PEDACT

Auto characterization of PEDs for digital references

towards iterative process optimisation

Project Nº: 43927229

Deliverable D5.2

Multi-criteria matrix for the characterization of existing PEDs

Authors: Mengjie Han, Beril Alpagut, Bahanur Nasya

URBAN EUROPE

This project is supported by the European Commission and funded under the Horizon 2020 ERA-NET Cofund scheme under grant agreement N° 875022

Document information Table

Project Data			
Project Acronym	PED-ACT		
Project Title	Auto characterization of PEDs for digital references towards iterative process optimisation		
Project n.	43927229		
Type of Project	Applied Research		
Topic identifier	Topic 2: Making PEDs happen: process innovation and business models		
Project duration	36 months		
Coordinator	Dalarna University, Sweden; Xingxing Zhang		
Deliverable Document Sheet			
Deliverable No.	D5.2		
Deliverable title	Multi-criteria matrix for the characterization of existing PEDs		
Description	D5.2 develops a multi-criteria matrix for the extraction of PED characteristics. The matrix will be used for benchmarking and matching the information of case cities.		
WP No.	5		
Lead beneficiary	Umeå University		
Author(s) & contributor(s)	Mengjie Han, Bahanur Nasya, Beril Alpagut, Juveria Shah, Promise Bassey Sunday, Batuhan Akkaya, and Gireesh Nair		
Type	Report		
Dissemination L.	X		
Due Date	X	Submission Date	X

Version	Date	Author(s)	Organisation	Comments
v1	2024-02	Mengjie Han	Dalarna University	Needs feedback from the PED experts/community
v2	2024-05	Mengjie Han, Beril Alpagut, Bahanur Nasya, Gireesh Nair and Batuhan Akkaya	Dalarna University, Demir Energy, and Wonderland	Ready to be **tested** with tasks and AI
v3	2024-06	Menjgie Han and Bahanur Nasya	Dalarna University and Wonderland	Expert input added

Copyright notice

© 2023 PED-ACT. ALL RIGHTS RESERVED. ANY DUPLICATION OR USE OF OBJECTS SUCH AS DIAGRAMS IN OTHER ELECTRONIC OR PRINTED PUBLICATIONS IS NOT PERMITTED WITHOUT THE AUTHOR'S AGREEMENT.

Disclaimer

This project receives funding under the Joint Programming Initiative (JPI) Urban Europe framework under project number 43927229.

It receives funding support from the strategic innovation program 'Viable Cities', which is financed by Vinnova, the Swedish Energy Agency and Formas (P2022-01000), The Scientific and Technological Research Council of Türkiye, and the Austrian Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the funding agencies involved.

Table of contents

List of Figures

Executive Summary

Deliverable D5.2 focuses on the development of a multi-criteria matrix to characterize PEDs. The deliverable highlights the creation of a matrix to facilitate benchmarking and matching information across case cities. It outlines the methodology for extracting PED characteristics using natural language processing (NLP) techniques by emphasizing the need for an approach to analyse unstructured textual data from various PED project websites and reports, most of which were available in 2023 on PED-EU-NET database. The keywords from NLP were further provided for the public to express their understandings and thoughts, to revise and comments to comprehend and complement the findings given by the NLP algorithms. This public consultation started with feedback loops within the DUT framework, the PED-ACT partner's experts and extended to the PED-EDU-NET network. Even though the process is still open, project and external partners are invited to use and revise the matrix, this version is the third revised version of the matrix. With this process, eight dimensions have been identified: **social**, **process**, **technical**, **governance**, **environmental**, **legal**, **financial**, and **managerial**. The evolving nature of PEDs and the increasing availability of survey data pinpoint necessities the need to further test and expand the matrix. This expansion would allow for a more understanding, precise and useful characterization of PEDs, incorporating new insights, technologies, and practices as they emerge. Thus, enhancing the matrix will enable it to remain relevant for analysing and comparing PEDs, and support their development and implementation in line with the latest advancements and stakeholder needs.

1. Introduction

Since the introduction of the SET-Plan Action 3.2 in 2018, there has been a growing number of PED-focused initiatives and projects at both the European Union and national levels [1]. These projects have contributed to a certain level of comprehension and practical implementation of PEDs on diverse environments. On the other hand, by observing key components of the PED definitions across fourteen distinct programs and initiatives, [2] noted that the main challenges for defining PED pose several primary challenges, including determining the boundaries of PEDs, establishing a method for calculating energy balance, addressing non-energy considerations, and evaluating qualitative requirements. It is also revealed that substantial knowledge gaps and inconsistencies in the PED concept, particularly in the varied definitions used, leading to frequent misinterpretations [2], which implies that different projects are established on different baselines and create a challenge for comparison. Characterization can provide valuable guidance in defining and aligning ongoing projects, facilitating the PED implementation by offering more practical tools and working directions, and allow for the establishment of benchmarks and performance metrics. It is a way for assessing the effectiveness of different strategies and for making improvements in PED replication. Many studies have started to address technical, economic, and social aspects of PED, but very limited study is found on characterizing PED [3]. It is thus necessary to investigate the different dimensions or multi-criteria related to PED

development, implementation, and operation to support broad understanding, successful decision-making, planning, and design.

A multi-criteria matrix is a tool used to organize and compare multiple dimensions of an object to allow for the evaluation of diverse aspects in a structured manner by enabling a comprehensive analysis of different variables. The matrix provides the objective comparison of various options or scenarios and facilitates stakeholder engagement by making the evaluation criteria transparent and allowing for the inclusion of diverse viewpoints. As such useful tool, it needs to cater to the needs of various stakeholders. **Efficient utilization of the matrix will provide a more inclusive and participatory environment** and ensure that a wide array of perspectives and concerns are considered. The different content needs also benchmarking and matching the information in PED learning. The matrix can help the PED project managers, to identify core topics in key areas for the particular PED and by engaging various stakeholders early in the planning process, it can help to build consensus and reduce or prevent potential conflicts, contributing to more community-centric urban solutions.

2. Method

Given the wide-ranging implementation of the PED projects throughout Europe, data generated from projects, most of which are available on the PED-EU-NET database, for extracting the latent criteria often exists in textual and unstructured formats, thus requiring an efficient approach to conduct this task. Qualitative methods have its limitations, primarily stemming from a lack of contextual understanding, which can lead to ambiguity and subjectivity. This method may encounter challenges in accurately interpreting and classifying instances subjective expressions on PED project. Its primary focus on determining the positive, negative, or neutral sentiment of a document may hinder its ability to capture various aspects effectively [4]. The characterization of PED is multifaceted, where a quantitative approach needs to be developed to facilitate scalable and automated analysis.

Topic modelling, as a **n**atural **l**anguage **p**rocessing (NLP) technique, plays a significant role in unveiling latent topics within extensive collection of textual data. It is a flexible and powerful method for recognizing topics and extracting valuable insights from unstructured data that, unlike tabular data, cannot be directly quantified. By using traditional statistical algorithms, topic modelling enables the discovery of underlying themes and patterns that might otherwise remain hidden, thus make it an effective and adaptable tool for gaining insights from large volumes of unstructured textual data [5]. Recently, the progress achieved in pre-trained large language models has enabled topic modelling to capture semantic relationships within data by leveraging the Transformer architecture's structural capabilities [6,7]. This structured approach, which has demonstrated analytical advantages in various application domains, can likewise contribute to a holistic comprehension of the multidimensional aspects inherent in PED (descriptions and reports).

With planning, deployment, and replication of 100 'Positive Energy Districts' by 2025 as a tool for achieving climate-neutral cities [1,8], the exploration of project sources, such as websites and

public documents, has become crucial for developing a common tool or database. Such type of textual data in the form of natural language is largely unstructured. Topic models should go beyond simple word frequency analysis and capture underlying semantic meaning of words and their associations within a document corpus [9], and identify numerous concepts that are embedded in a document collection, enabling the navigation of the collection according to a certain topic [10]. The method has attracted significant research interest in many fields, such as operations research [11], consumer sentiments toward brands [12], health communication [13], climate change [14], and AI applications [15]. It also serves as a powerful tool for understanding the success and growth factors in renewable energy projects [16] through the analysis of a large amount of research articles. By incorporating temporal information, topic modelling can further identity research trend in PED elements, such as green building [17] and renewable energy [18]. This work is an attempt to create a PED related framework for the complex work of auto-characterisation in form of a matrix.

3. Data

3.1. Website data

The data to be analysed to construct the multi-criteria matrix was collected with the objective of gathering relevant textual information from various online sources that are related to PED projects. The online sources are project websites of various PED projects at different stages of implementation. This source was chosen, as websites often provide up-to-date information on diverse categories, thus reflecting the latest trends, research outcomes and timely insights. The downside of such source is that the content is selected by project owner and has mainly or often promotional nature, but at the same time, the source is addressed to a wider audience and has therefore the proper language for the audience groups. Website text data may also include supplementary metadata, attached documents and hyperlinks to provide context-aware features that enrich the modelling process. Based on a compilation by JPI Urban Europe in February 2020, a total of 61 PED projects at different stages of implementation were initially identified. Subsequently, an additional search for relevant projects led to the discovery of 18 more project websites, bringing the total number of identified PED projects to 79. These projects were then filtered down to 52 based on the criterion that the content is in English language, furthermore the criterion of content length narrowed down to include only those websites with more than 200 words. This criterion was chosen to avoid that websites with very few words may not be able to provide sufficient information to capture the diversity of the PED projects and efficiently train the algorithms. Finally, the task could be performed with 33 websites. In addition, 22 PDF files including project description, reports, deliverables, and other useful information were added to the data material for training the algorithms.

3.2. Handling of data

A thorough analysis of textual data necessitates a comprehensive and methodical approach to data pre-processing. This critical initial step refines the data by removing extraneous characters

and non-alphabetic symbols that do not contribute to the analytical value. Furthermore, it addresses any discrepancies that might have emerged during the extraction of data from web sources. The pre-processing phase comprises several key procedures: tokenization, stemming, and lemmatization, each serving a distinct purpose in the text analysis pipeline. Tokenization is a foundational process of segmenting text data, including words and punctuation marks, into discrete entities. This segmentation provides basic elements required for subsequent analytical modelling. Stemming, on the other hand, simplifies words to their core forms by systematically stripping suffixes or affixes, thereby reducing the complexity of the linguistic data. Lemmatization, employing a more nuanced approach, involves a thorough vocabulary and morphological examination of words. This process aims to eliminate inflectional endings and retains the word's essential base form to hold significant semantic value.

In the pre-processing stage, the text corpus was imported and processed utilizing the robust functionalities of the Natural Language Toolkit (NLTK) and Scikit-learn libraries within the Python programming environment. Following the cleaning and tokenization processes, the text fragments were collated into a structured list. This list was then refined by eliminating English stop words, as defined by the NLTK corpus, to remove linguistically redundant elements. Furthermore, a manual exclusion list was created to omit terms that, while frequent across the dataset, contribute minimally to discerning the thematic essence of the texts. Namely, the exclusion list was enhanced by omitting ubiquitous core terms commonly encountered across project websites, including, for example, 'energy', 'districts', 'projects', and 'city'. This deliberate exclusion strategy is aimed at sharpening the analytical process on distinctive terms within the corpus. The final stages of pre-processing involved the execution of stemming and lemmatization techniques to consolidate words to their base forms, so that the dataset was simplified while preserving its semantic integrity.

3.3. Algorithms

Recent advancements in topic modelling algorithms have significantly led to their application in diverse scenarios for uncovering topics related to sustainable energy and urban development. Kumar & Ng's study utilizing LDA and NMF revealed an interesting finding: among the five topics identified in renewable energy projects, four of them were non-financial in nature [16]. This indicates that "softer" factors, such as government policies, public-private collaboration, community support and involvement, fiscal mechanisms and terms, and talent, have played a significant role in driving the growth of renewable energy. Additionally, the analysis of smart city risks revealed eight distinct thematic areas: social, institutional, partnership and resource management, scheduling and execution, financial, technology, political, and environmental [19]. In their research [19], the authors also demonstrated that NMF outperformed other methods as the topics generated by NMF were found to be more interpretable and in better alignment with previous literature.

In light of these developments, two traditional ones, Latent Dirichlet Allocation (LDA) and non-negative matrix factorization (NMF), that have been extensively used as efficient tools for modelling due to their flexibility and applicability [20,21] will be employed for the analysis in D5.2. The output from them is compared and merged to form the basis of the matrix. LDA [22,23] is an

unsupervised technique and a generative probabilistic model that represents documents as mixtures of topics, which enhances traditional method Probabilistic Latent Semantic Analysis (pLSA) by utilizing Dirichlet priors to estimate the distributions of document-topic and term-topic relationships. This probabilistic architecture allows LDA to capture the uncertainty and variability inherent in natural language data. The fundamental assumption for LDA is that documents sharing similar topics tend to employ similar groups of words. Each document is generated by first selecting a distribution over the topics. A topic assignment is then chosen for each word in the document, and the word is selected from the corresponding topic. By doing so, documents are created by a combination of topics, and the choice of words is influenced by the topics assigned to them.

NMF, on the other hand, is a dimensionality reduction technique for multivariate data and the concept of NMF was first introduced under the name positive matrix factorization [24], and became known as non-negative matrix factorization since 1999. Its distinctive feature is the generation of sparse, yet highly interpretable topic representations, thanks to the non-negativity constraint [25]. The high degree of sparsity means that each topic is composed of a few highly relevant words. In addition, NMF is computationally efficient compared to other algorithms since it requires fewer iterations to achieve a reasonable factorization. While NMF shares the same feature as LDA in terms of non-deterministic results, it can be reckoned that NMF can yield superior results. This is primarily because the NMF utilizes term frequency-inverse document frequency (TF-IDF) encoding, which takes into account the importance of words in the corpus rather than relying solely on raw word frequencies [26]. In addition, researchers generally concur that NMF, as a linear algebraic model, performs well with shorter texts [20]. The NLP output will be used as the basis for collecting public feedback and expert input in later stages.

4. Basis for the PED matrix

To form out the analyses a matrix, the values had to be grouped and with subtopics the meaning and importance of each group had to be defined. The modelling results from each of the algorithms was firstly summarized to extract a **topic name**, this naming was based on the top ten word in each topic group. When a topic could be further divided into multiple categories, **sub-topics** were formulated. Thus, topic groups have different hierarchical structures. Next, results from two algorithms were compared to merge the groups and to identify same topics to avoid redundancy. We followed the principle to group semantically similar words into one topic. Last, each of the top-rated words were examined to ensure that it has a real meaning and is consistent with its topic name. The results indicate that seven such topics can be drawn to form the basis of the matrix. The identified main topics are social development, implementation, technology, engagement, environment, policy, and finance. With these topics various subtopics are mapped as shown in Figure 1-7.

4.1. Social development

Social development can be described as the process of increasing the assets and capabilities of individuals to improve their wellbeing. It ~~that~~ emphasizes the different measures and strategies implemented to achieve the intergenerational balance between the economic, environmental, and social factors. This resonates with the PED concept to encourage civic engagement, inclusivity, and citizen participation as a pathway to building cohesive communities.

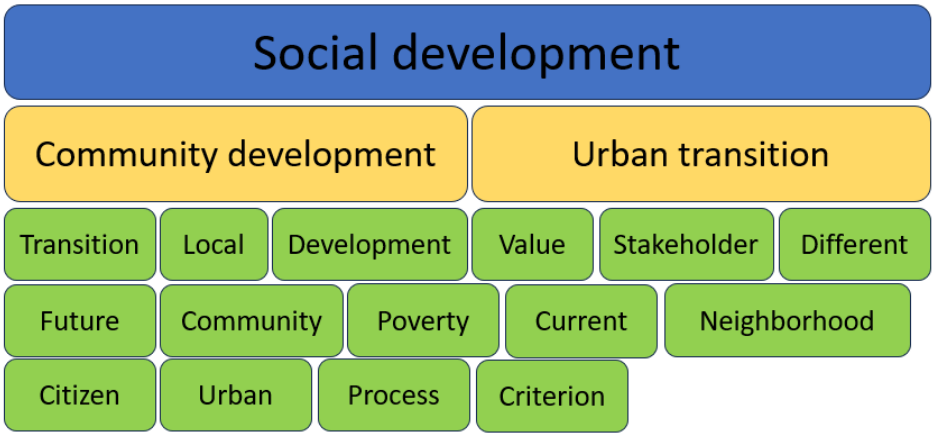

Figure 1: Social development

The sub-topic ***Community development*** is a vital component of social development. It speaks about empowering communities, enhancing their capacity to meet their own needs, and promoting active citizen participation. It recognizes the unique strengths, assets, and diversity within communities and seeks to mobilize these resources for sustainable change. It also entails involving community members in decision-making processes and creating opportunities for collaboration and collective action. Strategies need to be implemented to enhance social connections interaction, within communities. These efforts could contribute to the overall well-being and quality of life for residents within a PED.

Urban transition is the second dimension under social development. This reveals that social development is closely related to the evolution and transformation of urban areas. PEDs directly support the reduction of carbon emissions and energy consumption in urban areas and contribute to global and local climate goals. By generating renewable energy, improving energy efficiency, and enhancing energy flexibility, PEDs help cities move towards carbon neutrality. By reducing energy demand and increasing the share of locally produced renewable energy, PEDs enhance urban resilience against energy price fluctuations, supply disruptions, and the impacts of climate change. The engagement and participation of stakeholders in the transition lead to increased awareness and adoption of sustainable practices, which will create a sense of ownership and responsibility towards future urban environment.

Social development is a significant factor in creating a digital reference for PEDs. The social features, such as community growth and the forthcoming change, should be captured in the digital reference. It will shed light on how PEDs can promote inclusive neighbourhoods as well as how to make use of the social opportunities and address difficulties that present themselves

throughout the course of their growth and transition. Thus, PEDs cannot be solely characterized by their energy performance but should also prioritize the well-being and social structure of their communities.

4.2. Implementation

The ***implementation*** of PED projects is a complex process that requires collaboration between governments, urban planners, developers, businesses, and the community. PEDs can be successfully implemented by adhering to a clearly defined process, interactive construction, and creative innovation.

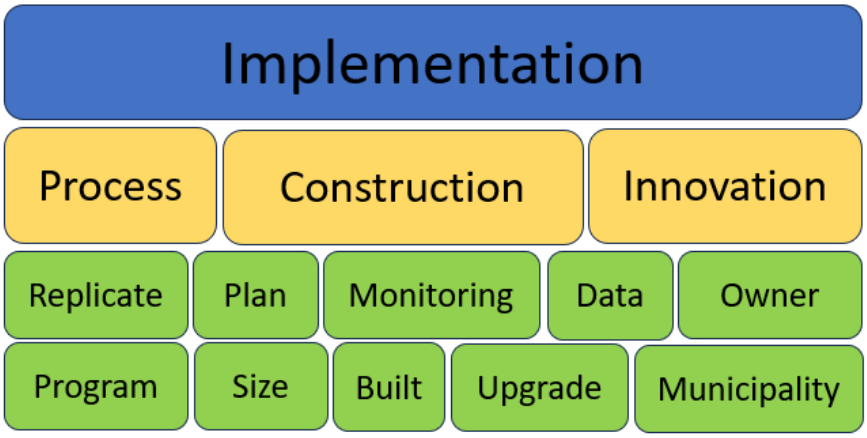

Figure 2: Implementation

Process as a sub-dimension of implementation focuses on systematic approach of adopting a PED across all stages from initial planning to ongoing management and future assessment. For the PED implementation process to be successful, cooperation and coordination between very different actors, expertise, and responsibilities, such as local government agencies, developers, energy providers, technology providers and the community is required. PEDs can promote efficient decision-making, resource allocation, and smooth implementation of urban transition goals by adhering to a clearly defined process that is dynamic and iterative. Optimized collaboration can prevent conflicts and mistakes.

A crucial sub-dimension that deals with the practical application of PEDs is ***construction***. It involves the actual design, development, realisation and installation of a building, the green and smart infrastructure, and the energy systems. The use of sustainable materials and environmental preservation need to be the priorities in construction processes. The PED construction needs to align with the local urban systems and development goals, as well as building regulations. It ensures that PEDs cannot be isolated projects but are part of a larger local culture and strategy to enhance urban connectivity and resilience. This perspective emphasizes the interdependencies between different urban systems, including transportation, water, waste management, and green spaces.

The third dimension is ***innovation***. To attain energy independence and sustainability, new technologies, strategies, and solutions need to be applied in the PED ecosystem. This covers the use of advanced monitoring and control systems, smart grid technology, and the design of energy-efficient buildings. It's crucial to start with a clear vision that addresses local issues in sustainability and resilience. This vision aims to transform urban areas that can inspire and be replicated elsewhere. Innovations often stem from challenging traditional approaches and practices. Envisioning new possibilities for urban living leads to innovative practices. Learning from other successful PED projects and urban sustainability initiatives worldwide can provide valuable insights for or trigger local innovation. Analysing case studies, engaging with international networks, and participating in knowledge exchange and education programs can also help identify innovative practices that could be adapted and implemented in a local context.

4.3. Technology

Technology emphasizes the crucial part in achieving PED goals. These districts leverage advanced technologies to optimize energy generation, distribution, and consumption.

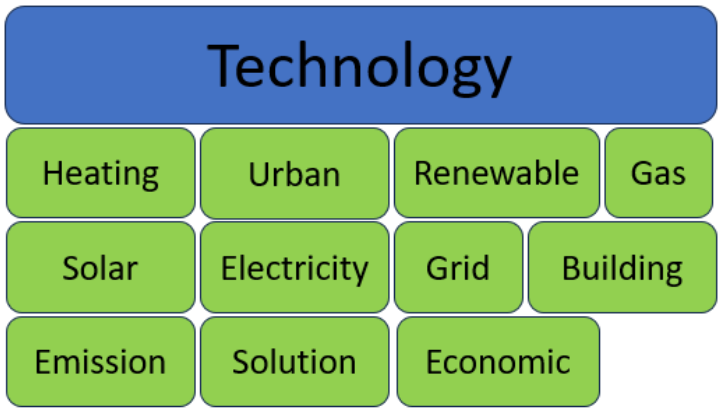

Figure 3: Technology

First, PEDs rely heavily on technologies that improve the **energy efficiency** of buildings and infrastructure. This includes high-performance insulation materials, energy-efficient design, HVAC systems that minimize energy use, and smart appliances that can adjust their operation based on real-time energy production and demand. The application of these technologies reduces the overall energy demand of the district. Second, a core component of PEDs is the **integration of renewable energy technologies** to meet and exceed the district's energy needs. This includes PV panels, solar thermal systems, wind turbines, geothermal systems, biomass, and biogas, etc. Third, **smart grids** equipped with advanced metering, sensors, and control systems allow for real-time monitoring and management of energy flows. It enhances the efficiency and reliability of the energy system. Forth, **energy storage** technologies, including battery storage and thermal storage, can balance supply and demand, storing excess energy produced during peak production times for later use. Last, **digital applications**, including the Internet of Things (IoT), building information and communication technology (ICT), artificial intelligence (AI), and blockchain, are crucial to enhance energy systems within PEDs. These technologies enable the

collection and analysis of vast amounts of data to facilitate smarter energy management decisions and promote energy sharing among residents.

4.4. Engagement

Engagement in PED is to the process of involving a wide range of stakeholders, such as residents, businesses, local authorities, utility companies, and other relevant parties, in the different phases of PED project. It emphasizes that the PED is designed and operated in a way that meets the needs and expectations of the community it serves, while also aligning with broader urban development goals.

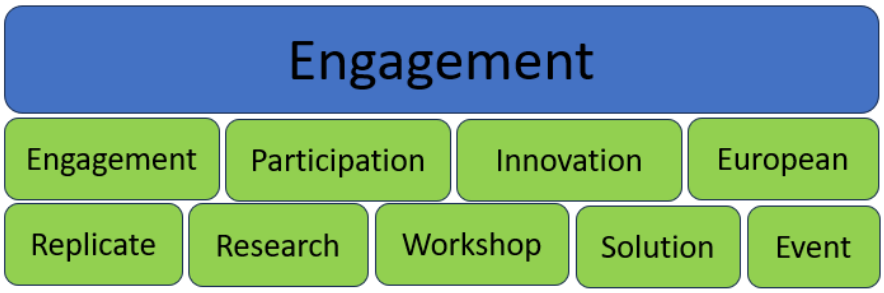

Figure 4: Engagement

An inclusive engagement involves creating opportunities for active participation for all stakeholders. This includes providing accessible information about PED project and potential impacts, possibilities to support the process or be a part of it, as well as creating forums for dialogue and feedback. Communication should be tailored to reach different audiences, using a variety of formats and languages as necessary to ensure inclusivity. Stakeholders should also be empowered in decision-making processes. Workshops, focus groups, social events and public consultations can be set up to allow community members to voice their opinions, offer insights, and contribute to the shaping of the project. These participatory processes should be designed to be welcoming and accessible, with efforts made to reduce barriers to participation such as scheduling conflicts, physical accessibility issues, and language barriers. Local entities often have deep understanding of the community's features and can serve as bridges between the PED project team and the wider community, which helps to mobilize participation.

In the engagement process, transparency is important for building trust and maintaining open two-directional communication. Information of challenges and setbacks should be open, and any decision-making should be well explained. Managers need to keep responsive to feedback and demonstrate how community input can affect project progress.

4.5. Environment

Environment emphasizes the significance of creating sustainable and environmentally friendly districts. PEDs seeks to impact on the environment by reducing energy consumption, particularly through improvements in energy efficiency and the switch to energy sources with lower carbon emissions.

Figure 5: Environment

In addition to energy technologies and material science, a smart and flexible energy management system is also required. By reducing dependency on external energy sources and enhancing green infrastructure, PEDs contribute to the climate resilience of urban areas. Resilience is necessary to be better prepared to withstand and recover from extreme weather events, climate change, and other environmental challenges. Moreover, common goals of district's energy encourage residence behaviours and active involvement to support environmental conservation.

4.6. Policy

Effective *policies* can create the necessary framework and support mechanisms to facilitate the planning, implementation, and scaling of PED initiatives. It should address legal and regulatory support as well as possible financial incentives.

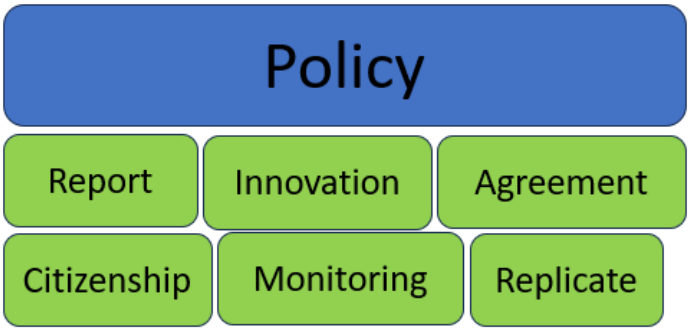

Figure 6: Policy

Without supportive ***policies***, many of the technical and operational innovations required for PEDs might face regulatory obstacles. Thus, appropriate policies encourage the adoption of new technologies and investment models by supporting various tasks in research and innovation. For example, policies that encourage or require the use of renewable energy sources should undergo evaluation in pilot studies prior to their broad implementation in new development areas. Effective policymaking can also establish mechanisms for monitoring and reporting progress to ensuring that PEDs contribute to broader territorial goals. Case studies and models from cities that have successfully implemented PEDs need to be examined. Monitoring and evaluation of PED policies need to be established, where data and feedback can be used to assess the effectiveness of

policy measures and make necessary adjustments. When introducing or revising policies, authorities need to have a comprehensive understanding of the current urban energy landscape and involve a broad range of stakeholders.

4.7. Finance

Finance is a critical factor that can either enable or constrain the successful development and implementation of PEDs. Securing financing is essential to cover initial investments and be able to start a PED project. Often the financial part is viewed as the most important pre-condition to start a PED project.

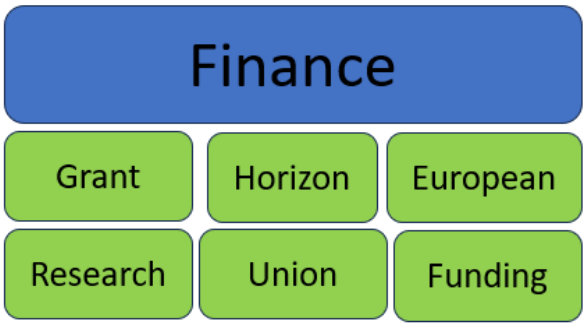

Figure 7: Finance

This PED matrix is derived from data of research projects. That is the reason why the results for this category are predominantly related to European research and innovation, therefore funding programs. Nonetheless, we are assured that the financial considerations are a more general dimension and need to be extended to a broader scope in the context of PED implementation. Financial supports for R&D projects enable the exploration of new ideas, the testing of pilot projects, and the adaptation of technologies to meet the unique needs of a PED.

Finance is not only covering initial costs for implementation, but also concerning long-term economic sustainability. This includes finding models that can sustain the operation and maintenance of the PED. Financial incentives or benefits can also encourage stakeholders to actively participate in PED initiatives. Subsidies, tax deduction, and other benefits can make the PED project more attractive to invest, to develop technologies and, or initiate innovations.

5. Aligning with public reflection

The seven topics as well as their sub-topics are ‘machine-based’ extraction from textual webpage data, which provides a basis for an initial understand of PED dimensions. However, the data reflects extensive information only from different projects and only two NLP algorithms were used. They heed to be further elaborated and commented by human inputs. Thus, a public survey and a series of expert discussion were included to enhance the matrix.

5.1 Mentimeter survey

It is essential to share and discuss the established base for the PED matrix as illustrated in Figures 1-7 with knowledgeable experts and potential users of the PED matrix. The primary topics outlined in these seven figures served as the foundation for the keywords for a public survey, which is to be presented in this section, to gather diverse opinions. The sub-topics (in yellow and green) were omitted in the survey to avoid providing the hint of answers. The initial objective is to guarantee a comprehensive understanding of the criteria's various topics among the public. Subsequently, the second ~~step~~ seeks to enrich and refine the topics by incorporating insights derived from the survey responses. The survey was conducted via Mentimeter[1] within the workshop organized by Expert Support Facility (ESF) of the JPI Urban Europe PED Programme 2 projects. The ESF Workshop was held on 25-26th of January 2024 in Brussels. The main idea for sharing the Mentimeter Survey with all the sister projects of PED-ACT, that were funded under JPI PED Programme 2, was to collect knowledge and experience that is being tested under different PED concept projects and their pilots where new PED developments are being considered. The respondents are experts from seven sister projects of PED-ACT, who have freedom to write free text regarding the keywords with light explanations. Word frequencies for each criterion are presented in Figures 8-14, which are compared to the sub-topics in Figures 1-7.

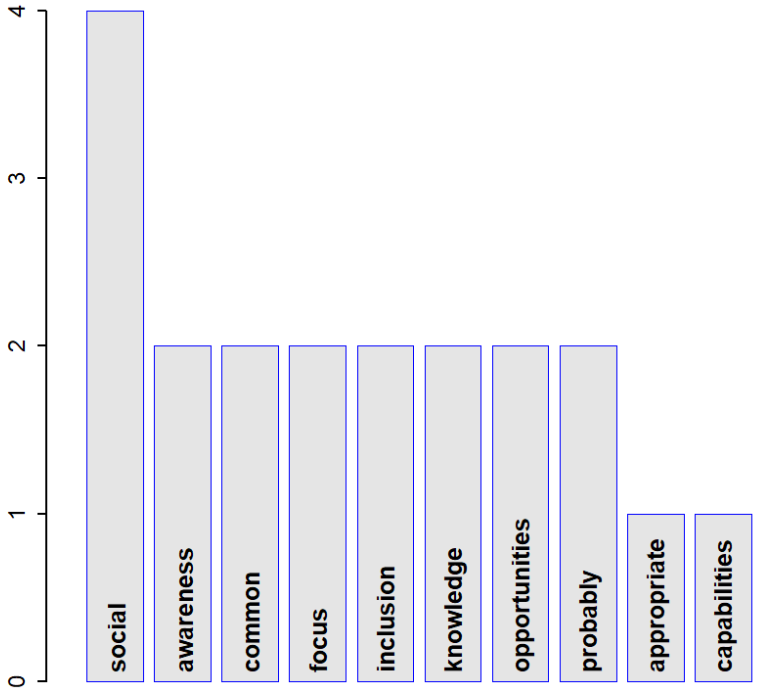

Figure 8: Word frequency of the answers (Question: What are the key points in social development?)

[1] https://www.menti.com/al6egdsjxorb

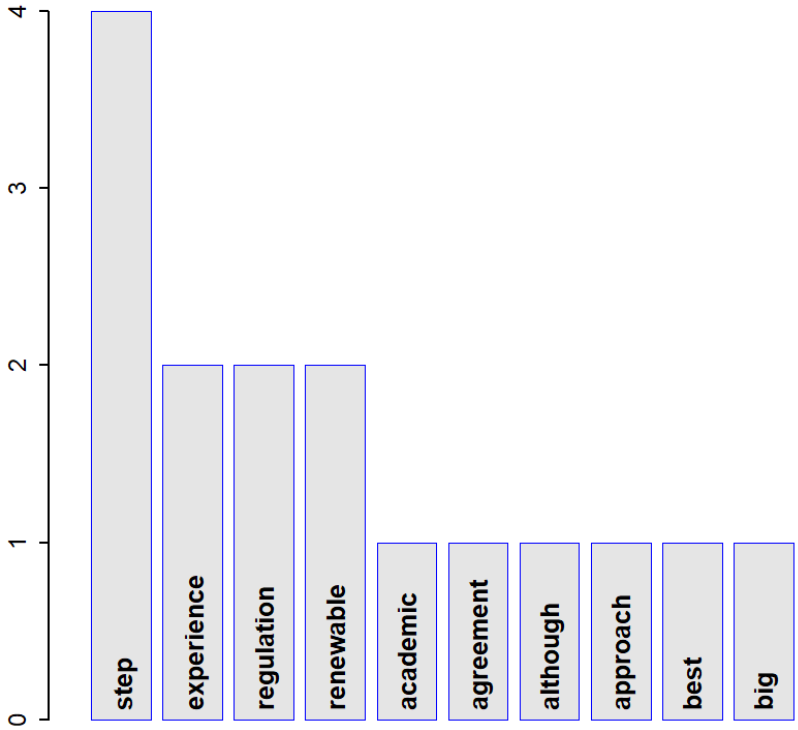

Figure 9: Word frequency of the answers (Question: What are the key points in implementation?)

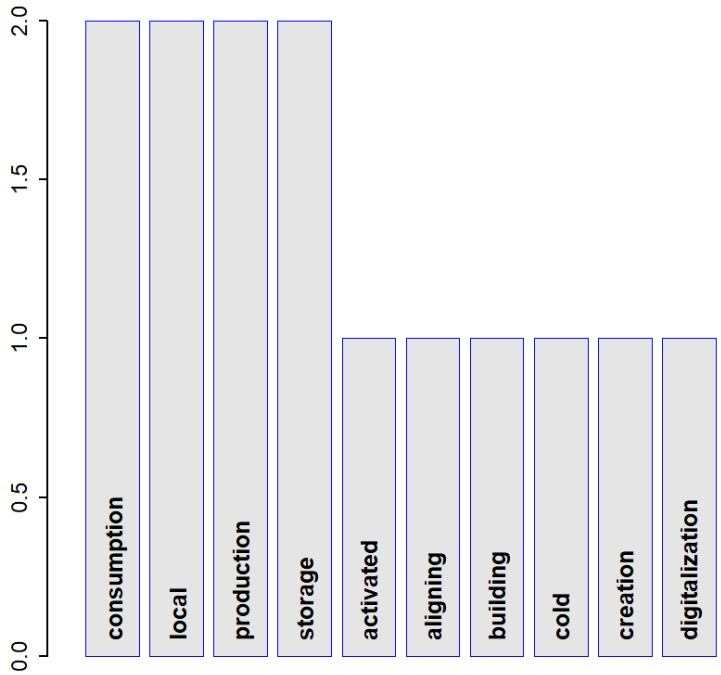

Figure 10: Word frequency of the answers (Question: What are the key points in technologies used in PEDs)

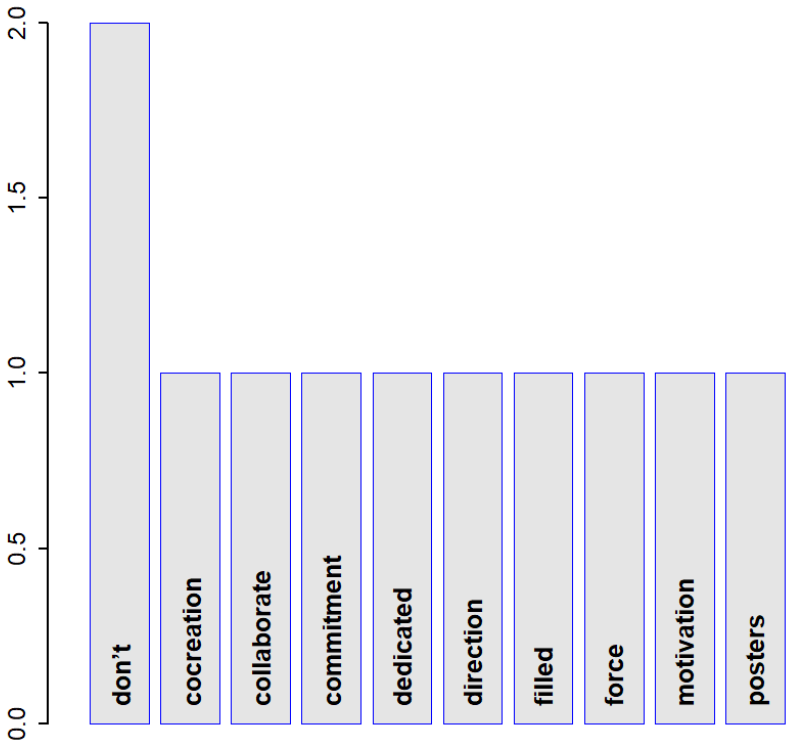

Figure 11: Word frequency of the answers (Question: What are the key points in stakeholder engagement?)

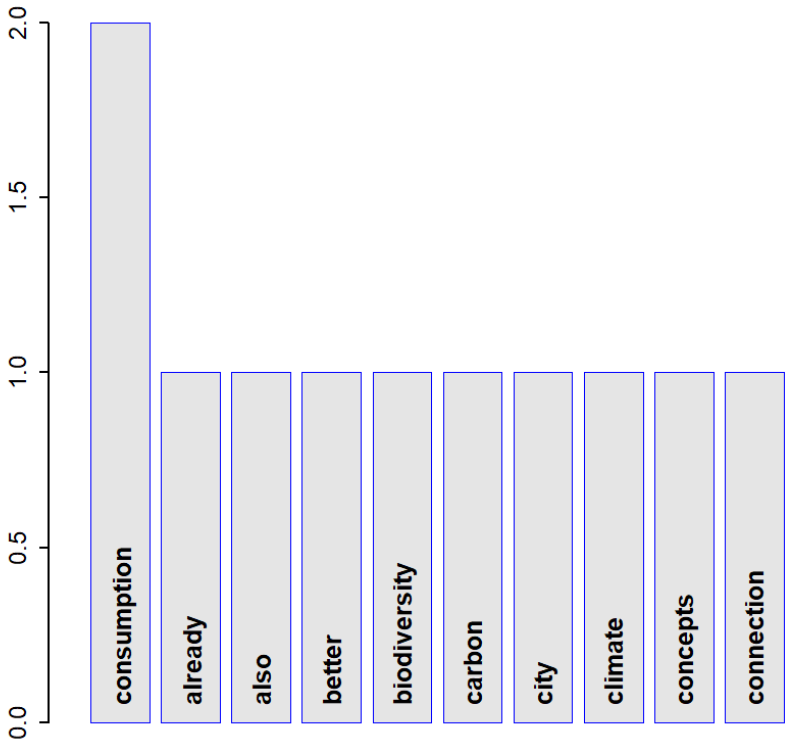

Figure 12: Word frequency of the answers (Question: What are the key points in environmental impact?)

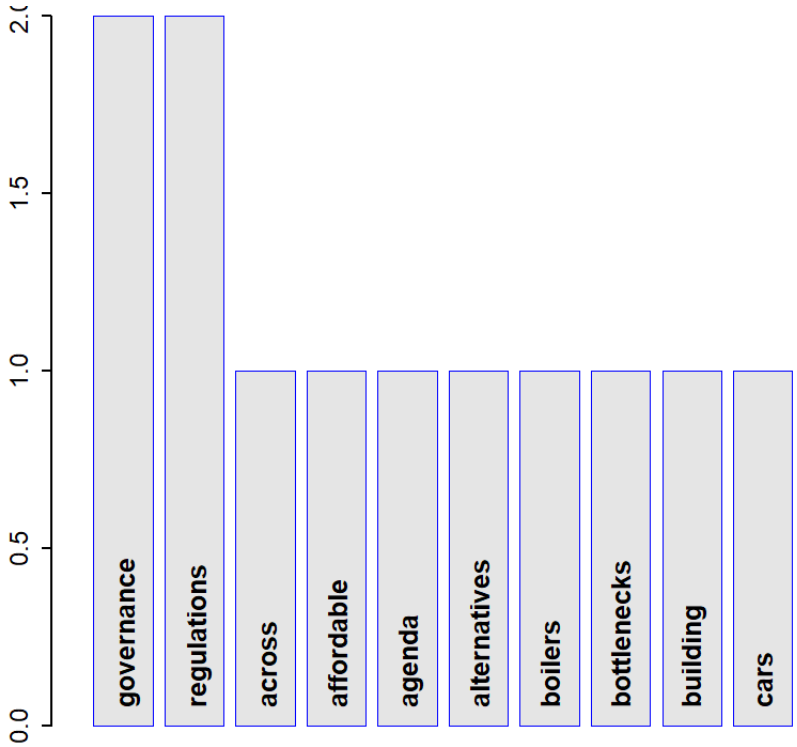

Figure 13: Word frequency of the answers (Question: What are the key points in policies?)

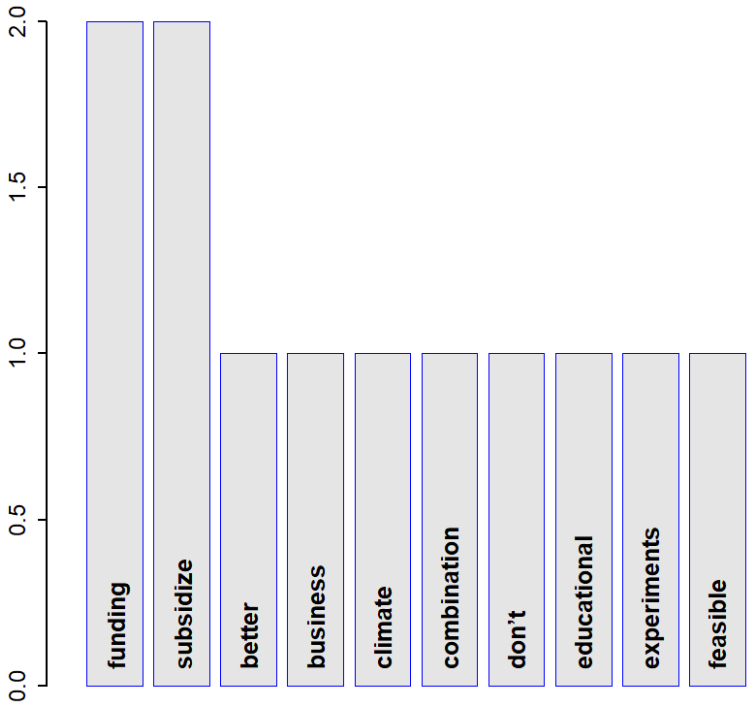

Figure 14: Word frequency of the answers (Question: What are the key points in finance?)

In general, public feedback complements the PED matrix formulated by NLP algorithms. The public demonstrates a robust understanding of the keywords and can contribute further insights to the matrix. The comparison between the keywords in Figures 1-7 and word frequency in Figures 8-14 remains preliminary and more comprehensive expert inputs need to be performed to further refine these insights in Section 5.2. In this D5.2, top-ranked words are analysed.

In Figure 8 for ***social development***, the survey result indicates a few more words: '**awareness'**, '**common'**, '**focus'**, '**inclusion'**, '**knowledge'**, and '**opportunities**. While 'awareness' refers to public understanding about the goals, technologies, practices, and benefits of PEDs, 'common' emphasizes aspects that are shared or jointly utilized by the community within a district, including language and strategy within a community. 'Focus' is the primary goals and strategies that are prioritized from a social perspective. It is also important for the society to be aware of the potential benefits and advantages that PED can offer.

Regarding ***implementation*** in Figure 9, feedback from the public emphasizes the significance of adopting a gradual, step-by-step approach during the transition from one scenario to another. Undertaking overly ambitious steps at once may encounter resistance from the population due to the sheer scale of change or the potential disruption. Consequently, it becomes important to identify and strategize subsequent phases by **employing living labs** and **pilot projects** as pivotal testing grounds within the initiatives. These platforms enable the iterative testing and refinement of concepts, technologies, and strategies in real-world settings to ensure that innovations are both practical and acceptable to the community. This **evidence-based progression** aligns each step with public expectations and contributes to building a solid foundation in the transformation.

in Figure 10, ***technologies*** to be deployed to PEDs should prioritize minimizing energy consumption, harmonizing consumption with production practices, and maximizing the generation of sustainable energy. Local environment should be originally considered for technological advancements and solutions need to be **tailored** to the specific ecosystem of the area.

In terms of ***engagement*** in Figure 11, it is essential not only to involve community members effectively but also to address the challenges associated with individuals who are reluctant to participate. One of the statements in the survey highlights the importance of respecting individual: **'Don't force them into a direction they don't want'[2].** A comprehensive engagement strategy should extend beyond incorporating willing participants. It must also devise thoughtful approaches to understand and address the reasons behind the disengagement of some individuals. This could involve creating more accessible and transparent channels for participation. Furthermore, it is crucial to recognize **different levels of interest and commitment** by offering multiple ways of engagement to allow people align with their comfort and interest levels.

Consumption is still a critical factor for PEDs to significantly influence ***environment***. Addressing consumption may extend to the broader energies, including water, materials, and waste. For

[2] It should be noted that the word frequency was counted by single words. The phrase 'don't force' should be concatenate.

PEDs to realize their full potential in making a positive environmental impact, it is also imperative to implement strategies that promote sustainable consumption behaviours among residents and businesses. This includes, among others, **facilitating a circular economy approach** to optimize resource use. By holistically tackling consumption, PEDs can contribute more effectively to reducing carbon footprints and enhancing resource conservation.

In the ***policy*** dimension, effective governance structures imply clear leadership and accountability that could bring together various stakeholders. Collectively contributing by the stakeholders ensure that the projects align with wider development strategies. Moreover, robust regulation frameworks are essential and necessary to embed guidelines and standards, **as well as incentives** to initiate energy projects. This includes policies that encourage the use of renewable energy, energy efficiency measures, sustainable practices, and the integration of green infrastructure. Regulations can also facilitate the deployment of innovative technologies and solutions and streamline the approval processes for new projects and financial support. Governance and regulatory frameworks should be **adaptive and responsive** to emerging challenges and opportunities. A way to incorporate feedback from pilot projects and living labs need to be prioritized to refine policies over time.

Subsidization acts as a key ***financial*** tool to facilitate PED. By providing financial incentives such as grants, low-interest loans, and tax deduction. Subsidies can lower the initial costs and barriers to implementing technologies and infrastructure upgrades. These financial supports are vital for encouraging private sector participation, as well as facilitating all sorts of innovation. Subsidization can also help cover the costs of research and development to make pioneering solutions more accessible and affordable, and aid urban living labs to demonstrate the viability and benefits to a wider audience. However, for subsidization to be effective and sustainable in the long term, it should be part of a comprehensive financial strategy that includes mechanisms for measuring impacts of subsidized PEDs.

5.2 Input of “PED-ACT experts”

The matrix in D5.2 is the basis for various tasks within the PED-ACT partnership, for example, to develop the PED manifesto, roadmaps, Climate City Contracts, Replication models, and much more within different project cases. For this reason, the “PED-ACTS” need to revise the categories and subcategories, taking into account the voting results from Mentimeter. A first round of survey with Mentimeter (5.1) with project external PED experts involved in PED projects (DUT ecosystem) has collected opinions and focus. Based on that, it is thus necessary to integrate expert input in a second round to reflect comprehensive dimensions and aspects of PED, at same time to include the matrix categories into the tasks, deliverables and outcomes. The “PED-ACT experts” group comprises WP leaders, task leaders, and case partners.

The difference between expert input and Mentimeter survey in Section 5.1 is that experts do not need to follow the original structure, but they can refer to it to freely organize the structure and fill in the elements. It is also necessary to use expert feedback to tweak the results. Thus, by integrating the Mentimeter survey and expert input, a final matrix has been proposed in the

following table. This table with the categories and subcategories will be tested in various PED-ACT tasks as a framework.

Social aspects Community, development, urban transition, awareness, social inclusion, opportunity, diversity, energy poverty, commons, stakeholder/community capacity building, just transition, future proofness, knowledge, focus, quality of life, social capital, trust, expectation (stakeholders, communities, and citizens), heritage protection, work-life balance, gender equality
Process aspects Process, construction, innovation, phase, skills/experience, new development, revitalisation, risk management, process mitigation, dealing with mistakes, sustaining trust, process steps, planning
Technical aspects Consumption, local, production, storage, renewable energy solutions, buildings, technology mix, maintenance, ICT, smart devices, user feedback, data generation, user training/guiding, existing energy infrastructure, new versus existing neighbourhood, protected buildings, decision on system boundaries, energy flexibility measures, passive and active solutions, timeframe and energy balances, existing technological regimes and path dependency, research
Governance aspects Co-creation, commitment, motivation, participation, replicate/scale up, political support/stewardship, decision-making, rights, ownership, definition of roles, awareness of governance structure, innovation potential, ROI on engagement, feedback loops for complaints, public support, transparency, resource allocation, climate goals, procurement policies
Environmental aspects Energy/resource consumption, efficiency improvement, carbon emissions, zero emission, climate neutrality, sufficiency, circular economy, sustainable behaviour, geographical boundaries, unions, sectors, clusters, climate adaptation, embodied energy, energy mix, sustainable mobility, land use, life cycle assessment, resilience
Legal aspects Governance, regulations, monitoring, agreement, ownership, security, funding (eligibility), IPR, procurement, ethics, GDPR
Financial aspects Funding, research, subsidies, business models, ROI (return on investment), non-financial resources, investment, crowd funding, taxes
Managerial aspects Entrepreneurship, organizational capacity building/mentoring learning/training, team building, sustaining capacities, sustaining maintenance, awareness creation

6. Conclusions

Deliverable D5.2 focused on the development of a multi-criteria matrix for characterizing PEDs. The utilised methodology proved to deliver useful outcomes. The extraction of PED characteristics using natural language processing (NLP) techniques allowed the analyses of the unstructured textual data from various PED project websites and reports. A round of manual selection of topics helped to create an initial version in seven categories and subdimensions as described in chapter 4. The dimensions were further provided as keywords for the public to express their understandings and thoughts, and comments to comprehend and complement the findings given by the NLP algorithms. Expert input was finally integrated to extend the matrix and align it with the experts tasks and work, so that the PED matrix can be utilised in different ways of an energy transition project.

Eight dimensions have been finally identified: *social*, *process*, *technical*, *governance*, *environmental*, *legal*, *financial*, and *managerial.* This deliverable highlights the creation of a matrix to map typical characteristics of a PED project and facilitate comparison, benchmarking and matching information across case cities. The evolving nature of PEDs and the increasing availability of survey data pinpoint the need to further expand the PED matrix. This expansion would allow for a more understanding and characterization of PEDs, incorporating new insights, technologies, and practices as they emerge. Thus, enhancing the matrix will enable it to remain relevant for analysing and comparing PEDs, and support their development and implementation in line with the latest advancements and stakeholder needs.

As a follow up action, the PED matrix will be used in different tasks in the PED-ACT project. This is a crucial step to validate the accuracy and importance of the different aspects. As a next step, it will also be shared for feedback and use with other PED-Actors. Potential changes in this process will lead to an update of this deliverable and create the foundation for more research and evaluation.

References

1. *SET-Plan ACTION N°3.2 Implementation Plan*; European Commission, 2018;
2. Albert-Seifried, V.; Murauskaite, L.; Massa, G.; Aelenei, L.; Baer, D.; Krangsås, S.G.; Alpagut, B.; Mutule, A.; Pokorny, N.; Vandevyvere, H. Definitions of Positive Energy Districts: A Review of the Status Quo and Challenges. *Sustainability in Energy and Buildings 2021* **2022**, 493–506.
3. Zhang, X.; Penaka, S.; Giriraj, S.; Sánchez, M.; Civiero, P.; Vandevyvere, H. Characterizing Positive Energy District (PED) through a Preliminary Review of 60 Existing Projects in Europe. *Buildings* **2021**, *11*, 318, doi:10.3390/buildings11080318.
4. Rambocas, M.; Pacheco, B.G. Online Sentiment Analysis in Marketing Research: A Review. *JRIM* **2018**, *12*, 146–163, doi:10.1108/JRIM-05-2017-0030.
5. Kuhn, K.D. Using Structural Topic Modeling to Identify Latent Topics and Trends in Aviation Incident Reports. *Transportation Research Part C: Emerging Technologies* **2018**, *87*, 105–122, doi:10.1016/j.trc.2017.12.018.
6. Praveen, S.V.; Vajrobol, V. Understanding the Perceptions of Healthcare Researchers Regarding ChatGPT: A Study Based on Bidirectional Encoder Representation from Transformers (BERT) Sentiment Analysis and Topic Modeling. *Ann Biomed Eng* **2023**, doi:10.1007/s10439-023-03222-0.

7. Grootendorst, M. BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure. *arXiv:2203.05794v1 [cs.CL]* **2022**.
8. Bossi, S.; Gollner, C.; Theierling, S. Towards 100 Positive Energy Districts in Europe: Preliminary Data Analysis of 61 European Cases. *Energies* **2020**, *13*, 6083, doi:10.3390/en13226083.
9. Uys, J.W.; Du Preez, N.D.; Uys, E.W. Leveraging Unstructured Information Using Topic Modelling. In Proceedings of the PICMET '08 - 2008 Portland International Conference on Management of Engineering & Technology; IEEE: Cape Town, South Africa, July 2008; pp. 955–961.
10. Blei, D.M.; Lafferty, J.D. Dynamic Topic Models. In Proceedings of the Proceedings of the 23rd international conference on Machine learning - ICML '06; ACM Press: Pittsburgh, Pennsylvania, 2006; pp. 113–120.
11. Gatti, C.J.; Brooks, J.D.; Nurre, S.G. A Historical Analysis of the Field of OR/MS Using Topic Models. **2015**, doi:10.48550/ARXIV.1510.05154.
12. Liu, X.; Burns, A.C.; Hou, Y. An Investigation of Brand-Related User-Generated Content on Twitter. *Journal of Advertising* **2017**, *46*, 236–247, doi:10.1080/00913367.2017.1297273.
13. Liu, Q.; Zheng, Z.; Zheng, J.; Chen, Q.; Liu, G.; Chen, S.; Chu, B.; Zhu, H.; Akinwunmi, B.; Huang, J.; et al. Health Communication Through News Media During the Early Stage of the COVID-19 Outbreak in China: Digital Topic Modeling Approach. *J Med Internet Res* **2020**, *22*, e19118, doi:10.2196/19118.
14. Dahal, B.; Kumar, S.A.P.; Li, Z. Topic Modeling and Sentiment Analysis of Global Climate Change Tweets. *Soc. Netw. Anal. Min.* **2019**, *9*, 24, doi:10.1007/s13278-019-0568-8.
15. Mustak, M.; Salminen, J.; Plé, L.; Wirtz, J. Artificial Intelligence in Marketing: Topic Modeling, Scientometric Analysis, and Research Agenda. *Journal of Business Research* **2021**, *124*, 389–404, doi:10.1016/j.jbusres.2020.10.044.
16. Kumar, M.; Ng, J. Using Text Mining and Topic Modelling to Understand Success and Growth Factors in Global Renewable Energy Projects. *Renewable Energy Focus* **2022**, *42*, 211–220, doi:10.1016/j.ref.2022.06.010.
17. Wu, Z.; He, Q.; Chen, Q.; Xue, H.; Li, S. A Topical Network Based Analysis and Visualization of Global Research Trends on Green Building from 1990 to 2020. *Journal of Cleaner Production* **2021**, *320*, 128818, doi:10.1016/j.jclepro.2021.128818.
18. Jiang, H.; Qiang, M.; Lin, P. A Topic Modeling Based Bibliometric Exploration of Hydropower Research. *Renewable and Sustainable Energy Reviews* **2016**, *57*, 226–237, doi:10.1016/j.rser.2015.12.194.
19. Gupta, K.; Zhang, W.; Hall, R.P. Risk Priorities and Their Co-Occurrences in Smart City Project Implementation: Evidence from India's Smart Cities Mission (SCM). *Environment and Planning B: Urban Analytics and City Science* **2021**, *48*, 880–894, doi:10.1177/2399808320907607.
20. Chen, Y.; Zhang, H.; Liu, R.; Ye, Z.; Lin, J. Experimental Explorations on Short Text Topic Mining between LDA and NMF Based Schemes. *Knowledge-Based Systems* **2019**, *163*, 1–13, doi:10.1016/j.knosys.2018.08.011.
21. Mifrah, S. Topic Modeling Coherence: A Comparative Study between LDA and NMF Models Using COVID'19 Corpus. *IJATCSE* **2020**, *9*, 5756–5761, doi:10.30534/ijatcse/2020/231942020.
22. Blei, D.M.; NG, A.Y.; Jordan, M.I. Latent Dirichlet Allocation. *Journal of Machine Learning Research* **2003**, *3*, 993–1022.
23. Blei, D.M. Probabilistic Topic Models. *Commun. ACM* **2012**, *55*, 77–84, doi:10.1145/2133806.2133826.
24. Paatero, P.; Tapper, U. Positive Matrix Factorization: A Non-Negative Factor Model with Optimal Utilization of Error Estimates of Data Values. *Environmetrics* **1994**, *5*, 111–126, doi:10.1002/env.3170050203.
25. Wang, J.; Zhang, X.-L. Deep NMF Topic Modeling. *Neurocomputing* **2023**, *515*, 157–173, doi:10.1016/j.neucom.2022.10.002.

26. Albalawi, R.; Yeap, T.H.; Benyoucef, M. Using Topic Modeling Methods for Short-Text Data: A Comparative Analysis. *Front. Artif. Intell.* **2020**, *3*, 42, doi:10.3389/frai.2020.00042.