IOP Conference Series: Earth and Environmental Science

PURPOSE-LED PUBLISHING™

PAPER • **OPEN ACCESS**

Leveraging Artificial Intelligence for Creating Material Passports from Building Stock Documents: A Workflow for Promoting Circularity in the Built Environment

To cite this article: Svenja Lauble *et al* 2025 *IOP Conf. Ser.: Earth Environ. Sci.* **1546** 012076

View the article online for updates and enhancements.

You may also like

- Assessment of Malaysian E-Passport PKI based on ISO 27000 Series International Standards
  Ahmad Iqbal Hakim bin Suhaimi, Nurulhuda Noordin and Mohd Faizul bin Ya'kub
- Integrating Digital Product Passports and BIM for Circular Construction: A Review of Challenges and Opportunities for Data-Driven Material Reuse and Sustainability Assessment
  Shayesteh Hamidkhaniha, Antonio Garcia Martinez and Mojtaba Sedaghat
- Circular material passports for buildings – Providing a robust methodology for promoting circular buildings
  V Göswein, S Carvalho, C Cerqueira et al.

This content was downloaded from IP address 129.13.194.163 on 09/01/2026 at 12:39

IOP Conf. Series: Earth and Environmental Science **1546** (2025) 012076 doi:10.1088/1755-1315/1546/1/012076

Leveraging Artificial Intelligence for Creating Material Passports from Building Stock Documents: A Workflow for Promoting Circularity in the Built Environment

Svenja Lauble[1*], Andreas Geiger[2] and Philipp Zielke[1]

[1] Institute of Technology and Management in Construction, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
[2] Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

*E-mail: Svenja.lauble@kit.edu

Abstract. To advance the concept of a circular built environment, it is essential to collate and examine material data from existing structures. Material passports—digital documents detailing materials, components, and their properties—are vital in this process. However, generating material passports from building stock documents remains underexplored, presenting challenges such as inefficient information exchange and the labour-intensive digitizing of building stock documents. This paper introduces a novel AI-enabled workflow to address these challenges by automating the extraction, organization, and enhancement of product and material information from building stock documents. Unlike traditional manual methods, the workflow leverages artificial intelligence to improve data extraction, streamline standardization, and integrate publicly available datasets into a centralized digital database. The processed data can be retrieved via the database in standardized data formats (e.g. Excel, CSV, building models) or via an API to facilitate digital data exchange and enable an assessment of the circular economy. The workflow's practical application is demonstrated in a German building case study, highlighting challenges in data quality and standardization during implementation. By automating and enhancing data handling, the workflow reduces manual effort, increases efficiency, and ensures higher data reliability. This approach enables the streamlined creation of material passports, empowering property managers with actionable insights for renovation and maintenance decisions while contributing to the circularity goals of the built environment.

1. Introduction

The construction sector is responsible for 50% of all raw materials extraction and 37% of CO_2 emissions in the EU [1, 2]. Therefore, buildings play a pivotal role in the EU's 2050 climate neutrality goal, with a pressing need for innovative tools and methodologies to evaluate and manage materials in existing buildings. Materials Passports (MPs) document materials for reuse and recycling [3]. They improve recyclability and guide future material development by

Content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

identifying embedded materials [4]. However, a lack of knowledge about building stocks hinders recycling efforts [5]. Emerging solutions include Building Information Modeling (BIM) for MP generation, deconstructability, and environmental assessments [4, 6]. GIS aids material management [7], while scan-to-BIM and computer vision enhance material recovery and identification [8]. Despite these advancements, tools face issues like proprietary constraints, poor integration, and fragmented workflows [9, 10]. Data inaccessibility and platform incompatibilities further hinder MP creation [11].

This work aims to integrate building stock data into an MP using AI, addressing limitations in current tools. Preliminary results show improved data accessibility and workflow efficiency, supporting the transition to a circular economy. Section 2 introduces MPs and AI. Section 3 outlines methodology, Section 4 explains data integration, and Section 5 presents case study results.

2. Review of the Literature

2.1 Material Passport for Building Stock

The concept of MP has gained attention for its potential to drive sustainable practices across industries. [12] identified key benefits, including promoting secondary material markets, enhancing material traceability, reducing environmental footprints, and enabling market differentiation. [13] highlighted their role in Building Renovation Passports, comparing three development initiatives. [14] outlined a process for creating Material Passports using an IFC file: generating a bill of materials (BOM), categorizing materials, classifying them, and gathering chemical property data. The process also involves assessing health and safety hazards, evaluating circularity, determining lifetime, and recyclability. [15] proposed an alternative for incomplete IFC files: creating an object list, mapping objects to construction products, and developing Environmental Product Declarations (EPDs). Additional circularity data can then be added to complete the passport (Figure 1).

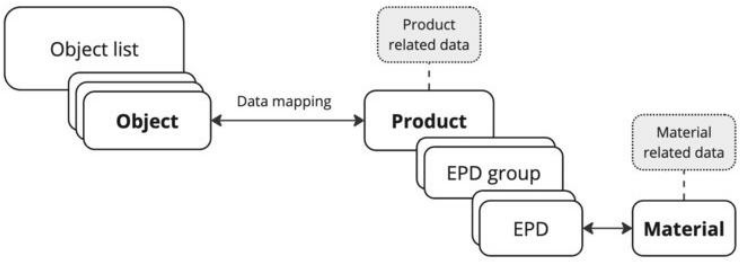

Figure 1. Workflow from object list to material passport [15]

Creating an MP based on an incomplete IFC requires extracting data from building documents, which may be analog or digital (Table 1). These include building profiles with essential details like construction year and use, historical maps for structural insights, and expert reports identifying material hazards. Product data sheets and technical documents provide material properties, while energy certificates relate to efficiency and sustainability. Visual data, such as analog/digital plans, photographs, 3D models, point clouds, and photogrammetry, offer detailed representations of materials. However, most existing buildings lack IFC files, limiting comprehensive material reuse insights.

IOP Conf. Series: Earth and Environmental Science **1546** (2025) 012076 doi:10.1088/1755-1315/1546/1/012076

Table 1. Overview of documents that may be available in the built environment [16].

	Analogue data	Digital data
Text-based data	Building profile, historical documents in paper format, product data sheets, energy certificate, technical documents, expert reports, e.g. pollutants, listed building regulations (if listed), area calculations, manual measurement, consumption data,	Building profile, product data sheets, energy certificate, technical documents, expert reports, e.g. pollutants, listed building regulations (if listed), area calculations, manual measurement, consumption data
Visual data	Maps and cadastral plans, PDF plans at a scale of 1:50 and 1:100, photos	Maps and cadastral plans, 3D city models, GIS, Google Streetview, PDF plans at a scale of 1:50 and 1:100, photos, 360-degree images, point clouds/ photogrammetry, aerial photos, depth scans, IFC models

2.2 Relevant Artificial Intelligence Technologies

To extract product and material information from building documents, methods for processing visual and text-based data are essential. Optical Character Recognition (OCR) digitizes text, while Natural Language Processing (NLP) extracts and tags relevant information from sources like energy certificates and technical specifications. Recent advancements in NLP, particularly with large language models (LLMs) [17, 18], enable efficient information extraction from building documents and online resources. NLP is being validated for construction use cases, such as analysing unstructured reports [19].

Computer Vision (CV) processes visual data from images and videos [20, 21] and performs tasks like image classification, object localization, and semantic segmentation. In building analysis, CV aids in segmenting objects, such as walls in floor plans, for detailed evaluation. Integrating OCR, NLP, and CV technologies creates a robust framework for automating the extraction of critical textual and visual information, supporting material passport creation.

3. Workflow Creation Methodology

3.1 Workflow Creation Methodology

In developing the workflow, our approach was based on the V-Model [22, 23]. The initial four steps for the decomposition and definition were executed: (1) User Requirements, defining available documents and preferred data format results, (2) System Specification, defining technical and functional aspects, (3) Verification Plan, defining metrics for testing (e.g. number of recognized building elements, percentage differences in volume and area), and (4) Verification Procedure, providing critical feedback after testing. However, the subsequent integration and verification stages - (5) Fabricate, (6) Inspect, (7) Verify, (8) Integrate System, and (9) Validate User Requirements - were omitted and will require further investigation in future research.

Initial prototype versions of the services for floor plan recognition from raster data and, based on this, for generating the 3D building model are used to validate the workflow developed here. The focus is on the workflow itself; statements about the recognition rate and quality of the available documents are not the focus.

3.2 Description of the Data Set

To validate the workflow, an office and administration building located in Germany were selected as a test case in the research project NaiS ("Nachhaltige intelligente Sanierungsmaßnahmen"). The selected office building was constructed in four phases from the 1950s to 1993, with a total floor area of approximately 5,500 m^2. The following data is available for this building. Floor plans and elevations are provided in PDF format. In addition, photographs documenting building deficiencies are included in the PDF report, along with additional descriptive images in JPEG format. Energy consumption data for both electricity and gas are also available.

4. Presentation of the Workflow

The clear definition of user requirements is crucial for the success of the workflow. The requirements were documented in a series of workshops with around 30 participants, including industry and IT experts in the field of building stock digitization, and were gradually refined.

1. Data Processing: All existing building data must be seamlessly processed and integrated into the workflow.
2. User Involvement: Users should actively participate in the process, with the ability to review and correct data while understanding its quality.
3. Data Accessibility: Digitized data should be available in a simple Excel spreadsheet for easy access and as an IFC model for advanced processing by property portfolio managers.
4. Expert Verification: Qualified professionals and strategic partners should validate, supplement, and reintegrate incomplete or outdated data to ensure a reliable foundation for further analysis.
5. Preliminary Assessment: Property managers should be able to perform an initial evaluation of data quality and refurbishment potential on the platform, enabling sustainability assessments before detailed analyses.

The process diagram describes an end-to-end workflow, which is subdivided into three principal domains: data processing, project management, and dashboard (Figure 2).

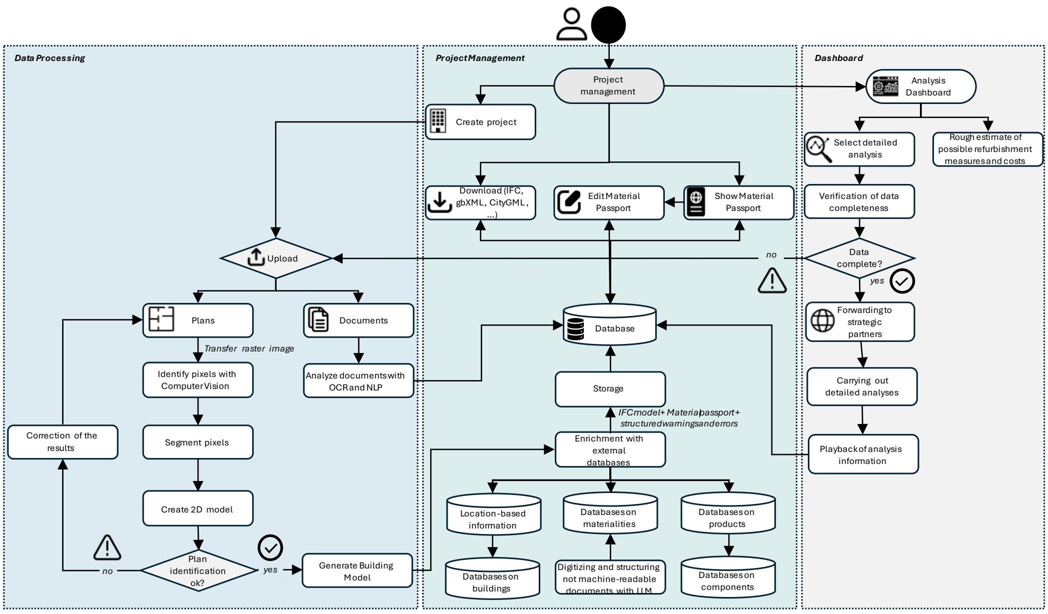

Figure 2. Overview of the workflow for Creating Material Passports from Building Stock Documents

The creation of a project is initiated by the uploading of plans, including floor plans and elevations, and documents, such as the energy performance certificate, expert reports, or construction specifications, into the system. The uploaded plans are subjected to CV analysis, whereby the system initially starts with a rasterization, scales down and rotates the plan. Then the system identifies the pixels and then segments them into walls, windows and doors. The user has the option of reworking and correcting any poorly recognized or unreliable segments. A two-dimensional model is then generated from this segmentation with vectorization. Furthermore, the materials are derived from the legend, in addition to the geographical location and other pertinent information from the plan header, if available. Concurrently, the documents are subjected to analysis through the utilization of OCR and NLP techniques, with the objective of extracting and processing the textual data. Subsequently, the user verifies the accuracy of the plan identification. In the event of a positive outcome, an IFC model is generated. In the event of an unsuccessful plan identification, the results are corrected, and the plans and documents are either re-uploaded or corrected once more before the process continues. The digitized information from the plans and documents is now stored in a structured database, which includes product and material data as well as an IFC model.

In the context of project management, the resulting material passport can be either displayed or edited, and the results of the information processing can be downloaded. When displayed, the respective sources of the information are indicated with an associated degree of uncertainty pertaining to potential corrections. The material passport is stored in conjunction with the IFC model within a database. The database also contains structured warnings and errors that may occur during processing, in addition to the IFC models and material passports. The digitized information is augmented with supplementary external data sources, comprising information on locations, materials, products, components and buildings. The location-based data set can include information on geodata, climate data, solar or geothermal potential, and the demolition atlas, which can be used for the analysis of resources. Furthermore, data pertaining to the material register, energy prices and CO_2 emission factors can be incorporated into the building. Databases such as the Environmental Product Declaration (EPD), the Old Building Atlas (Altbauatlas), Tabula (Typology Approach for Building Stock Energy Assessment) and building price databases are linked to the component, product and material. To comprehensively digitize the data from external data sources and make it available in a structured manner, non-machine-readable documents are processed using an LLM and kept up to date by generating APIs.

In the final stage of the process, the dashboard enables users to undertake a comprehensive examination of the data. The first step stage is to ascertain the completeness of the data for individual analyses. The dashboard indicates which data are still relevant for the specified analyses and thus provides guidance on how to improve data quality. In the event of data being absent, a notification is generated, and the incomplete data is returned for reprocessing. Upon completion of the data set, it can be forwarded to the designated strategic partners for detailed analysis. Furthermore, it is possible to ascertain which potential refurbishment measures would be typical for this building and to estimate the associated costs.

These three areas provide the property manager, as well as software companies engaged in detailed analyses pertaining to sustainability and refurbishment, with an initial data foundation upon which to conduct assessments. One strategic objective would be to analyse data from numerous refurbishment projects as a data broker to demonstrate and evaluate statistical values for the building stock and the projects. To engender trust, it is deemed essential that this platform is hosted by a neutral organization with a non-profit objective.

5. Discussion based on the Case Study

For the case study considered in this article, the focus is on the digitization (see also Figure 3, left section 'Data Processing') of analog floor plans and the automated generation of a 3D representation for each floor plan. Existing floor plans can be processed as PDF or image files. The basis for recognizing the floor plans is an AI model that has been trained and validated using around 1,600 labelled images for the three classes of wall, window and door. The result of the recognition are the areas of the trained building element classes from which semantically enriched 2D polygons are generated. Based on these 2D polygons, 3D building elements are generated in a further process step and a 3D building floor is created. Initial tests were organized with project partners based on the refurbishment project mentioned above. For this purpose, the floor plan generated by the AI and the resulting IFC model were compared with an IFC Revit model created based on the floor plan. Figure 3 shows the areas of the recognized building element classes on the left-hand side. Walls are shown in orange, the windows in blue and the doors in green. On the right side the generated IFC model is shown. The results of this comparison can be seen in Table 2. The walls presented a difference of 18,6% additional volume in the 127 walls in the final IFC model generated by the AI compared to the 61 walls in the reference model. The presence of columns in the plan divides the walls into many pieces of varying thickness, which adds additional volume. It was expected that the results would show a larger number of walls and a slightly larger volume. The height of the doors and windows were set to the average height values for the windows and doors from the reference model. The area of the base length and the height were used for the comparison, since the volume of the elements depends on the model and types used. The results were the recognition of almost all the doors, with a discrepancy of 7,4% in the final area; and 8 additional windows recognized with a discrepancy of 12,3% of the area as a result. Two of the additional windows were façade elements which, if included in the reference model, would result in a smaller difference.

Figure 3. Initial floor plan (left) to digital building model (right)

Table 2. Results of the comparison between the AI-generated model and the reference model

Elements	AI-generated model	Reference model	Deviation
Walls			
Quantity	120	61	59
Total volume	253.58 m^3	214.41 m^3	18.3%
Windows			
Quantity	32	24	8
Total volume	85.53 m^2	77.39 m^2	10.5%
Doors			
Quantity	30	30	0
Total volume	70.3 m^2	59.1 m^2	19%

The IFC model generated during the tests could be successfully read and processed in the tools of the project partners and various CAD applications. However, it has also been shown that additional elements within the floor plans need to be considered for recognition. These include the room stamps information which is relevant for further use of the building model. In addition, the various areas of a plan such as the plan header, legend, views or sections must also be labeled. This differentiation is necessary, for example, to derive information such as floor heights, parapet and window heights or the building height. It also ensures that building elements are not incorrectly recognized in other parts of the plan and increases reliability and performance, as not all plan areas have to be considered during floor plan recognition.

With the 3D building model generated with AI, however, it must also be considered that a comparable accuracy to modelling in a CAD system cannot be achieved. The accuracy is sufficient for the use case considered here, namely the refurbishment of buildings and the assessment of existing buildings. Product characteristics can be identified through the legend in the floor plans and mapped with external product information, such as U-values and thermal properties. This enhances the digital building model for further analysis and material passport generation. Material classes may also be determined via the Old Building Atlas using the year of construction. Ongoing research is exploring how product information can be extracted from additional document types to improve mapping.

6. Conclusion

This study introduces an innovative workflow designed to address the significant gaps in the creation of MP for existing building stocks. By employing AI for data processing, the workflow addresses the challenges inherent in the digitization and integration of disparate building documentation, thereby enhancing data accessibility and interoperability. The findings illustrate that AI is capable of efficiently processing floor plans and associated documentation to generate material passports, although there are inherent limitations in terms of accuracy when compared to manual CAD modelling. The case study corroborates the workflow's capacity to generate semantically enhanced building models and underscores the potential for integrating external datasets to augment the material passport's comprehensiveness.

Despite the promising initial results, this study highlights areas that require closer and more detailed investigation. The focus is on the overall workflow, so the issues of data quality and recognition accuracy for specific building elements are not discussed in detail.

The proposed approach paves the way for a transition to a circular built environment, equipping property managers with the tools to make informed decisions regarding renovation and sustainability. By addressing both technical and operational barriers, this workflow contributes to the development of more efficient tools for material management and circularity assessments. In conclusion, this study demonstrates a viable pathway for advancing the EU's circular economy objectives and underscores the critical role of digitalization in achieving sustainable practices in the construction sector.

Acknowledgements

The authors acknowledge the financial support by the Federal Ministry for Economic Affairs and Climate Action of Germany in the project “Nachhaltige intelligente Sanierungsmaßnahmen” (project number 01MN23005). The generative AI tool OpenAI ChatGPT (2025) was used in the research and the drafting of the manuscript.

IOP Conf. Series: Earth and Environmental Science **1546** (2025) 012076 doi:10.1088/1755-1315/1546/1/012076

References

[1] Benachio GLF, do Carmo Duarte Freitas M, Tavares SF. 2020. Circular economy in the construction industry: a systematic literature review. *J. Clean. Product.* 260, 121046. https://doi.org/10.1016/j.jclepro.2020.121046.
[2] UNEP (2024). Building Materials And The Climate: Constructing A New Future. https://wedocs.unep.org/handle/20.500.11822/43293 (date of access 19/09/2024).
[3] Building As Material Banks (BAMB) (2016). D1 Synthesis of the State-Of-The-Art: Key Barriers and Opportunities for Materials Passports and Reversible Building Design in the Current System. http://www.bamb2020.eu/wp-content/uploads/2016/03/D1_Synthesis-report-onState-of-the-art_20161129_FINAL.pdf.
[4] Honic, M., Kovacic, I., & Rechberger, H. (2019). Improving the recycling potential of buildings through Material Passports (MP): An Austrian case study. *Journal of Cleaner Production,* 217, 787–797. https://doi.org/10.1016/j.jclepro.2019.01.212.
[5] Brunner, P. H. (2011). Urban mining a contribution to reindustrializing the city. *Journal of Industrial Ecology,* 15(3), 339–341. https://doi.org/10.1111/j.1530-9290.2011.00345.x.
[6] Koutamanis, A. (2024). From Building Information Modelling to Digital Twins: Digital Representation for a Circular Economy. In *A Circular Built Environment in the Digital Age* (pp. 3–20). Springer. https://doi.org/10.1007/978-3-031-39675-5_1.
[7] Tsui, T., Wuyts, W., & Van den Berghe, K. (2024). Geographic information systems for circular cities and regions. In *A Circular Built Environment in the Digital Age,* 21. https://doi.org/10.1007/978-3-031-39675-5_2.
[8] Armeni, I., Raghu, D., & De Wolf, C. (2024). Artificial Intelligence for Predicting Reuse Patterns. In *A Circular Built Environment in the Digital Age,* 57. https://doi.org/10.1007/978-3-031-39675-5_4.
[9] Valluru, P., Karlapudi, J., Menzel, K., Mätäsniemi, T., & Shemeikka, J. (2020). A semantic data model to represent building material data in AEC collaborative workflows. *Boosting Collaborative Networks 4.0,* 133–142. https://doi.org/10.1007/978-3-030-62412-5_11.
[10] Çetin, S., Gruis, V., Rukanova, B., Tan, Y. H., & De Wolf, C. (2022). A conceptual framework for a digital circular built environment: The data pipeline, passport generator, and passport pool. *ICSBE* (Vol. 2, pp. 97–106).
[11] Pal, K. (2023). Assessing the Data Exchange of Material Passports between Key Stakeholders.
[12] Munaro, M. R., & Tavares, S. F. (2021). Materials passport's review: challenges and opportunities toward a circular economy building sector. *Built Environment Project and Asset Management,* 11(4), 767–782. https://doi.org/10.1108/BEPAM-02-2020-0027.
[13] Sesana, M. M., & Salvalai, G. (2018). A review on Building Renovation Passport: Potentialities and barriers on current initiatives. *Energy and Buildings,* 173, 195–205. https://doi.org/10.1016/j.enbuild.2018.05.027.
[14] Hoosain, M. S., Paul, B. S., Raza, S. M., & Ramakrishna, S. (2021). Material passports and circular economy. In *An introduction to circular economy,* 131–158. https://doi.org/10.1007/978-981-15-8510-4_8.
[15] Concular GmbH (2024). Internal source.
[16] Lauble, S., Bonekämper, H., Zielke, P., & Wu B. (2024). Improving Sustainability Assessment in German Building Stock: User Insights and Potential for Artificial Intelligence. *European Conference in Product and Process Modelling 2024.*
[17] Alammar, J. (2018). Visualizing a neural machine translation model (mechanics of seq2seq models with attention). *Visualizing Machine Learning One Concept at a Time Blog.*
[18] Kasneci, E., Sessler, K., Küchemann, S., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. https://doi.org/10.1016/j.lindif.2023.102274.
[19] Jagannathan, M., Roy, D., & Delhi, V. S. K. (2022). Application of NLP-based topic modeling to analyze unstructured text data in annual reports of construction contracting companies. *CSI Transactions on ICT,* 10(2), 97–106. https://doi.org/10.1007/s40012-022-00355-w.
[20] Paneru, S., & Jeelani, I. (2021). Computer vision applications in construction: Current state, opportunities & challenges. *Automation in Construction,* 132, 103940. https://doi.org/10.1016/j.autcon.2021.103940.
[21] Wu, B., & Maalek, R. (2023). Renovation or redevelopment: the case of smart decision-support in aging buildings. *Smart Cities,* 6(4), 1922–1936. https://doi.org/10.3390/smartcities6040089.
[22] Graessler, I., & Hentze, J. (2020). The new V-Model of VDI 2206 and its validation. *Automatisierungstechnik,* 68(5), 312–324. https://doi.org/10.1515/auto-2020-0015.
[23] Regulwar, G. B., Deshmukh, P. R., Tugnayat, R. M., et al. (2010). Variations in V model for software development. *International Journal of Advanced Research in Computer Science,* 1(2), 134–135.