Proceedings of the ACM Web Conference 2023, WWW '23 Companion, Association for Computing Machinery, New York, NY, USA, 2023, p. 731–741. doi:10.1145/3543873.3587613.

[11] M. Z. Hauschild, R. K. Rosenbaum, S. I. Olsen (Eds.), Life Cycle Assessment: Theory and Practice, Springer International Publishing, 2018. doi:10.1007/978-3-319-56475-3.

[12] J. M. Keil, M. Abdallah, E. Beeh, K. Bugelnig, J. Haubrich, S. S. Kumtamukkula, T. Lorenz, M. Löbbecke, N. C. Neumann, K. Nottensteiner, D. Peters, R. Reiser, I. V. Rodriguez Brena, R. Sturm, B. Thiele, S. Torstrick-von der Lieth, MaTiC-M Knowledge Graph Competency Questions, 2024. doi:10.5281/zenodo.10730784.