ONLINE FIRST ARTICLE
doi.org/10.36253/ijas-16758

Statistica Applicata - Italian Journal of Applied Statistics

Received: 20.02.24/ **Accepted:** 09.06.25
Published online: 02.09.25

WATER FUTURES THROUGH A GENDER LENS: RETHINKING LOCAL PATHWAYS TO SUSTAINABILITY

Alessia Forciniti [1], **Emma Zavarrone, Martha Friel**
Department of Humanities, IULM University, Milan, Italy

Abstract *This study investigates gendered perspectives in the context of Nature-Based Solutions (NBS), focusing on local water management in Milan. A sample of audio-video interviews with governance stakeholders was transformed into a textual corpus, annotated with metadata on speaker roles and gender. A combination of natural language processing techniques and social network analysis was employed to identify thematic groups and recursive semantic patterns. Results revealed clear gender-based differences: women emphasize citizen participation, collaboration with science and culture, and focus on problem definition, viewing sustainable development as a gradual transformative process. Men, in contrast, prioritize specialized skills, territorial authority, and a circular flow of sustainable actions involving multiple actors. Key themes identified include urban territory, industries, politics, and civic participation, highlighting how gender shapes strategic approaches to local sustainable development.*

Keywords: *Sustainability, Gender-based Management, NLP, Machine Learning, Social Network Analysis*

1. INTRODUCTION

One of the most widely recognized definitions of sustainability - as "a means of achieving intergenerational equity" - was introduced in 1987 by the Brundtland Commission (WCED, World Commission on Environment and Development, 1987) in the report Our Common Future. This conceptualization goes beyond merely reducing poverty or improving production and consumption models. It embodies a holistic vision of sustainability as a dynamic concept that embraces the complex interconnections between humans, the environment, and economic systems, emphasizing technological innovation and ecological transition. This paradigm shift has led governments and institutions, at both national and international levels, to implement a blueprint for a better and more sustainable future for all. A clear definition of sustainability has been incorporated into a series of goals, initiatives, and regulations aimed at addressing critical challenges related

[1] Alessia Forciniti: alessia.forciniti@iulm.it (Corresponding author)

to sustainable development. The main objectives are encapsulated in the United Nations (UN) Agenda 2030 (United Nations, 2015), which defines 17 Sustainable Development Goals (SDGs) integrating several interconnected areas of interest. Adopted by all UN Member States in 2015, the SDGs represent a global commitment to improving quality of life while preserving the planet and maintaining the balance between social, economic, and environmental dimensions. The SDGs call for rapid and transformative change to mitigate the effects of unsustainable practices - such as urban expansion, pollution, and environmental degradation - which contribute to global warming, climate change, depletion of natural resources, and disruptions in the hydrological cycle, leading to extreme events such as floods and droughts.

A strong relationship emerges between the SDGs and ecosystem-based approaches, which have gained prominence as effective complements or, in some cases, alternatives to conventional technical strategies (Eggermont et al., 2015). These approaches draw on the functions and services of natural ecosystems to address global challenges in a sustainable manner, including within the framework of Nature-Based Solutions, which use natural processes to manage environmental challenges such as water scarcity, pollution, and flooding. Examples include mitigating climate change by reducing greenhouse gas emissions and increasing resilience to climate-related disasters (SDG 13); conserving biodiversity through the establishment of protected areas and the restoration of degraded ecosystems (SDG 15); sustainably managing water resources by improving water quality, reducing scarcity, and mitigating flood risks (SDG 6); and fostering local economic opportunities through initiatives such as ecotourism and sustainable agriculture (SDG 8; SDG 11). Moreover, by enhancing access to clean air, water, and green spaces, ecosystem-based approaches can also improve human health and well-being (SDG 3). Beyond environmental benefits, these approaches can also address social challenges. Ecosystem-oriented strategies generate co-benefits across ecological and social domains, aligning with broader debates on sustainability and equity (Bocci and Mishra, 2021; Leone, 2019). In this regard, the interplay between ecological functions and gendered practices further highlights both opportunities and challenges for sustainable governance. Involving women and other underrepresented groups in decision-making and ensuring equitable access to resources can help dismantle gender barriers, build more inclusive and resilient communities, and simultaneously promote gender equality, economic development, and social justice, fostering a more balanced and equitable society. Establishing a connection between sustainability theory and local governance is essential to un-

derstanding how policy and management decisions are operationalized through context-specific practices, such as those proposed in this study within Milan's local water management system. Specifically, this research explores how the interlinkages among multiple SDGs are reflected in, and supported by, the application of Nature-Based Solutions within the metropolitan context of Milan.

2. NATURE-BASED SOLUTIONS, GENDERED GOVERNANCE AND RESEARCH QUESTIONS

In the literature, *Nature-Based Solutions* (NBS), introduced by the European Commission, are defined as "solutions that are inspired and supported by nature, which are cost-effective, simultaneously provide environmental, social, and economic benefits, and help build resilience. Such solutions bring more, and more diverse, nature and natural features and processes into cities, landscapes, and seascapes, through locally adapted, resource-efficient, and systemic interventions" (European Commission, 2023).
Despite their growing relevance, the implementation of NBS still faces significant challenges due to governance barriers, policy frameworks, and local contextual issues (Ershad Sarabi et al., 2019). While EU directives encourage NBS integration, national and municipal applications remain uneven, revealing governance gaps that limit their effectiveness in addressing urban sustainability (Zingraff-Hamed et al., 2020). One aspect considered crucial is the co-creation and/or co-participation process (Wamsler et al., 2020) by various actors with different competencies, which the literature identifies as useful for developing an "ecology of knowledge".
Structural inequalities often constrain meaningful involvement in decision-making, particularly regarding gender. Disparities in participation and leadership roles can weaken local management processes and sustainable resource governance. This issue is particularly evident in the underrepresentation of women in sustainable development practices, even though gender equality, as embodied in SDG 5, emphasizes the interconnectedness of cultural, symbolic, social, and economic developments, increasingly recognizing the critical role of women's empowerment in achieving comprehensive and lasting progress. Furthermore, women's participation in governance represents an essential dimension for a democratic society (Yadav, 2023). Numerous studies confirm that integrating women into decision-making structures enhances the inclusiveness, transparency, and long-term effectiveness of governance systems (Cornwall and Rivas, 2015; World Bank, 2019). Countries with higher gender equality tend to show stronger economic growth,

more stable institutions, and reduced corruption (Dollar et al., 2001). Women's participation is also associated with improved environmental sustainability, as they frequently prioritize conservation, social welfare, and intergenerational equity (Chattopadhyay and Duflo, 2004; Nellemann et al., 2011). Nevertheless, women's contributions have historically been undervalued, producing persistent gaps in both policy and practice (UN Women, 2020).

Building on the broader context of NBS, this study specifically focuses on governance aspects and gender inclusion in the managerial practices of designing and implementing NBS in the Italian context, with a particular emphasis on the city of Milan. By examining how women participate in decision-making processes, contribute to policy development, and influence the integration of ecological, social, and economic objectives, the research aims to provide insights into the interplay between inclusive governance and the effective deployment of NBS in urban environmental management. Such a focus allows for a better understanding of how gender-sensitive governance can enhance both sustainability outcomes and social equity.

By analyzing the dialogue with local policymakers, the paper maps whether and through which visions the presence of women influences governance perspectives in terms of objectives, actions, and targets. This study examines the dynamics of women's empowerment within local governance, presenting a case study focused on integrated water resource management (IWRM) tasks. In the Italian context of Lombardy, IWRM is regulated through legislation such as Regional Law 15 March 2016, No. 4 (Regione Lombardia, 2015), and Regional Regulation No. 7, 2017 (Regione Lombardia, 2017), which were enacted to address the hydrological impacts of land-use changes, emphasizing hydraulic and hydrological conservation principles. At the same time, various organizations have emerged to support municipalities in implementing these measures (Ershad Sarabi et al., 2019; Zingraff-Hamed et al., 2020).

Given the growing attention in the literature to gendered governance and intersectional participation, the research questions were designed to capture both individual and relational dimensions of gender in decision-making processes. The paper aims at examining how gendered visions influence managerial practices and how interactions between men and women shape collaborative strategies, institutional behavior, and social dynamics within NBS governance. Moreover, the underrepresentation of women in leadership roles-an empirical reflection of the real governance structure in the Milan metropolitan area-should be acknowledged as an inherent contextual limitation rather than a design bias. Although this study adopts

a binary categorization of gender (women and men), reflecting the composition of the collected data and the structure of the sample, it recognizes that gender is a socially constructed and potentially non-binary dimension. This awareness aligns the theoretical framework of the study with contemporary gender research while maintaining methodological consistency with the available empirical material.

Accordingly, the research is guided by two main questions:

- RQ_1: What are the key managerial directions of women involved in developing NBS for water management in Milan?
- RQ_2: How do gendered perspectives interact in decision-making and governance processes-considering male-male, female-male, and female-female dynamics-and how do these intersections relate to broader social and organizational factors at the local level?

To explore gender-based conceptual nuances, the study analyzed a corpus of interviews on local water management through a combined application of Natural Language Processing and Social Network Analysis. This integrative approach made it possible to identify thematic clusters and recurrent linguistic structures that reveal the semantic patterns underlying managerial visions of sustainability.

Section 2 describes the methodological approach; Section 3 presents data collection; Section 4 presents the findings; Section 5 discusses the results and concludes.

3. METHODS

A multi-step model was developed, and the methodological approach is outlined below.

1. **Step 1:** *Transformation of audio-videos in textual data collection.*
   An automatic transcription service called *Happyscribe* was used to convert the interviews into text. This tool offers an accuracy rate between 85% and 99% and supports 120 languages. In our case, it was used for Italian language transcription.
   Additionally, a manual metadata annotation process was performed for each transcribed interview. Specifically, metadata including gender, professional role, and thematic macro-category (innovation, ecological transition, people, environment), was manually annotated for each transcript as structured attributes. This coding enabled the cross-referencing of linguistic patterns

with socio-demographic and organizational variables, thereby allowing the integration of text-based semantics and relational structures within the governance network.

2. **Step 2:** *Gender-based sub-corpora*.
   Based on the metadata annotation described above, the second step relied on the structured attributes to conduct gender-based filtering of the texts. This information made it possible to separate the corpus into two sub-corpora: one composed of dialogues with male governance representatives and the other with female governance representatives.

3. **Step 3:** *Textual Analysis and Social Network Analysis.*
   In the third step, Textual Analysis is combined with Social Network Analysis (SNA) (Wasserman and Faust, 1994). More specifically, Textual Analysis is the preliminary step for both network analysis and the subsequent word embedding.
   Before performing the network and embedding procedures, all interview transcripts were pre-processed (i.e., (Benzécri, 1981; Bolasco, 2013)) to ensure textual uniformity and remove linguistic noise. The pre-processing pipeline included tokenization, lowercasing, lemmatization, and the removal of punctuation, numerical tokens, and Italian stopwords (ISO 639-1). A domain-specific stopword list was also created to filter high-frequency functional terms that did not contribute to semantic differentiation and were commonly used as part of questions and answers (e.g., the pillars: innovation, people, environment, ecological_transition, Milan). After the text preprocessing phase, a frequency threshold of four occurrences ($n_i = 4$) was adopted for vocabulary extraction. This threshold was selected to exclude extremely rare words that may introduce noise or represent idiosyncratic expressions without semantic relevance, while retaining terms that occur frequently enough to contribute to stable and interpretable co-occurrence patterns. The minimum frequency is adopted to balance vocabulary coverage and model reliability, ensuring that meaningful lexical units are represented without inflating sparsity in the term matrix. The lexical table - constructed from the extraction of the vocabulary of the most frequent terms - called **Document Term Matrix (DTM)**, was realized based on a term-frequency weighting scheme, where n_{ij} counts the number of times the term i^{th} *(i = 1, ..., q)* appears in the document j^{th} *(j = 1, ..., p)*.
   The pre-processing protocol was implemented prior to both the clustering

and embedding stages, following identical criteria settings across procedures. Maintaining methodological consistency in text pre-processing ensured the comparability and integration of the lexical, semantic, and network-based analyses.

After the textual procedure, to detect co-occurrence structures and relational patterns among concepts, SNA was performed. Instead of relying on the more common methods of topic modelling used in literature (i.e.,(Blei et al., 2003; Hofmann, 1990)), some studies have found that the textual network analysis can be very effective in determining thematic groups can be very effective (Gerlach et al., 2018; Paolillo and Forciniti, 2021).

To this aim, the **DTM** was converted into a binary matrix **T**, where the generic element t_{ij} equals 1 if the i^{th} term appears in the j^{th} document, and 0 otherwise. This approach removes the weight of relationships between words and texts, allowing for feature extraction based solely on the presence or absence of words in the texts.

To detect recurrent pairs of terms appearing side by side in a certain order as semantic trajectories, the **T** matrix was multiplied by its transpose, $\mathbf{TT}^T$, obtaining a new *term x termx matrix*. This co-occurrence matrix, called **A**, tracks each couple of terms, as a relationship.

The **A** matrix recalls the properties of SNA, where **A** represents an adjacency matrix with terms in both rows and columns, and the cells that record the relationship between pairs of terms. In our case of a binary matrix, the adjacency matrix simply shows which terms are related and which are not, without weighting the strength of this relationship. The adjacency matrix in SNA may be formally represented by a graph G (V, E), where V is a finite set of nodes (represented by the words), and E is a finite set of linkages (Borgatti et al., 2009) (represented by the co-occurrences).

To delve deeper into the structure of relationships between terms and to identify sub-groups of densely interconnected nodes that are poorly connected to other parts of the network (Fortunato, 2010; Newman, 2006; Newman and Girvan, 2004) (often interpreted as thematic groups), *community detection* methods are often suggested in literature. Community detection is useful to find groups of nodes called communities, clusters or modules that share similar properties or functions. The literature offers diverse community detection algorithms (Clauset et al., 2004; Lancichinetti and Fortunato, 2011; Newman, 2006; Newman and Girvan, 2004; Zhou et al., 2017), each with its own strengths and weaknesses depending on the specific dataset.

Commonly used community detection algorithms often rely on optimization criteria.
In this study, modularity-based methods were considered for optimization. These methods aim to maximize the number of edges within communities while minimizing the number of edges between them. Q is a widely-used metric to evaluate the modularity. Proposed by Newman and Girvan (2004) (Newman and Girvan, 2004), Q optimizes group divisions, leading to a hierarchical structure. Its value quantifies the extent to which a network can be divided into distinct, non-overlapping communities, calculated as:

$$Q = \frac{1}{2h} \sum_{i,i'} \left[a_{ii'} - \frac{k_i k_{i'}}{2h} \right] \delta(c_i, c_{i'}) \tag{1}$$

Suppose to divide the adjacency matrix $\mathbf{A}$ into two communities, the total number of edges in the network is denoted by h, the degree of vertex i by k_i, and the delta of membership value of term i to community c_i by δ.
Q value ranges from 0 to 1. A value of $Q = 0$ indicates a random network structure, while a value closer to 1 signifies a strong community structure. Typically, empirical values of Q fall between 0.3 and 0.7 (Newman and Girvan, 2004).
In our model, the *fast-greedy* algorithm was employed to achieve the best possible Q (Clauset et al., 2004). Fast-greedy is as an agglomerative hierarchical clustering method for very large networks. Its main advantage lies in its stopping criteria which determines the number of groups based on optimization of Q (Newman, 2006; Newman and Girvan, 2004).
For our research objectives, the algorithm was separately performed on two gender-based co-occurrence matrices.
The combined use of the approaches based on textual and social network analysis, presented in step 3, has allowed RQ_1 to be answered.

4. **Step 4:** *Word embedding*.
Recent advancements in language modeling have enabled the learning of dense vector representations of words, capturing fine-grained semantic and syntactic regularities through vector arithmetic. Semantic vector space models of language find space in various applications, including information retrieval, document classification, question answering, named entity recognition ((Pennington et al., 2014). In this paper, Natural Language Process-

ing (NLP) strategies were adopted. A word embedding approach was performed to extract linguistic and semantic regularities from sub-corpora to investigate how the gender influences the managerial perspectives of stakeholders involved in local governance.
An unsupervised learning algorithm called *GloVe* (Global Vectors for Word Representation) (Pennington et al., 2014) was implemented.
For the embedding phase, the GloVe algorithm was applied to the processed corpus to model semantic relationships among terms. In model, training is performed on aggregated global word-word co-occurrence statistics from a corpus, and the resulting representations showcase interesting linear substructures of the word vector space. GloVe combines the strengths of two main model families for learning word vectors: a) global matrix factorization methods, such as Latent Semantic Analysis (LSA) (Deerwester et al., 1990), and b) local context window methods, such as the skip-gram model (Mikolov et al., 2013), which produce the multi-clustering idea of distributed representations(Bengio et al., 2003). In this model, an analogy structure can be encoded as follows:

e.g., vector ("King") - vector("Queen") = vector("Man") - vector ("Woman").

Current methods face significant limitations. While methods like LSA effectively utilize statistical information, they struggle with word analogy tasks, suggesting a suboptimal vector space structure. On the other hand, methods like skip-gram excel at analogy tasks but underutilize global co-occurrence statistics due to their reliance on local context windows.
In this framework, GloVe uses the strengths of both models, overcoming their well-known drawbacks to produce a vector space with a meaningful substructure. It achieves a 75% accuracy on the word analogy dataset (Pennington et al., 2014) and outperforms models on similarity tasks and named entity recognition benchmarks. The model leverages statistical information by training only on the nonzero elements in a word-word co-occurrence matrix, rather than on the entire sparse matrix or individual context windows in a large corpus. GloVe is based on a log-bilinear model with a weighted least-squares approach aimed at modeling the logarithm of the words' co-occurrence probability. The ratios of co-occurrence probabilities can be used to determine semantic relationships through the linear substructures of the word vector space. These linear substructures are captured by nearest neighbors, though Euclidean distance (or cosine similarity) between two word vectors that provides an effective method for measuring the linguistic

or semantic similarity of the corresponding words. Sometimes, the nearest neighbors according to this metric reveal rare but relevant words that lie outside an average human's vocabulary. Similarity metrics typically produce a single scalar value to quantify the relationship between two words. However, this oversimplifies the complex relationships that can exist between words. For instance, the words "man" and "woman" are both related as human beings, but they also represent contrasting concepts. To capture these nuances, a model should associate multiple values to word pairs. A natural approach is to use the vector difference between the two word vectors. GloVe is designed to ensure that these vector differences effectively capture the semantic and syntactic relationships between terms. The underlying concept distinguishing "man" from "woman" can be expressed by other word pairs such as "King" and "Queen." As shown previously, mathematically, the vector differences between these pairs may be:

e.g., *vector("Man") - vector ("Woman") or vector ("King") - vector ("Queen").*

Aspects of meaning can be extracted directly from co-occurrence probabilities.
Consider the matrix of word-word co-occurrence counts be denoted by $\mathbf{A}$, whose entries $\mathbf{A}_{ij}$ tabulate the number of times word j occurs in the context of word i. Let $A_i = \sum_k A_{ik}$ be the number of times any word appears in the context of word i. The probability that word j appears in the context of word i equals:

$$P_{ij} = P(j|i) = A_{ij}/A_i \tag{2}$$

Suppose we are interested to a specific concept, for which we might the words i and j.
The relationship between words can be examined by studying the ratio of their co-occurrence probabilities with various probe words, denoted as k (Pennington et al., 2014). The ratio of their co-occurrence probabilities is given by:

$$P_{ik}/P_{jk} \tag{3}$$

For words k related to i but not j, the ratio will return high values. Analogously, for words k related to j but not i, the ratio should be small. For words k that are either related to both i and j, or to neither, the ratio should be close to 1. In the algorithm steps, soft constraints for each word pair were defined using:

$$w_i^T w_j + b_i + b_j = \log(A_{ij}) \tag{4}$$

where w_i is the vector for the main word, w_j represents the vector for the context word, and b_i, b_j are scalar biases for the main and context words.
And a cost function is determined as:

$$J = \sum_{i=1}^{V} \sum_{j=1}^{V} f(A_{ij})(w_i^T w_j + b_i + b_j - \log A_{ij})^2 \tag{5}$$

where f denotes a weighting function which help us to prevent learning only from extremely common word pairs.
More properly, GloVe is based on the following function:

$$f(A_{ij}) \quad = \quad \begin{cases} \left(\frac{A_{ij}}{A_{max}}\right)^{\alpha} & \text{if } A_{ij} < A_{max} \\ 1 & \text{otherwise} \end{cases} \tag{6}$$

To summarize, GloVe's training objective is to learn word vectors so that their dot product approximates the logarithm of word co-occurrence probabilities both for similarity and analogy. Since the logarithm of a ratio equals the difference between logarithms, this approach connects co-occurrence ratios with vector differences, as explained in detail above. As these ratios capture semantic relationships, this information is encoded in the vector space. Consequently, GloVe's word vectors excel at word analogy tasks, similar to those in *word2vec* (Mikolov et al., 2013).
This study focuses on analyzing linguistic regularities through word analogies. By factorizing the co-occurrence matrix using the GloVe algorithm and then employing the local context window method of the skip-gram model, cosine similarity is used to identify these analogical relationships.

The structure of GloVe vectors often exhibits analogy-like relationships, such as:

$$\textit{vector("King") - vector("Man") + vector("Woman")} \approx \textit{?}$$

The application of the GloVe algorithm on the co-occurrence matrix $\mathbf{A}$ was replicated using the same set of parameters for both the male sub-corpus ($\mathbf{A_M}$), and the female sub-corpus ($\mathbf{A_F}$) in order to ensure methodological consistency and comparability between the two models.
This last step replies to RQ_2.
Figure 1 provides an overview of the methodological workflow, as a summary of the sequential steps and analytical procedures detailed in this Section 2.

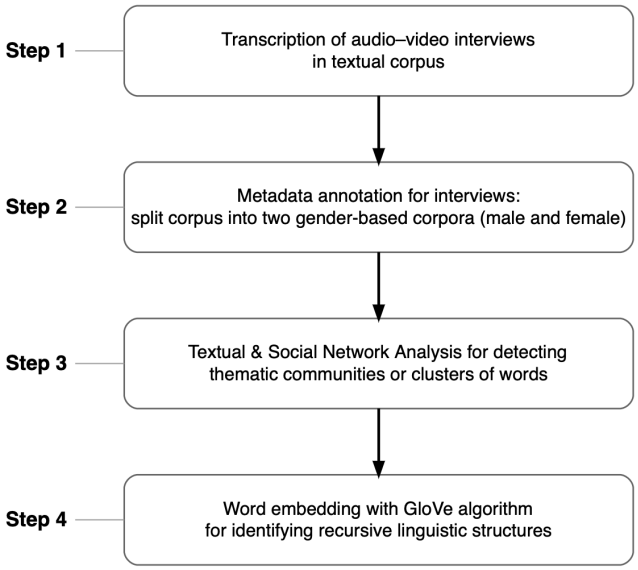


Figure 1: Methodological flowchart

To evaluate the model's ability to capture semantic relationships between words, cosine similarity between word vectors was measured in the GloVe embeddings, where **a** and **b** denote two word vectors in the embedding space. As is well known, cosine similarity

$$\cos(\theta) = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\|\|\mathbf{b}\|}$$

quantifies the alignment between two vectors in their multidimensional space: values close to 1 indicate high semantic similarity, values near 0 indicate no correlation, and negative values indicate dissimilarity. In this study, the model was trained using a minimum word occurrence threshold of 4, and 100-dimensional vectors. The learning rate was set to 0.05, with 25 training epochs, following parameter ranges commonly adopted in similar text-mining studies to ensure semantic stability and interpretability of the resulting embeddings (Pennington et al., 2014).
The co-occurrence matrices used for GloVe training were built from the preprocessing described in Step 3, maintaining consistency with the lexical data used for clustering procedure used in network analysis.

4. DATA

4.1. DATA COLLECTION

The data are derived from a non-probability sample of 101 video-audio interviews conducted in October 2022. The interviewees were women and men involved in local water management in the metropolitan city of Milan and adjacent geographical areas. It should be noted that the use of a non-probability sample limits the generalizability of the findings. This sampling strategy was chosen to include key stakeholders directly engaged in decision-making processes related to water governance within the study area and was also determined by the availability of the individuals involved in the research. Managers and operators working on projects and initiatives for ecological regeneration were included in this investigation, taking into account their roles and gender to achieve a balanced sample. Information concerning their roles and gender was added as metadata for each statistical unit.
Although the analysis adopts a binary categorization of gender-namely women and men-in line with the composition of the sample and the structure of the data, the study acknowledges that gender is a complex and socially constructed dimension that extends beyond the binary framework. This acknowledgment aligns the theoretical framework of the study with recent developments in gender research, which emphasize the existence of non-binary and intersectional identities influencing experiences, perceptions, and forms of participation in sustainability governance. The methodological alignment with a binary structure derives from a matter of practical convenience, as no non-binary gender identities were explic-

itly declared by the individuals included in the sample. Such recognition does not alter the methodological structure of the present work but aims to align its theoretical premises with an inclusive approach. Furthermore, the underrepresentation of women in leadership and executive positions within the sample reflects the actual gender imbalance observed in the local governance structure, rather than a methodological shortcoming of the research design. This empirical condition, while limiting the diversity of managerial perspectives, is itself a relevant finding that illustrates existing structural asymmetries in decision-making roles. Acknowledging this imbalance is therefore essential to understanding how gendered dynamics influence both participation and strategic orientation in sustainability practices.

The interviews were structured around four macro-categories: (a) *innovation*, (b) *ecological transition*, (c) *people*, and (d) *environment*. *Innovation* denotes processes or technologies designed to meet human needs while reducing environmental impact and fostering social and economic equity and well-being. This concept often involves creating sustainable business models, such as circular economies or waste-reducing systems, as well as developing renewable energy solutions. The *ecological transition* focuses on minimizing the environmental footprint of human activities through sustainable production and consumption practices, the use of renewable energy sources, cleaner transportation methods, agricultural techniques that protect soil and water quality, and the reduction of chemical fertilizers and pesticides. It also includes technological advancements that support sustainable development and the implementation of supportive policies. *People* are pivotal to sustainable development, encompassing key factors such as education, healthcare, access to clean water and sanitation, and opportunities for economic and social justice. The *environment*, on the other hand, consists of chemical, physical, and biological factors that affect human health and long-term ecological needs. The four pillars-innovation, ecological transition, people, and environment-are therefore treated as crucial semantic dimensions of water-related sustainable development, following the framework inspired by the sustainability model adopted by CAP Group. Founded in 1928, CAP Group is one of the main companies managing integrated water services in Milan and surrounding municipalities. The company is recognized as the first monoutility in the country's heritage and is actively involved in minimizing hydraulic risk and addressing urban flooding issues. The relevance and coherence of these four dimensions were subsequently validated through a pilot study in which the research team participated in thematic workshops focused on sustainability assessment for water management and urban

ecological interventions. Insights emerging from this preliminary stage helped refine the interpretative framework and confirm the suitability of the selected categories for the main phase of thematic area extraction. Each interview question was therefore coded according to these dimensions, which were operationalized as semantic domains reflecting sustainability and governance priorities relevant to the Milan metropolitan context. The coding framework was refined through iterative readings of the transcripts and systematic comparison across the research team to ensure conceptual coherence and reliability. Furthermore, interactions between gender and professional role were considered within each macro-category to capture how gendered perspectives intersect with functional responsibilities in shaping attitudes toward innovation, ecological transition, social inclusion, and environmental protection. These cross-dimensional relationships provided additional depth to the interpretation of governance dynamics and informed the subsequent analytical steps.

4.2. DATA DESCRIPTION

The sample consists of 57.43% men and 42.57% women. Meta-data analysis highlights a prevalence of men in apical roles or decision-making positions. Men hold 90.48% of CEO positions, 85.71% of General Director, and 69.23% of founders. Women are more prevalent in staff positions within geological support systems or in roles related to stakeholder engagement in environmental protection. These positions represent 93.21% of women's roles. Roles as coordinators (78.57%) and chairs of public or private organizations (53.85%) are also frequent.

The areas in which the various professionals are most engaged in the different roles show that innovation is the least frequent domain among the various professions, with values below 1%. The ecological transition is primarily addressed by CEOs (63.08%) and founders (54.45%). The topic of people or citizen engagement is represented by 66.56% of coordinators and 47.98% of geology staff. Environmental issues are strongly linked to the roles of organization chairs (72.61%). The descriptive analysis highlights that while men are more engaged in ecological transition practices, women in governance are more attentive to citizen participation and environmental changes.

5. RESULTS

The interview corpus is composed of a total of 1,957 distinct lexical items (types) and a total of 3,866 word occurrences (tokens). The lexical diversity, assessed using the *type-token ratio* (TTR), a widely employed but rather rudimentary met-

ric, is 50.62%, demonstrating a substantial range of vocabulary used to depict water-related well-being from both male and female viewpoints. Gender-based TTR analysis revealed a +1.9% difference favoring the male sub-corpus, suggesting greater heterogeneity in male managerial discourse. It should be noted, however, that the TTR is among the simplest and most coarse-grained measures for investigating corpus characteristics and, in this case study, it was employed merely as an exploratory proxy. To deepen the understanding of the textual features, metrics quantifying unique word occurrences within the corpus, termed *hapax legomena*, were employed. These metrics, commonly utilized to analyze authorial vocabulary, historical language trends, or subcultural lexicons, facilitated the identification of a female perspective on sustainable development. This perspective is predominantly defined by semantic components related to collaborative ideation, strategic planning, and joint participation. Statistical measures of gender-segmented lexical indices present a female subculture characterized by a consistent linguistic expression of shared theoretical frameworks, with the female sub-corpus showing a mean hapax of 0.09. Conversely, the male subculture demonstrated a higher tendency to engage with common topics through varied linguistic constructs, as evidenced by an average *hapax* value of 0.45.

The sparsity of the entire corpus is equals 72.5% on a lexical matrix of 915 features,($\mathbf{DTM}_{T101x915}$). The sparsity decreases within gender-specific subcorpora, where internal discursive homogeneity increases. Specifically, in the matrix of the male sub-corpus ($\mathbf{DTM}_{M58x329}$), sparsity is reduced to 66.30%, while in the matrix of the female sub-corpus ($\mathbf{DTM}_{F43x586}$), it drops further to 57.19%.

5.1. THEMATIC CLUSTERS ON WATER-BASED MANAGEMENT

The *fast-greedy* algorithm was applied on the co-occurrence matrices of the two sub-corpora: $\mathbf{A}_{M329x329}$ for male matrix, and $\mathbf{A}_{F586x586}$ for female matrix.
Several communities of frequently co-occurring terms were detected with a different number of communities or thematic clusters, depending on whether it is the network of male stakeholders or female stakeholders.
The performance of algorithm on matrix $\mathbf{A}_M$ showed a good modularity optimization with Q equals 0.40. In order to identify the two main thematic clusters characterizing the male stakeholders' sub-corpus, Table 1 presents the two communities detected.

As a complement to the semantic overview presented in Table 1, Table 2 reports the list of the 50 most frequent terms associated with each community. This detailed lexical mapping provides transparency for the interpretation of the

Table 1: Semantic communities identified in the male sub-corpus

Community	Description
Community 1: Technical Innovation and Operational Efficiency	Centred on projects, processes, and materials related to water purification and circular economy. This cluster reflects a pragmatic and engineering-oriented discourse emphasizing innovation, ecological transition, and the optimization of technical resources within municipal operations.
Community 2: Institutional Participation and Territorial Governance	Focused on citizens, associations, and local regulation frameworks. This community combines managerial language with participatory themes, highlighting collaboration between public and private actors, stakeholder involvement, and social responsibility in the governance of sustainability.

clusters and allows readers to trace how the thematic labels were derived.

In Table 2, first thematic group presents a broad perspective on sustainable development, emphasizing issues such as plastic pollution, reliance on disposable resources, the necessity for system purification, and circular economy approaches. The second group focuses on the role of local authorities, including municipalities and other stakeholders, highlighting sustainable well-being as a process that transfers responsibility to local entities. Specifically, the first community highlights themes related to public administration and local development. Terms such as municipality, event, platform, process, project, activity, and improve reflect an interest in enhancing public management and administrative efficiency. Additionally, words like startup, ecological, economy, plastic, single-use, and sand illustrate the intersection between innovation and sustainability. The findings indicate that sustainable local governance requires a balance between technological advancements and environmentally conscious practices. The second community revolves around concepts of civic engagement and social dynamics. Terms such as key, citizen, territory, relationship, association, participation, social, involvement, pact, value, private, stakeholder, sustainability, proximity, and responsible suggest a strong emphasis on community engagement and social responsibility. The male sub-corpus is focused on regulatory frameworks and institutional roles, recognizing the value of community involvement, they prioritize different mech-

Table 2: Communities of male sub-corpus

Community	Words
Community 1: Technical Innovation and Operational Efficiency	comune (municipality/town hall), evento (event), piattaforma (platform), processo (process), progetto (project), specifico (specific), attività (activity), migliorare (improve), materiale (material), startup (startup), depurazione (purification), transizione (transition), ecologico (ecological), problema (problem), economia (economy), circolare (circular), plastica (plastic), design, monouso (single use), sabbia (sand), kg
Community 2: Institutional Participation and Territorial Governanc	chiave (key), cittadino (citizen), territorio (territory), relazione (relationship), outing (outing), spazio (space), percorso (path/route), quartiere (neighborhood/district), partecipazione (participation), associazione (association), condivisione (sharing), regolamento (regulation), sociale (social), profit (profit), attivo (active), coinvolgimento (involvement), invitare (invite), patto (pact/agreement), valore (value), privato (private), soggetto (subject), direzione (direction), test (test), stakeholder (stakeholder), sostenibilità (sustainability), consiliatura (council term), vicinanza (proximity/closeness), responsabile (responsible)

anisms for achieving it.
Figure 2 shows a visual representation of the word co-occurrence networks as an exploratory visualization technique that supports the interpretation of results by making communities, hubs, nodes, and bridges visible. Nodes are organized into two thematic clusters: (i) NBS thematic group 1, capturing project-operational and material (techno-environmental) aspects; and (ii) NBS thematic group 2, aggregating participatory governance and the regulatory framework (place-based). The spatial arrangement highlights intra-cluster densifications. In the purple area (group 1), two sub-poles emerge: (a) materials and production processes (plastic (plastica), single use (monouso), kg, material (materiale), design (design), process (processo), project (progetto), platform (piattaforma)); (b) transition and environment (circular (circolare), transition (transizione), treatment (depurazione), ecological (ecologico), activity (attività)). In the orange cluster (group 2), we observe: (a) participation and regulation (participation (partecipazione), sharing (condivisione), engagement (coinvolgimento), stakeholder (stakeholder), regulation (regolamento), responsible (responsabile), social (sociale), value (valore)); (b) territory and proximity (neighborhood (quartiere), territory (territorio), citizen (cittadino), actor (soggetto), space (spazio), proximity (vicinanza), council (consiliatura), pact (patto), invite (invitare)). This indicates that institutional and social dimensions structure actors, procedures, and territorial proximity (consistent with a managerial framing). Bridging nodes occupy intermediate positions linking the two groupings: on the governance side (territory (territorio), neighborhood (quartiere), sharing (condivisione), key (chiave), path (percorso), value (valore)) and on the technical-operational side (economy (economia), activity (attività), specific (specifico), design (design), treatment (depurazione)). These "corridors" connect governance and territorial aspects with techno–environmental ones. Peripheral lexical units pertain to technical and procedural items (*e.g.*, problem (problema), process (processo), plastic (plastica), single use (monouso), regulation (regolamento)), whereas a central nucleus (territory (territorio), neighborhood (quartiere), sharing (condivisione) with economy (economia), treatment (depurazione), activity (attività)) constitutes a cross-cutting overlap area with a distinctly operational and territorial profile, consistent with socio-technical co-production.

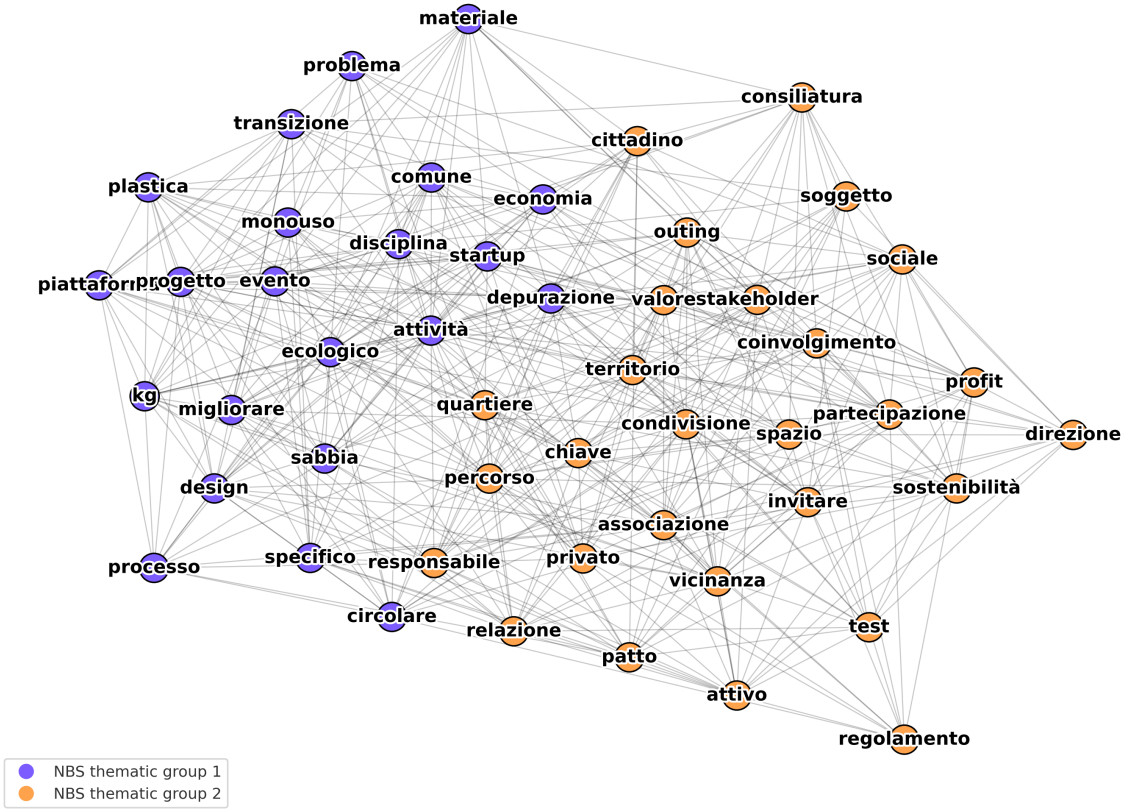


Figure 2: Network of male co-occurrences

The performance of matrix $\mathbf{A}_F$ shows a Q value of 0.34, indicating a value falling in optimal range. The algorithm for the female network returned three communities as reported in Table 3.

As in the previous case, the 50 most frequent terms for female sub-corpus are displayed into Table 4.

In Table 4, the first thematic cluster is characterized by a generalist view of the hydric resources using terms such as pollution, plastics, and sewage systems. The second group highlights the need to build a path for services and innovation. The third group identifies the welfare, recalling themes as environment, climate change, urban protection, care for people and planet. Female view is mainly related to a process of simplification, transformation, and digitalization of systems aimed at innovation and recycling-based economies. Citizen and cooperation are

Table 3: Semantic communities identified in the female sub-corpus

Community	Description
Community 1: Circular Innovation	Focused on experimentation, circular economy, and ecological design. This cluster reflects a managerial and creative discourse emphasizing sustainable production models and technological transition.
Community 2: Civic Participation and Local Engagement	Centred on citizens, associations, and territorial services. The discourse highlights collaboration, public involvement, and active participation in local sustainability initiatives.
Community 3: Environmental Protection and Scientific Awareness	Oriented toward environmental safeguarding, climate change, and natural resources. This cluster integrates scientific terminology and advocacy for ecological responsibility.

crucial in the sewage treatment as well as the geological aspect. Specifically, the first community focuses on concepts related to sustainable development and technological innovation. Terms such as sector, experiment, economy, circular, transition highlight the importance of sustainability and ecological solutions. Words like plastic, design, startup, ecological, purification indicate a specific concern for recycling and responsible resource use. The second community is characterized by a focus on civic participation and community dynamics. Terms such as citizen, project, association, participation, involvement, private reflect the emphasis on engaging citizens and local associations in territorial and social management. The third community is oriented towards environmental protection and ecological awareness. Words such as protection, solidarity, cultural, geology, natural, environment, change, climatic, water, water-related, planet emphasize the commitment to sustainability and territorial protection. Female perspective shows that technological advancements significantly support the transition to a circular economy. However, challenges such as high initial costs and resistance to change still limit widespread adoption. Startups that implement sustainable practices gain consumer trust and regulatory benefits, reinforcing the idea that sustainability can drive economic growth. Civic engagement emerged as a key factor in the success of environmental initiatives. Projects with strong citizen participation tend to have a more lasting impact, especially when local associations play an active role. Bu-

Table 4: Communities of female sub-corpus

Community	**Words**
Community 1: Circular Innovation	settore (sector), sperimentare (experiment), economia (economy), circolare (circular), transizione (transition), problema (problem), plastica (plastic), design (design), startup (startup), ecologico (ecological), depurazione (purification)
Community 2: Civic Participation and Local Engagement	cittadino (citizen), progetto (project), chiave (key), associazione (association), elemento (element), piano (plan), servizio (service), attivo (active), territorio (territory), relazione (relationship), partecipazione (participation), quartiere (neighborhood), percorso (path), spazio (space), coinvolgimento (involvement), direzione (direction), privato (private)
Community 3: Environmental Protection and Scientific Awareness	tutela (protection), solidarità (solidarity), innovatore (innovator), culturale (cultural), geologia (geology), naturale (natural), ambiente (environment), cambiamento (change), climatico (climatic), acqua (water), idrico (water-related), riso (rice), protezione (safeguard), contenimento (containment), contesto (context), alleanza (alliance), scientifico (scientific), sottosuolo (subsoil), fondo (bottom), pianeta (planet)

reaucratic obstacles and funding limitations remain significant challenges. Public-private partnerships offer a potential solution, combining community involvement with institutional support to create more effective sustainability programs. Environmental protection is closely linked to cultural awareness. Societies with a strong ecological identity tend to adopt more sustainable behaviors. Water management plays a crucial role in climate adaptation, with innovative purification technologies proving essential. The study also highlights that grassroots initiatives often outperform top-down policies in environmental conservation. A key policy recommendation is to integrate local knowledge into governance strategies, ensuring that sustainability efforts are both effective and culturally relevant. The findings emphasize the need for innovation, active participation, and localized approaches to achieve a more sustainable and inclusive environmental transition. The integration of circular economy principles, particularly in waste management and resource optimization, appears to be more emphasized in the female sub-corpus, whereas the male sub-corpus focuses more on administrative efficiency and service management. The second community revolves around concepts of civic engagement and social dynamics. Terms such as key, citizen, territory, relationship, association, participation, social, involvement, pact, value, private, stakeholder, sustainability, proximity, and responsible suggest a strong emphasis on community engagement and social responsibility. Both sub-corpora acknowledge the importance of participatory governance, but with different nuances. The female sub-corpus tends to emphasize collaboration, mutual support, and inclusive decision-making, while the male sub-corpus is more focused on regulatory frameworks and institutional roles. This suggests that while both perspectives recognize the value of community involvement, they prioritize different mechanisms for achieving it.

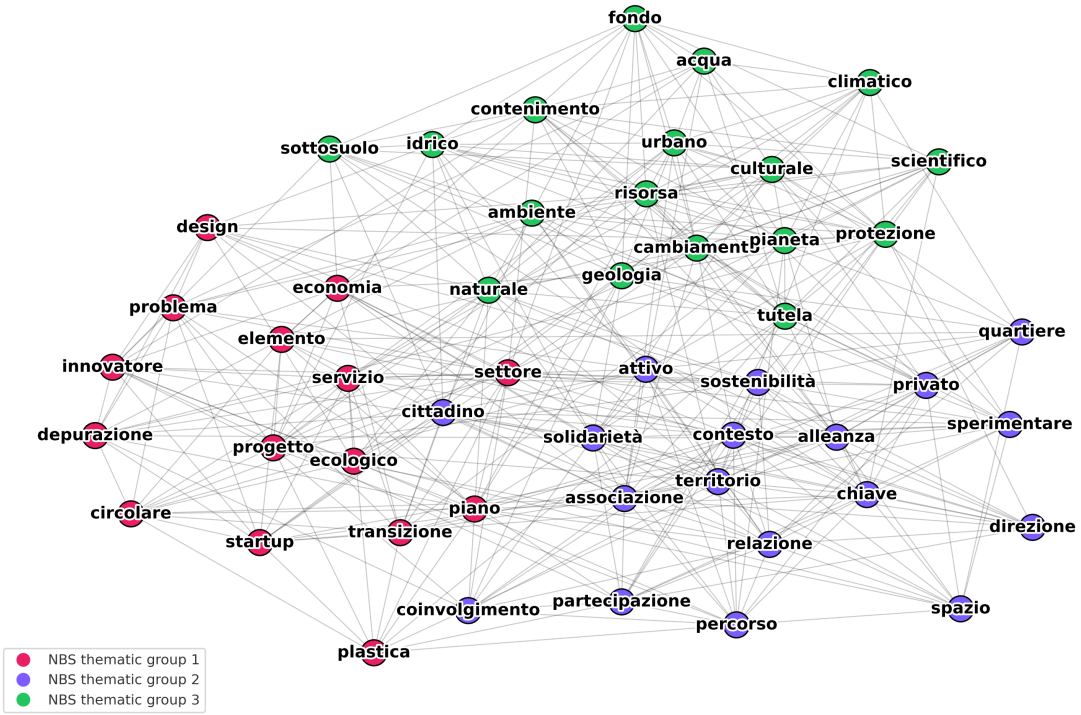


Figure 3: Network of female co-occurrences

Fig. 3 offers a synoptic reading of the female managerial vocabulary through a network composed of three partially overlapping thematic nuclei, reflecting three semantic domains: a technical-operational domain (NBS thematic group 1), a governance and participation domain (NBS thematic group 2), and an environmental and scientific domain (NBS thematic group 3). The spatial distribution indicates greater affinity and conceptual co-occurrence: the technical-operational nucleus (group 1) gathers terms related to processes, tools, and project deliverables. This area includes *project*, *design*, *transition*, *circular*, *treatment*, *ecological*, *service*, and *plan*, with extensions toward material elements such as *plastic* and performance concerns such as *economy* and *problem*. The proximity among *project*, *design*, and *transition*/*circular* suggests a lexicon oriented to operational transformation, where innovation (*startup*, *innovator*) is connected both to technological components and to efficiency criteria. The governance and participation nucleus (group 2) concentrates actors and social practices-*participation*, *engagement*, *association*, *citizen*-together with terms that provide direction and framing

(*territory*, *neighborhood*, *relation*, *key*, *sustainability*, *context*). The relatively compact position of *participation* with *engagement* and *association* signals a semantics of mobilization, while the links to *territory* and *neighborhood* indicate the local grounding of practices. The environmental-scientific nucleus (group 3) articulates the physical, ecological, and knowledge-cultural components: *water*, the *hydric* element, *resource*, *protection*, and *safeguard* converge with *geology* and macro references to *climate change*. This cluster maintains a robust internal backbone, while the "geo-climate" branch points to foundational knowledge and scales of reference. The three thematic clusters translate into conceptual "corridors" where the technical area dialogues with the environmental one, highlighting the translation of ecological constraints into project choices. The social dimension is foregrounded, suggesting that the integration of social and ecological instances constitutes one of the densest areas of discourse. The overall topology shows a densely interconnected center, where cross-cutting terms (*sector*, *active*, *territory*, *environment*) coexist, and a sparser periphery hosting more specialized or focused elements (*innovator*, *scientific*, *containment*). This center-periphery gradient makes visible how the NBS agenda is structured around a core of shared concepts, from which more specific sub-languages branch out.

A comparative analysis of the two tables reveals significant differences in thematic focus between the female and male sub-corpora. The gender-based analysis highlights different perspectives on how governance interprets the water issue and intends to plan future strategic resolutions. In the thematic group 1, both perspectives recognize the importance of innovation and environmental sustainability. However, the female sub-corpus places a stronger emphasis on circular economy principles and resource management, whereas the male sub-corpus prioritizes local administration and service efficiency. In thematic group 2, both sub-corpora express a commitment to civic participation and community involvement, but female perspective highlights cooperative and supportive initiatives, while the male sub-corpus is more concerned with regulations and institutional dynamics, focusing on roles and responsibilities within governance structures. Notably, the thematic group 3 appears only in the female sub-corpus and explicitly addresses environmental protection and ecological awareness. This suggests a greater concern for sustainability and environmental advocacy among female respondents, an aspect that is absent in the male sub-corpus. Overall, these findings highlight complementary approaches to sustainable development and territorial governance. The female sub-corpus demonstrates a stronger orientation toward ecological awareness and social well-being, whereas the male sub-corpus places

greater emphasis on administrative management and institutional efficiency. Understanding these differences can help develop more inclusive and effective sustainability strategies that integrate both environmental and governance perspectives. Based on the findings, women in local governance demonstrate a distinct perspective that prioritizes transformation where welfare that stems from a harmonious blend of social and culture dimensions. These insights underscore the importance of women's empowerment in local governance and their unique contributions in shaping sustainable and equitable communities.
From a graphical standpoint, the architectures of the two networks differ markedly. The male semantic network consists of two dense blocks with many cross-links, yielding a predominantly techno-operational reading; by contrast, the female semantic network exhibits three main blocks that support a techno-social and environmental and scientific interpretation. In the male managerial view, centrality is driven by the interface between territorial sociality and techno-processual activity (e.g., economy, activity, specific, treatment). The female view is more polar, with cross-cutting nodes that connect territory and environment with economy and design, as well as context governance and sustainability. The thematic peripheries also diverge: in the male semantics they relate to material and regulatory aspects, whereas in the female semantics they highlight specialist issues (e.g., containment, science) and problematizations linked to organizational elements (e.g., alliance, solidarity).

5.2. RECURSIVE STRUCTURES IN MANAGEMENT LANGUAGE

To detect the women's vision and determine gender-based differences word embedding was performed on sub-corpora.
The same word vector differences were adopted to capture similarities or dissimilarities between genders.
From the vocabularies of both male and female sub-corpora, eleven shared lemmas with a standardized centrality score above 0.50 (see Table 5) were identified. To evaluate the model's ability to capture semantic relationships between words, cosine similarity between word vectors was measured. For analogy analyses (Table 5), the resulting vector was computed, and words in the vocabulary were ranked and filtered based on their cosine similarity to this vector. Using word embedding techniques, the NBS emerged as a conceptual framework encompassing urban environments, industrial operations, governance, and civic participation. To better understand how these elements shape discourse, the selected terms relate to locality (local, urban), advantages (benefit, proximity), governance (politi-

cal, citizen), industry (industrial), and spatial aspects (quarter, service, territory). The analysis of word vector representations highlights gender-based differences in the perception of sustainability and urban spaces. Male stakeholders tend to emphasize regulatory aspects, policy frameworks, and sector-specific issues such as biomethane as a renewable energy source and industrial ecology. In contrast, female stakeholders focus more on activities related to resource recovery, business engagement, and urban space preservation, often with a forward-looking perspective. This distinction is evident in their approach: men concentrate on technical solutions like waste-to-energy processes, whereas women are more concerned with territorial challenges, such as environmental issues in Po Valley watershed. Despite these differences, both groups converge on the significance of shared urban spaces, resource management, and active citizen participation in governance.

Table 5 reports the similarity between each word and a target vector, ranked according to the highest cosine similarity of linguistic substructures extracted. Therefore, the words were ranked from most similar to least similar, providing an ordering based on the semantic relationships learned by the vectors: words with high empirical probability or high similarity (cosine value ≥ 0.60) were reported.

To evaluate the model's ability to capture the alignment between cosine similarity and semantic similarity among words, a gold standard was constructed based on the topology of the semantic network (Section 3). More precisely, the labels of the thematic clusters identified in the network analysis group conceptually coherent words: words with a link were manually labeled as positive (true label = 1), while those without a link to other words were labeled as negative (false label = 0). The correct word of analogy analyses is considered a positive prediction based on classification metrics such as *accuracy*, *precision*, *recall*, and *F1-score* (Table 6).

As reported in Table 6, the classification metrics for the male and female subcorpora highlight differences in the model's predictive performance. For the male sub-corpus, accuracy, precision, recall, and F1-score reach relatively high and balanced values (accuracy = 0.733, precision = 0.818, recall = 0.818, F1-score = 0.818), indicating that the model is able to correctly identify positive and negative cases with good overall reliability. In contrast, the female sub-corpus shows lower precision (0.417) and F1-score (0.588), although recall remains high (1.0), suggesting that while the model successfully identifies most true positives, it also produces a higher number of false positives, leading to reduced precision. These results reveal a gender-related discrepancy in model performance. The male subcorpus exhibits consistently higher metrics across accuracy, precision, recall, and

Table 5: Word vector representations and linguistic sub-structure(s) in gender-based sub-corpora

Word vectors	Male sub.	Sim.	Female sub.	Sim.
Sustainability − local + urban	Discipline	0.78	Recovery	0.61
Sustainability − local + benefit	Reciprocal	0.68	Business	0.65
Service − proximity + political	Biomethane	0.56	To undertake	0.75
Territory − proximity + political	Biomethane	0.70	Cartographic	0.70
Territory − local + citizen	Europe	0.59	Resource	0.72
	Symbiosis	0.58	Safeguard	0.53
	Industrial	0.58		
Political − local + industrial	Waste to energy	0.60	Neighboring	0.64
	Purifier	0.60	Po Valley	0.59
	Participatory	0.59		
Sustainability − industrial + neighborhood	Sharing	0.68	Association	0.58
			To involve	0.55
Territory − local + benefit	Plastic	0.67	Plant	0.67
	Mud	0.61	Winter	0.63
	Transfer	0.61	Project	0.61
	Ecological	0.60	Industrial	0.58
	Circular Economy	0.53	Ecological	0.52

[1] "sub" is referred to linguistic sub-structures.

[2] "Sim." indicates the cosine similarity.

F1-score, reflecting a better alignment between the predicted labels and the gold standard. In the female sub-corpus, the reduced precision and F1-score indicate that the semantic embeddings capture the target relations less effectively, possibly due to differences in word distributions or conceptual associations within the female sub-corpus. Overall, Table 6 illustrates that the model performs more robustly for the male sub-corpus compared to the female sub-corpus, highlighting the importance of considering gender-specific patterns in semantic representation.

6. DISCUSSION AND CONCLUSION

This paper investigated women's active participation in local governance structures related to sustainable development. This paper investigated women's active

Table 6: Metrics for male and female sub-corpora

Sub-corpus	Accuracy	Precision	Recall	F1-score
Male	0.733	0.818	0.818	0.818
Female	0.500	0.417	1.000	0.588

participation in local governance structures related to sustainable development. By studying the integrated management of water resources in Milan, gender-based managerial visions for the development of NBS were proposed, demonstrating how the gender of stakeholders can influence the determinants guiding urban transformation processes based on sustainable principles. To identify gender-based perspectives, a combination of NLP approaches was proposed to develop an innovative strategy aimed at analyzing stakeholders' speech. A collection of interviews with local operators and managers active in Milan and surrounding areas was conducted. Their interviews were transformed into a textual corpus, and each individual speech was annotated with metadata regarding the speaker's gender and professional role. Initially, the data collection highlighted a disparity in top-level positions related to sustainable development management in Milan. Men predominantly occupy higher-level roles such as CEO, General Director, and Founder, while women are more commonly found in positions related to geological support systems, stakeholder engagement, and environmental protection. The gender-based metadata enabled the creation of two sub-corpora on which a linguistic analysis was conducted to capture the managerial visions of men and women, and how objectives, strategies, and actions related to water resource management differ. The first analytical step focused on defining differing conceptualizations connected to the topics. The textual network approach proved effective in revealing distinct thematic groups related to sustainable development practices between male and female governance. Female governance views sustainable engagement as a progressive transformation, emphasizing citizen involvement and highlighting welfare, solidarity, science, and culture as fundamental and interconnected pillars. Conversely, male governance perceives sustainable development as a linear process, progressing from the circular economy to the intermediation of territorial authorities, acting as a bridge. The second analytical step involved the use of a word embedding strategy, which uncovered interesting linguistic combinations based on the use of common words. The results underscored the deter-

mining semantics that describe the strategic visions of governance. The semantic depth provided by word embedding further enriched the thematic framework and revealed the presence of recursive structures characterizing managerial visions in different domains, such as urban territory, industrial activity, politics, and citizen participation. Significant differences in the linguistic and semantic structures of the male and female sub-corpora emerged. The female sub-corpus associates "sustainability" with concepts like "recovery" and "business," emphasizing well-being, local development, and community involvement. In contrast, the male sub-corpus links "sustainability" to "discipline" and "biomethane," reflecting a more technical and industrial perspective. These semantic differences highlight broader cultural and social variations, with women focusing on social and environmental aspects, while men prioritize technology and industry. The findings also brought to light a future-oriented female participation that prioritizes specific issue identification over immediate solution finding. Women demonstrated awareness of transformational processes involving diverse contexts and actors and view solidarity, science, culture, and welfare as crucial factors for achieving full sustainable development. Water resources therefore have implications across multiple sustainability dimensions. Male visions focus on support from local authorities and the circular economy, emphasizing current strategies, technical competence, specialization, European policies, industrial ecology, and renewable resources. The findings suggest that gendered factors influence how concepts like "sustainability," "territory," and "politics" are linguistically represented and prioritized.

The results also revealed differences in the way gendered perspectives interacted within governance processes. Exchanges among male participants tended to reinforce technical and performance-oriented framings, reflecting a shared managerial rationality. In mixed-gender dialogues, however, discussions often became more integrative, combining procedural and social dimensions, particularly when women held coordination or facilitation roles. Interactions among female participants were generally characterized by collaborative and consensus-building language, emphasizing inclusivity and mutual learning. These patterns indicate that gendered communication does not operate in isolation but through relational dynamics that mirror the broader organizational culture and local governance arrangements.

The observed gendered linguistic patterns can be interpreted in light of existing theoretical perspectives on gender and governance (e.g.,(Piserà, 2024; van Daalen K.R et al., 2024)). Differences in discursive orientations may reflect structural factors, such as the underrepresentation of women in senior decision-making

roles, which can shape access to institutional power and influence linguistic framing. Consistent with this, the sample analyzed in the present study reveals a gender imbalance, with male participants disproportionately represented in executive and managerial positions within local governance bodies. At the same time, cultural norms and role-based expectations likely contribute to divergent rhetorical strategies: male participants tended to emphasize managerial and technical rationales, whereas female participants more frequently referred to social inclusion, collaboration, and relational dimensions of governance. These tendencies align with prior research suggesting that gendered communication styles are often embedded within broader organizational cultures and governance structures. The empirical evidence collected in this study reflects these dynamics. In fact, within the examined sample, men were largely concentrated in executive and managerial positions, especially within technical and infrastructural departments concerned with hydraulic risk management, maintenance, and regulatory compliance. Their narratives frequently revolved around efficiency, innovation, and the optimization of resources-concepts closely associated with the *innovation* and *ecological transition* dimensions identified in the analysis. By contrast, female participants, who were more often engaged in planning, communication, and citizen participation units, tended to articulate discourses emphasizing inclusion, awareness-raising, and the social value of sustainability initiatives. Their language recurrently integrated references to cooperation, care, and long-term ecological responsibility, thus aligning more strongly with the *people* and *environment* macro-categories. Taken together, these findings suggest that gendered linguistic differences may not simply derive from communicative styles, but rather from differentiated institutional roles and access to decision-making spheres. In this sense, discourse becomes a reflection of structural positioning within local governance, revealing how responsibilities, expertise, and interactional contexts influence the way sustainability and water management are conceptually framed. Therefore, the findings should be interpreted with caution. The study demonstrates how several contextual and individual-level variables may act as confounding factors influencing linguistic patterns, including participants' professional roles, institutional affiliation, and exposure to participatory decision-making environments. While the current study controlled for role distribution in sampling, future research could further disentangle the effects of structural and cultural determinants from those of individual agency by integrating mixed-method approaches and larger samples. Such developments would strengthen the explanatory potential of the gender-governance nexus highlighted in this analysis.

Some limitations must be acknowledged. The sample is relatively small (101 interviews), which affects the generalizability of the results. Future studies should expand the sample size and include additional stakeholders such as residents, political authorities, and other governance actors to capture a more comprehensive view. From a research perspective, further work could also explore how governance dynamics and gender inclusion differ across various spatial contexts - for instance, by comparing metropolitan areas such as Milan with smaller urban or peri-urban settings. Such comparative analyses would help identify whether scale and local governance structures influence participatory processes and the effectiveness of NBS.

Methodologically, the transcription process-despite a high accuracy rate-may have introduced minor errors that could influence the linguistic analyses. Although the study identified gender-based differences in discourse, causality cannot be inferred; further longitudinal or experimental research would be required to understand how gender shapes decision-making in governance over time. Furthermore, while the study's use of NLP approaches such as word embedding and textual network analysis provided meaningful insights, these methods rely on statistical co-occurrence patterns and may overlook subtle contextual nuances or the influence of other demographic variables, such as age, education, and professional background. Future research should employ more advanced neural network techniques to enhance semantic and linguistic analysis. Methods such as named entity recognition, sentiment analysis, and tone detection could offer additional depth in identifying demographic and gender-specific differences in governance discourse.

Finally, beyond the specific context of Milan, the proposed methodological framework can be adapted and scaled to other urban and regional settings characterized by complex governance systems. Its combined use of textual analysis, word embeddings, and social network analysis provides a flexible structure that can be replicated with different types of stakeholders and institutional configurations. Adapting the approach to other contexts would primarily require recalibrating the linguistic corpus and the metadata variables according to the local governance structures, languages, and thematic domains of sustainability. For instance, the same analytical pipeline could be applied to investigate gendered perspectives in water governance in other metropolitan areas or extended to related sectors such as climate adaptation, energy transition, or urban regeneration. Such comparative applications would make it possible to assess the robustness and transferability of the framework, while highlighting context-specific features that influence discursive and governance patterns. In this sense, the methodological design presented

here is intended as a transferable analytical tool rather than a case-specific model, capable of supporting cross-case investigations into the gender-governance nexus and the social dimensions of sustainability-oriented policies.

References

Bengio, Y., Ducharme, R., Vincent, P., and Janvin, C. (2003). A neural probabilistic language model. In *Journal of Machine Learning Research*, 3: 1137–1155.

Benzécri, J. (1981). *Pratique de l'analyse des données. Linguistique & Lexicologie*. Springer.

Blei, M., Ng, A., and Jordan, M. (2003). Latent dirichlet allocation. In *The Journal of Machine Learning Research*, 3: 993–1022.

Bocci, C. and Mishra, K. (2021). Forest power: The impact of community forest management on female empowerment. In *Ecological Economics*, 187.

Bolasco, S. (2013). *L'analisi automatica dei testi. Per fare ricerca con il text mining*. Carocci Editore.

Borgatti, S.P., Mehra, A., Brass, D.J., and Labianca, G. (2009). Network analysis in the social sciences. In *Handbook of Research Methods in Social and Behavioral Sciences*, 892–895.

Chattopadhyay, R. and Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in india. In *Econometrica*, 72 (5): 1409–1443.

Clauset, A., Newman, M.E.J., and Moore, C. (2004). Finding community structure in very large networks. In *Phys. Rev. E.*, 70: 066111.

Cornwall, A. and Rivas, A.M. (2015). From âgender equalityâ and âwomenâs empowermentâ to global justice: reclaiming a transformative agenda for gender and development. In *Third World Quarterly*, 36 (2): 396–415.

Deerwester, S., Dumais, S., Furnas, G., et al. (1990). From âgender equalityâ and âwomenâs empowermentâ to global justice: reclaiming a transformative agenda for gender and development. In *Journal of the American Society for Information Science*, 41: 391–407.

Dollar, D., Fisman, R., and Gatti, R. (2001). Are women really the âfairerâ sex? corruption and women in government. In *Journal of Economic Behavior & Organization*, 46 (4): 423–429.

Eggermont, H., Balian, E., Azevedo, J., et al. (2015). Nature-based solutions: new influence for environmental management and research in europe. In *GAIA - Ecol. Perspect. Sci. Soc.*, 24: 243–248.

Ershad Sarabi, S., Han, A., Romme, B., et al. (2019). Key enablers of and barriers to the uptake and implementation of nature-based solutions in urban settings: A review. In *Resources*, 8.

European Commission (2023). *Nature-based solutions â EU-funded nbs research projects tackle the climate and biodiversity crisis*. Publications Office of the European Union.

Fortunato, S. (2010). Community detection in graphs. In *Physics Reports*, 486: 75–174.

Gerlach, M., Peixoto, T., and Altmann, E. (2018). A network approach to topic models. In *Science Advanced*, 4: 1–12.

Hofmann, T. (1990). Probabilistic latent semantic analysis. In *Proceedings of the XV Conference on Uncertainty in Artificial Intelligence*. Stockholm, Sweden.

Lancichinetti, A. and Fortunato, S. (2011). Limits of modularity maximization in community detection. In *Phys. Rev. E.*, 84 (6): 066122.

Leone, M. (2019). Women as decision makers in community forest management: Evidence from nepal. In *Journal of Development Economics*, 180–191.

Mikolov, T., Yih, W., and Zweig, G. (2013). Linguistic regularities in continuous space word representations. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (HLT-NAACL 2013)*, 746–751. Association for Computational Linguistics, Atlanta, Georgia.

Nellemann, C., Verma, R., and Hislop, L. (2011). *Women at the frontline of climate change: Gender risks and hopes*. United Nations Environment Programme.

Newman, M. (2006). Modularity and community structure in networks. In *Proceedings of the National Academy of Sciences of the United States of America*, 103: 8577–8582.

Newman, M. and Girvan, M. (2004). Finding and evaluating community structure in networks. In *Phys. Rev. E.*, 69: 026113.

Paolillo, M. and Forciniti, A. (2021). L'impatto del covid-19 sull'opinione pubblica: una strategia di analisi per lo studio della comunicazione su twitter. In A. Favretto, A. Maturo, and S. Tomelleri, eds., *L'impatto sociale del Covid-19*, 310–318. Franco Angeli, Milano.

Pennington, J., Socher, R., and Manning, C. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543. Association for Computational Linguistics, Doha, Qatar.

Piserà, S. (2024). Gendered language and board diversity: Evidence from european banks. In *International Journal of Finance & Economics*, 29 (3): 3083–3102. doi:10.1002/ijfe.3039.

Regione Lombardia (2015). *Legge Regionale 15 marzo 2016, n. 4, Revisione della normativa regionale in materia di difesa del suolo, di prevenzione e mitigazione del rischio idrogeologico e di gestione dei corsi d'acqua*.

Regione Lombardia (2017). *Regolamento Regionale 23 novembre 2017, n. 7, Regolamento recante criteri e metodi per il rispetto del principio dell'invarianza idraulica ed idrologica ai sensi dell'articolo 58 bis della legge regionale 11 marzo 2005, n. 12 (Legge per il governo del territorio)*.

UN Women (2020). Women's leadership and political participation. `https://www.unwomen.org/en/what-we-do/leadership-and-political-participation`.

United Nations (2015). *Resolution adopted by the General Assembly on 25 September 2015. Transforming our world: the 2030 Agenda for Sustainable Development, A/RES/70/1*. United Nations.

van Daalen K.R et al. (2024). Bridging the gender, climate, and health gap: The road to cop29. In *The Lancet Planetary Health*, 8 (12): e1088–e1105. doi: 10.1016/S2542-5196(24)00233-9.

Wamsler, C., Alkan-Olsson, J., Björn, H., Falck, H., et al. (2020). Beyond participation: When citizen engagement leads to undesirable outcomes for nature-based solutions and climate change adaptation. In *Clim. Chang.*, 158: 235–254.

Wasserman, S. and Faust, K. (1994). *Social Network Analysis*. Cambridge University Press, Cambridge.

WCED, World Commission on Environment and Development (1987). *Report of the World Commission on Environment and Development: Our Common Future*. New York, USA.

World Bank (2019). *Women, Business and the Law 2019: A Decade of Reform*. World Bank Group.

Yadav, S. (2023). Women empowerment and representation in local governance. In *Indian Journal of Law and Legal Research*, 5.

Zhou, C., Feng, L., and Zhao, Q. (2017). A novel community detection method in bipartite networks. In *Physica A.*, 492: 1679–1693.

Zingraff-Hamed, A., Huesker, F., et al. (2020). Governance models for nature-based solutions: Seventeen cases from germany. In *Nature-based solutions in river landscapes*. Springer.