



FLORISTIC AND VEGETATION DATA ANALYSIS IN THE AGE OF ARTIFICIAL INTELLIGENCE

FEOLI ENRICO

Department of Life Sciences, University of Trieste, Italy
feoli@units.it

(RECEIVED 27 FEB 2026; RECEIVED IN REVISED FORM 11 MAR 2026; ACCEPTED 7 APR 2026)

ABSTRACT – The structuring of ecological data has long been central to biodiversity research. In Italy, Sandro Pignatti’s promotion of the first national database of Italian flora represented a landmark achievement in organizing ecological knowledge. Building on this legacy, the present paper explores how Artificial Intelligence (AI) can be employed to obtain and analyze data, thereby complementing traditional approaches to vegetation analysis. Rather than offering a technical account of AI, the paper demonstrates its practical value through two case studies conducted with Microsoft Copilot, a free AI tool available in recent Windows systems. The first case study classifies Italian regions based on the chorological types of their flora. The second examines oak woodlands in Central-Southern Italy, considering structural data such as NDVI, biomass, and percentage of leaf cover. Both examples align with Pignatti’s research lines in phytogeography and plant ecology, and are intended to be critically assessed by researchers with expertise in these fields. The aim is not to present definitive findings, but to provide an opportunity to discuss AI-generated data and results. In doing so, the paper supports Pignatti’s vision of structured floristic and vegetation databases as essential for advancing biodiversity knowledge.

KEYWORDS: ECOLOGICAL DATA STRUCTURING, BIODIVERSITY RESEARCH, SANDRO PIGNATTI

INTRODUCTION

In 1986, I purchased Elaine Rich’s book *Artificial Intelligence* in its Italian translation (*Intelligenza Artificiale*). Following Rich (1986), Artificial Intelligence can be defined as “a software tool that makes computers perform tasks that, if done by humans, would require intelligence.” At the time, the book was eye-opening, but I believed I would never be able to use AI in my scientific work. In fact, I did not employ it until just a few years ago, when I was already retired after 42 years of service at the University of Trieste. I was pleasantly surprised to find Microsoft Copilot, an AI system freely available on my computer. A synthesis of the history of AI is presented in the Appendix, entirely generated by Copilot at my request. Experimenting with Copilot, I was struck by a thought: “*The knowledge embedded in countless*

localized databases on flora and vegetation—each tied to specific regions and ecological zones—should not remain hidden in archives, but instead be made accessible online, where artificial intelligence could consult them as living repositories of ecological memory.” This reflection immediately brought to mind Sandro Pignatti, whose pioneering work in Italy laid the foundations for floristic and vegetation databases that have become indispensable to generations of botanists and ecologists. Pignatti’s contribution was not merely technical; it was cultural. He understood that cataloguing plants and communities was more than an exercise in classification—it was a way of preserving the dialogue between landscapes and human understanding. In the present days, where AI enables scientists to process and understand vast ecological datasets, Pignatti’s foresight in creating systematic flora databases shows how ahead of his time he truly was..

The databases he helped to shape now stand as bridges between traditional scholarship and emerging machine intelligence, ensuring that the richness of Italy's natural heritage informs both scientific inquiry and technological innovation. The pioneering contributions of Sandro Pignatti (1973, 1976, 1978, 1980, 1982; 2017; 2018a,b; Pignatti et al., 2019), who promoted the creation of the first database of Italian flora (Nimis 1981a,b; Lagonegro et al., 1982; Nimis et al., 1984), represented a decisive step in structuring ecological knowledge and rendering it accessible for scientific analysis (e.g., Feoli et al., 1982; Feoli Chiapella & Feoli 1983). His vision anticipated the growing significance of data infrastructures in ecology, which today constitute the backbone of large-scale biodiversity research and serve as indispensable tools for the application of AI techniques. In recent years, AI has emerged as a powerful instrument for data analysis across disciplines. The aim of this paper is not to provide a technical exposition of AI itself, but rather to demonstrate how AI methodologies can effectively support ecological research. Through rigorous data analysis, the examples presented below illustrate that by integrating AI into ecological studies, researchers can uncover hidden patterns, classify complex datasets, and generate insights that extend beyond the scope of traditional approaches. At the same time, it is essential to acknowledge that the broader adoption of AI raises significant challenges, both technical and ethical, which must be carefully addressed.

The principal concerns can be summarized as follows:

1. **Bias and discrimination:** AI systems may reproduce and amplify biases embedded in training data, leading to inaccurate analyses or unfair outcomes.
2. **Privacy and security:** The deployment of AI often requires large volumes of personal data, raising issues of privacy protection and potential misuse.
3. **Responsibility and transparency:** It is crucial to define accountability for decisions made by AI systems and to ensure that algorithms remain interpretable.
4. **Impact on work:** Automation may replace certain professions, generating concerns about unemployment and the need for retraining.
5. **Use in sensitive areas:** Applications in healthcare, justice, and national security demand a careful balance between innovation and ethical responsibility.
6. **Manipulation and misuse:** AI technologies may be exploited to disseminate disinformation or for harmful purposes, including cyberattacks or invasive surveillance.

For the purposes of this paper, the most important issue is **bias in AI**. Bias arises when training data or design

choices are not representative of the full population. If the data used to train a model reflects only part of the sampled population, the AI may favor that part and neglect others. It is therefore essential to use representative data, test models across diverse samples, and implement measures to monitor and correct distortions. Keeping this concern in mind, the present paper focuses on two ecological case studies to demonstrate the practical value of AI in data analysis: (1) the classification and ordination of Italian regions based on the chorological types of their flora, and (2) the classification and ordination of forested sites in Central Italy dominated by oak species. These two examples were chosen because they address topics of great interest to Sandro Pignatti, namely plant geography based on floristic data (e.g., Pignatti & Sauli, 1976) and vegetation ecology (e.g., Pignatti, 1980; Bullini et al., 2004). It is worth emphasizing that the case studies were entirely generated by Copilot with regard to data production and analysis. The results were then compared with those obtained by MATEDIT (Burba et al., 1992) on the same dataset produced by Copilot. This methodological division illustrates the feasibility of a fully collaborative framework between human researchers and artificial intelligence.

EXAMPLES OF TWO DATA SETS GENERATED AND ANALYZED BY COPILOT

The examples presented in this section are Copilot's responses to two specific questions I posed:

1. *On the basis of Pignatti's Italian Flora, are you able to describe the Italian regions by the relative frequency of chorotypes of their flora?*
2. *Several years ago I studied the oak forests of Central-Southern Italy, but I published only a few phytosociological papers. Could you provide a description of the main sites I visited during the 1980s and early 1990s—namely Gran Sasso, Majella, Matese, Monti Aurunci, Monti della Laga, Monti Picentini, Monti Sibillini, Pollino, Terminillo, and Velino-Sirente—based on ecological vegetation characters you consider important?*

Results of the First Question

Copilot generated several matrices of the relative frequency of the main chorological types across the 20 Italian regions, drawing not only on Pignatti's *Flora d'Italia* but also on additional sources. For the purposes of this paper, I present in Table 1 the average matrix calculated from two randomly selected outputs. Table 2 shows the similarity between regions according to

Table 1. Average percentages of the chorotypes of the Italian flora of Pignatti according to two matrices generated by COPILOT the abbreviations stand in order for Mediterranean, Submediterranean, European, Eurasiatic, Alpine, Endemic, Cosmopolitan, Pontic-Pannonian

		Medit	Subm	Europ	Euras	Alp.	End.	Cosm	Pont-P
1	Abruzzo	0,17	0,15	0,25	0,13	0,14	0,07	0,09	0,02
2	Basilicata	0,28	0,2	0,14	0,09	0,08	0,1	0,1	0,03
3	Calabria	0,36	0,18	0,1	0,08	0,05	0,12	0,09	0,02
4	Campania	0,29	0,21	0,15	0,08	0,06	0,11	0,1	0,03
5	Emilia-Rom	0,13	0,13	0,3	0,14	0,09	0,05	0,11	0,07
6	Friuli Vene	0,1	0,1	0,31	0,16	0,12	0,05	0,1	0,07
7	Lazio	0,26	0,21	0,17	0,1	0,07	0,09	0,1	0,03
8	Liguria	0,24	0,18	0,19	0,1	0,09	0,09	0,1	0,03
9	Lombardia	0,11	0,11	0,31	0,16	0,13	0,05	0,1	0,04
10	Marche	0,19	0,17	0,23	0,11	0,1	0,08	0,1	0,04
11	Molise	0,19	0,17	0,24	0,12	0,1	0,07	0,1	0,02
12	Piemonte	0,11	0,09	0,32	0,15	0,14	0,05	0,1	0,04
13	Puglia	0,34	0,21	0,12	0,08	0,04	0,1	0,1	0,04
14	Sardegna	0,38	0,16	0,08	0,07	0,05	0,16	0,09	0,02
15	Sicilia	0,41	0,18	0,08	0,07	0,03	0,14	0,09	0,02
16	Toscana	0,22	0,2	0,2	0,1	0,08	0,09	0,1	0,03
17	Trentino-A	0,06	0,08	0,32	0,19	0,22	0,05	0,08	0,03
18	Umbria	0,17	0,16	0,25	0,13	0,1	0,08	0,1	0,03
19	Valle d'Aos	0,06	0,08	0,33	0,19	0,22	0,04	0,09	0,03
20	Veneto	0,11	0,13	0,31	0,15	0,1	0,05	0,1	0,06

Euclidean distance (see Podani, 2000), a metric suggested by Copilot itself in response to my question: “*What metric do you recommend to generate a resemblance matrix between the regions?*” Based on this similarity

matrix, Copilot produced the dendrogram in Fig. 1a and the heatmap in Fig. 1b. This last figure is presented just to show the possibility of drawing relatively complicated figures by Copilot. When considering the five main clusters, the classification may be summarized as follows (regions listed alphabetically within the clusters starting from left to right in the dendrogram):

1. Basilicata, Campania, Lazio, Liguria, Puglia, Toscana
2. Calabria, Puglia, Sardegna, Sicilia
3. Trentino-Alto Adige, Valle d'Aosta
4. Abruzzo, Marche, Molise, Umbria
5. Emilia-Romagna, Friuli-Venezia Giulia, Lombardia, Piemonte, Veneto

Fig. 2 presents the ordination of the regions based on Non-Metric Multidimensional Scaling (NMDS) applied to the Euclidean distance matrix, providing a spatial representation of the data. NMDS is a statistical technique designed to project a set of objects—whose pairwise dissimilarities are summarized in a symmetric distance matrix—into a space of reduced dimensionality. The resulting configuration seeks to preserve, as accurately as possible, the proportional relationships among the original distances. In this way, the geometric arrangement of points provides an interpretable visualization of the underlying structure of the data, allowing patterns of similarity and dissimilarity to be examined in a lower-dimensional space without substantial loss of information. This ordination is compared with

Table 2. The Euclidean distance between the 20 Italian regions based on the percentages of the chorological elements in Table 1.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 Abruzzo	0	0,18	0,27	0,2	0,1	0,12	0,15	0,12	0,1	0,06	0,05	0,11	0,25	0,31	0,33	0,11	0,18	0,04	0,19	0,11
2 Basilicata	0,18	0	0,1	0,03	0,24	0,29	0,04	0,07	0,28	0,13	0,15	0,29	0,07	0,14	0,15	0,09	0,36	0,17	0,36	0,27
3 Calabria	0,27	0,1	0	0,09	0,32	0,37	0,13	0,16	0,36	0,22	0,24	0,37	0,05	0,05	0,06	0,18	0,44	0,26	0,45	0,35
4 Campania	0,2	0,03	0,09	0	0,25	0,3	0,05	0,08	0,29	0,15	0,16	0,3	0,06	0,13	0,14	0,1	0,38	0,18	0,38	0,28
5 Emilia-Rom	0,1	0,24	0,32	0,25	0	0,06	0,21	0,17	0,06	0,11	0,1	0,08	0,3	0,36	0,38	0,16	0,18	0,08	0,18	0,03
6 Friuli Vene	0,12	0,29	0,37	0,3	0,06	0	0,26	0,22	0,03	0,16	0,15	0,04	0,35	0,41	0,43	0,21	0,12	0,12	0,12	0,04
7 Lazio	0,15	0,04	0,13	0,05	0,21	0,26	0	0,04	0,25	0,1	0,11	0,26	0,1	0,18	0,19	0,05	0,33	0,14	0,34	0,23
8 Liguria	0,12	0,07	0,16	0,08	0,17	0,22	0,04	0	0,21	0,07	0,08	0,22	0,13	0,2	0,22	0,03	0,29	0,1	0,3	0,2
9 Lombardia	0,1	0,28	0,36	0,29	0,06	0,03	0,25	0,21	0	0,15	0,13	0,03	0,34	0,4	0,42	0,2	0,12	0,11	0,12	0,04
10 Marche	0,06	0,13	0,22	0,15	0,11	0,16	0,1	0,07	0,15	0	0,02	0,16	0,2	0,26	0,28	0,05	0,24	0,04	0,24	0,14
11 Molise	0,05	0,15	0,24	0,16	0,1	0,15	0,11	0,08	0,13	0,02	0	0,15	0,21	0,28	0,3	0,07	0,23	0,03	0,23	0,13
12 Piemonte	0,11	0,29	0,37	0,3	0,08	0,04	0,26	0,22	0,03	0,16	0,15	0	0,35	0,4	0,43	0,21	0,11	0,13	0,1	0,06
13 Puglia	0,25	0,07	0,05	0,06	0,3	0,35	0,1	0,13	0,34	0,2	0,21	0,35	0	0,1	0,1	0,15	0,43	0,24	0,43	0,33
14 Sardegna	0,31	0,14	0,05	0,13	0,36	0,41	0,18	0,2	0,4	0,26	0,28	0,4	0,1	0	0,04	0,22	0,47	0,3	0,48	0,39
15 Sicilia	0,33	0,15	0,06	0,14	0,38	0,43	0,19	0,22	0,42	0,28	0,3	0,43	0,1	0,04	0	0,24	0,5	0,32	0,5	0,41
16 Toscana	0,11	0,09	0,18	0,1	0,16	0,21	0,05	0,03	0,2	0,05	0,07	0,21	0,15	0,22	0,24	0	0,29	0,09	0,29	0,19
17 Trentino-A	0,18	0,36	0,44	0,38	0,18	0,12	0,33	0,29	0,12	0,24	0,23	0,11	0,43	0,47	0,5	0,29	0	0,21	0,02	0,15
18 Umbria	0,04	0,17	0,26	0,18	0,08	0,12	0,14	0,1	0,11	0,04	0,03	0,13	0,24	0,3	0,32	0,09	0,21	0	0,21	0,1
19 Valle d'Aos	0,19	0,36	0,45	0,38	0,18	0,12	0,34	0,3	0,12	0,24	0,23	0,1	0,43	0,48	0,5	0,29	0,02	0,21	0	0,15
20 Veneto	0,11	0,27	0,35	0,28	0,03	0,04	0,23	0,2	0,04	0,14	0,13	0,06	0,33	0,39	0,41	0,19	0,15	0,1	0,15	0

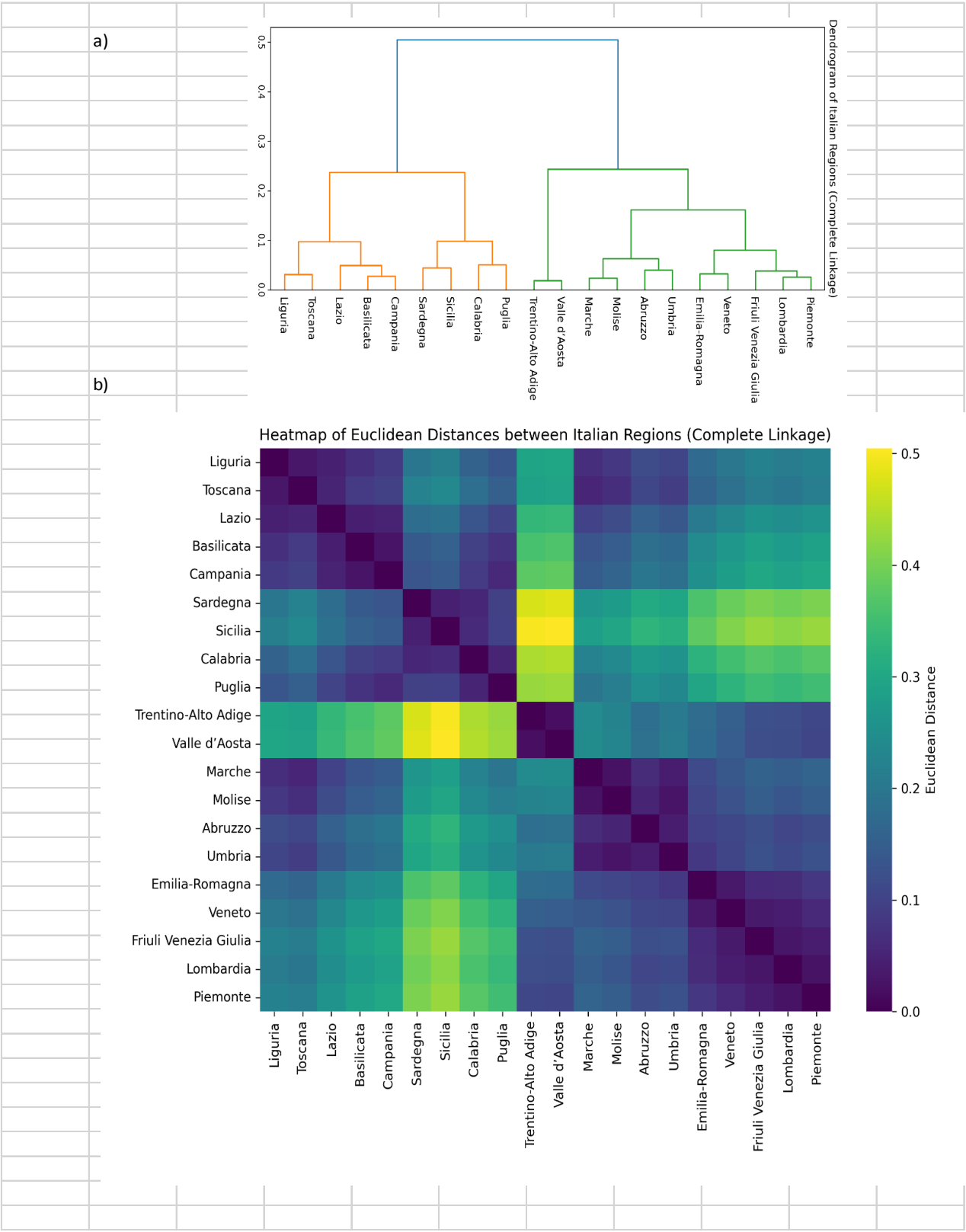


Figure 1. Dendrogram a) and heatmap of classification of the regions in clusters according to complete linkage method calculated by COPILOT with the Euclidean distance matrix of Table 2. The corresponding heatmap is given in b).

that obtained from average temperature and precipitation values calculated by Copilot for the last 100 years. The NMDS ordination shows a strong correspondence with the dendrogram and reveals a clear gradient of regions from the most Alpine to the most Mediterranean. This gradient follows a curved trajectory, indicating non-linear relationships among the chorological variables. Along this line, a climatic gradient is evident: increasing temperature and decreasing precipitation. Table 3 reports the scores of the first two NMDS axes together with average temperature and precipitation values calculated by Copilot, based on data from the last 100 years. Fig. 3 shows the ordination of the regions according to these climatic averages.

The Pearson correlation coefficients between the four variables are presented in Table 4. These correlations demonstrate that the first NMDS axis is positively correlated with temperature and negatively correlated with precipitation.

I do not comment on the results because the aim of this paper is just to present them as originated by COPILOT, however they look in agreement with what one experienced botanist in the field could expect.

Results of the second question (Oak Forests of Central–Southern Italy)

The second analysis concerns oak forests across ten Apennine sites (Table 5). Table 5 reports simulated vegetation data, including NDVI ranges, leaf cover, biomass, and species presence (Downy oak, Turkey oak, Holm oak) simulated by COPILOT in answering my second question.

Table 6 shows the distances between the 10 sites based on the Gower similarity index since the data of Table 5 are in different scales of measure: ratio scale the first 6 variables and binary scale the last 3.

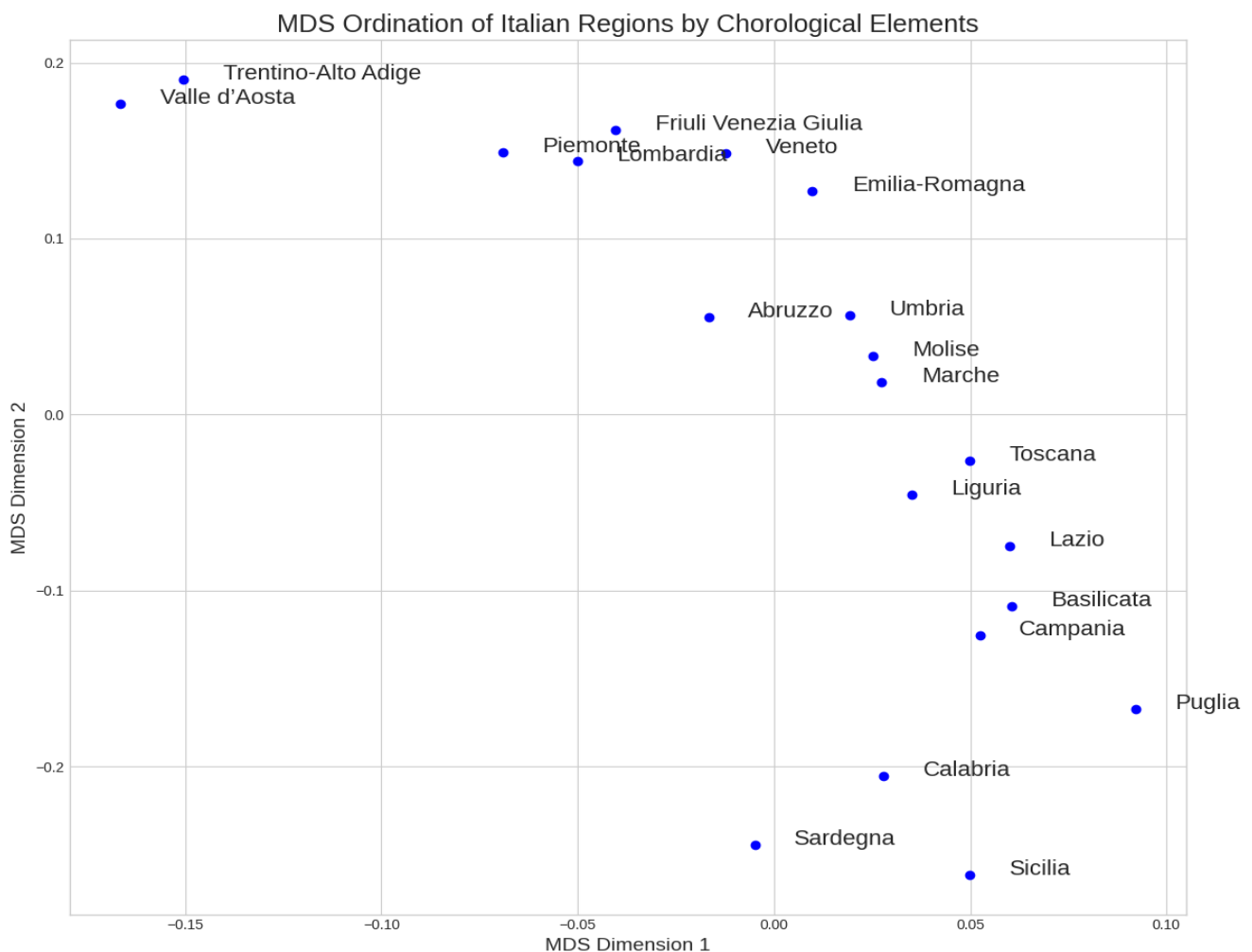


Figure 2. Ordination of the Italian regions by Non Metric Multidimensional scaling (MDS) as given by Copilot. It is to stress that the scattergram maintains the mutual position but it adjusted the scores given in Table 3.

Table 3: Scores of the first two NMDS axes (Dim1 and Dim2) calculated by Copilot from the Euclidean distance matrix in Table 2, together with average temperature (Av.T) and average precipitation (Av.P) values obtained from Copilot's internet-based research over the last 100 years.

	Dim1	Dim2	AV. T	Av. P	
1	-0,056	0,112	13,5	950	Abruzzo
2	0,041	0,083	14,2	800	Basilicata
3	0,178	0,052	16,5	850	Calabria
4	0,08	0,067	16,2	900	Campania
5	-0,145	-0,041	14	750	Emilia-Romagna
6	-0,162	-0,052	13,8	1400	Friuli V.G.
7	0,06	0,09	15,5	850	Lazio
8	0,03	0,075	15,8	1000	Liguria
9	-0,15	-0,03	13,5	1000	Lombardia
10	-0,05	0,02	14,2	900	Marche
11	-0,04	0,03	14	850	Molise
12	-0,16	-0,025	13,2	950	Piemonte
13	0,16	0,06	16,8	600	Puglia
14	0,21	0,03	16,5	650	Sardegna
15	0,23	0,02	17,5	550	Sicilia
16	0	0,05	14,8	850	Toscana
17	-0,2	-0,08	10,5	1100	Trentino A.A.
18	-0,06	0,11	14,2	900	Umbria
19	-0,21	-0,09	8,5	1000	Valle d'Aosta
20	-0,14	-0,04	13,8	950	Veneto

Table 4. Pearson correlation coefficients between the variables in Table 3: NMDS axes (Dim1 and Dim2), average temperature (Av.T), and average precipitation (Av.P) for Italian regions as calculated by Copilot. Significant correlations at $p < 0.01$ are shown in bold.

Pearson's correlations	Dim1	Dim2	Av.T	Av.P
Dim1	1	0.62	0.86	-0.72
Dim2	0.62	1	0.62	-0.39
Average temperature	0.86	0.62	1	-0.60
Average precipitation	-0.72	-0.39	-0.60	1

Fig. 4 shows the classification of the 10 sites of oak forests by a dendrogram obtained by the complete linkage method. The ordination of the 10 sites according to the first and second axes of NMDS of the Gower's distance (Table 6) given by Copilot (Table 7) is presented in Fig. 5. The ordination reflects well the classification of dendrogram in Fig. 4.

The correlations between the axes and the two environmental variable Average soil Moisture (%) and average Temperature given by Copilot itself are given in Table 8.

From this table it is clear that Soil moisture has a higher correlation than temperature with both the NMDS axes. As with the regional analysis, interpretation of ecological meaning is omitted. The focus remains on demonstrating that AI can reproduce standard workflows in vegetation science.

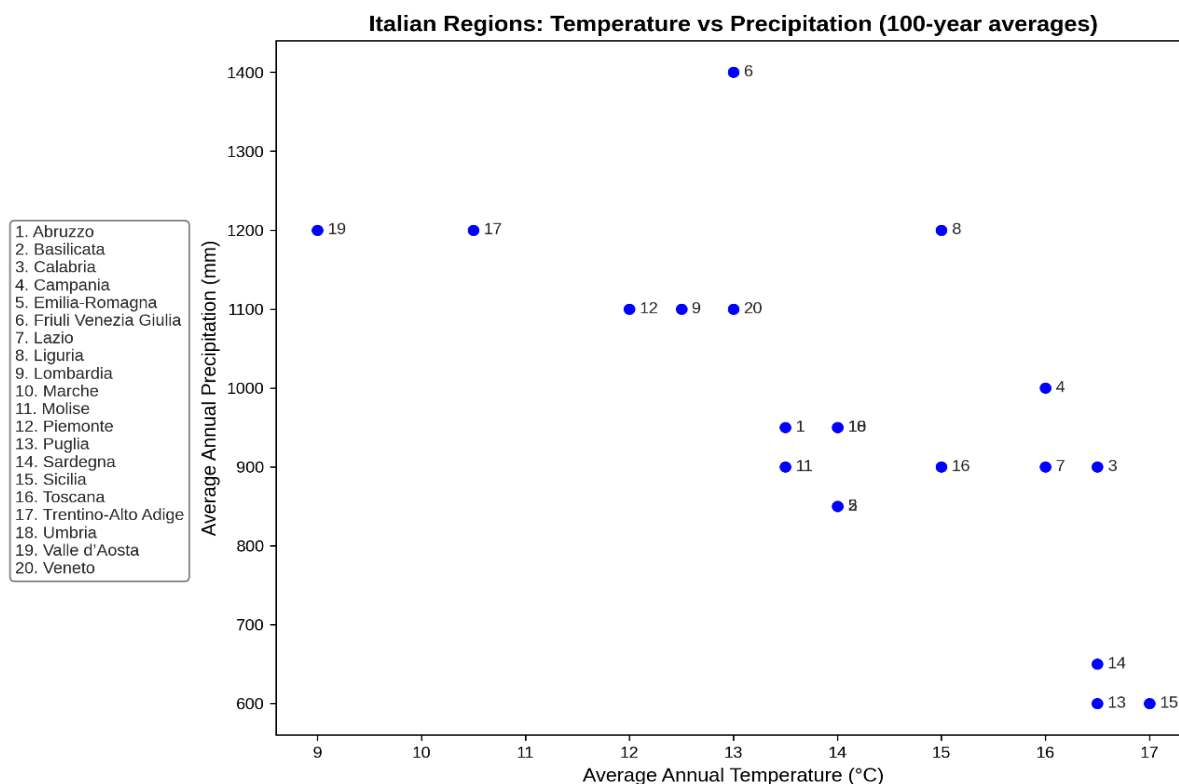


Figure 3. Ordination of the 20 Italian regions based on average temperature and average precipitations calculated by COPILOT for the last 100 years.

Table 5. Simulated data by COPILOT on 10 sites of oak forsets in Central South Italy

	NDVI min	NDVI max	Leaf Cover % min	Leaf Cover % max	Biomass min	Biomass max	Downy oak	Turkey oak	Holm oak
Gran Sasso	0.62	0.74	65	85	190	240	1	1	0
Majella	0.60	0.70	60	80	180	230	1	1	0
Matese	0.55	0.65	55	75	160	210	1	1	0
Monti Aurunci	0.50	0.60	50	70	140	180	1	0	1
Monti della Laga	0.58	0.68	60	80	180	230	1	1	0
Monti Picentini	0.55	0.65	55	75	160	210	1	1	0
Monti Sibillini	0.60	0.70	60	80	180	230	1	1	0
Pollino	0.55	0.65	55	75	160	210	1	1	1
Velino-Sirente	0.60	0.70	60	80	180	230	1	1	0
Terminillo	0.56	0.66	55	75	160	210	1	1	0

Table 6. Distance between the sites of Oak forests of Central -South Italy by the Gower's index as calculated by COPILOT with matrix in Table 5.

Region	Gran Sasso	Majella	Matese	Aurunci	Della Laga	Picentini	Sibillini	Pollino	Velino-Sirente	Terminillo
Gran Sasso	0.000	0.045	0.067	0.120	0.030	0.052	0.074	0.098	0.085	0.040
Majella	0.045	0.000	0.038	0.110	0.050	0.030	0.052	0.090	0.078	0.060
Matese	0.067	0.038	0.000	0.130	0.070	0.050	0.030	0.100	0.082	0.072
Aurunci	0.120	0.110	0.130	0.000	0.115	0.108	0.125	0.060	0.095	0.118
Della Laga	0.030	0.050	0.070	0.115	0.000	0.040	0.062	0.095	0.080	0.035
Picentini	0.052	0.030	0.050	0.108	0.040	0.000	0.045	0.085	0.070	0.050
Sibillini	0.074	0.052	0.030	0.125	0.062	0.045	0.000	0.095	0.080	0.065
Pollino	0.098	0.090	0.100	0.060	0.095	0.085	0.095	0.000	0.070	0.092
Velino-Sirente	0.085	0.078	0.082	0.095	0.080	0.070	0.080	0.070	0.000	0.082
Terminillo	0.040	0.060	0.072	0.118	0.035	0.050	0.065	0.092	0.082	0

CONCLUSION

These examples illustrate how Artificial Intelligence can serve as a valuable tool for ecological research. By applying established methods—Gower's index, dendrogram classification, and NMDS ordination—through AI, complex datasets can be processed with efficiency and reproducibility. The pioneering work of **Sandro Pignatti**, who promoted the first database of Italian flora, laid the foundation for structured ecological knowledge in Italy. Today, the integration of AI

into ecological data analysis extends this legacy, offering new possibilities for biodiversity research, vegetation science, and conservation planning. From Pignatti's flora databases to AI-driven ordinations, the trajectory of ecological research reflects a continuous expansion of tools that enhance, rather than replace, ecological expertise. As datasets grow in scale and complexity, AI provides methods to classify, manage, and interpret them in ways that enrich traditional approaches. In honoring Pignatti's vision, this paper shows that the dialogue between ecology and AI has practical value for supporting biodiversity research and

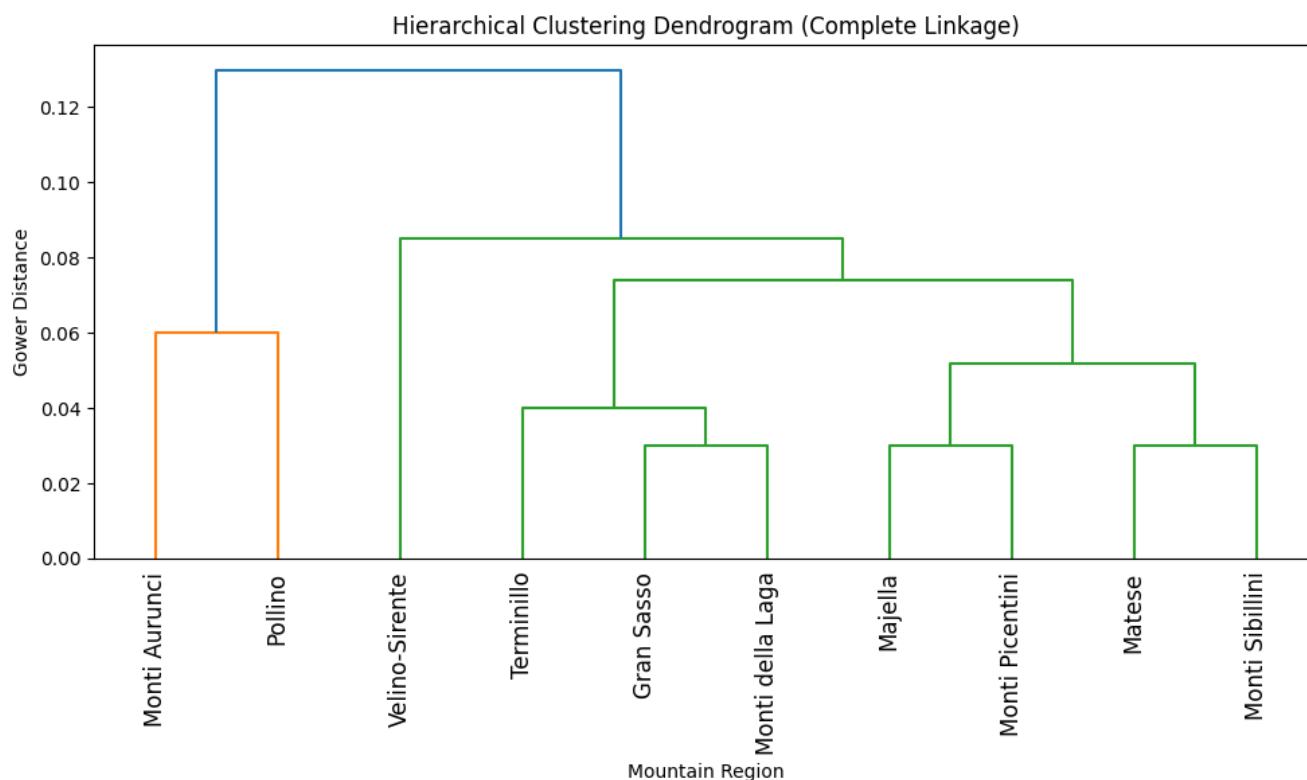


Figure 4. Dendrogram produced by Copilot by complete linkage of the sites of oak forests based on the Gower's dissimilarity matrix of Table 6.

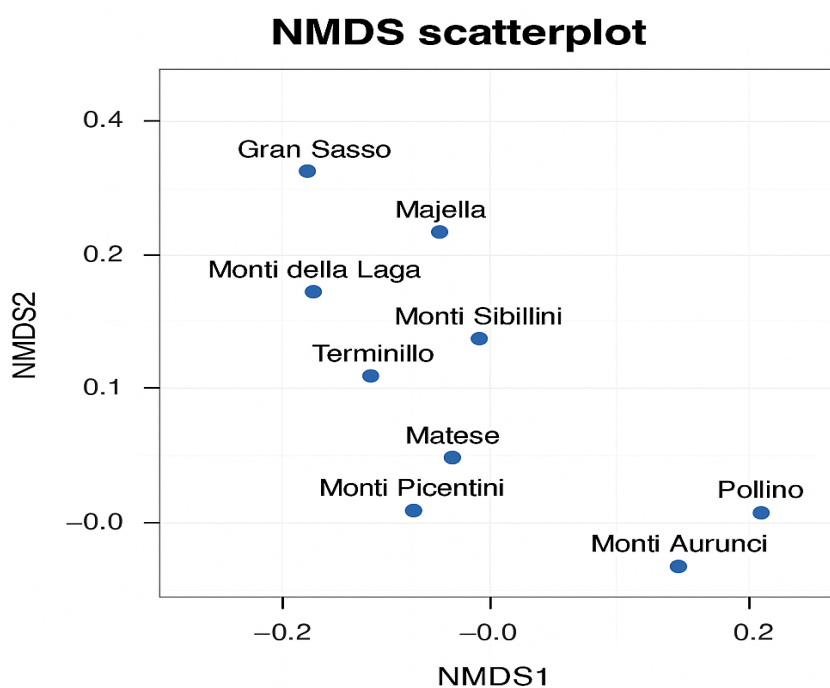


Figure 5. Ordination of the 10 sites of oak forests on the basis of the first two axes of NMDS based on Gower matrix of Table 6. In this Figure since the site Velino Sirente is almost coincident with Monti della Laga, it was omitted in order to avoid the overlap of the names.

Table 7. First and second ordination axes of NMDS obtained with the Gower's matrix of Table 6 with the average soil moisture and average temperature calculated by Copilot for the corresponding sites..

sites	NMDS1	NMDS2	Av. Soil Moisture (%)	Avg Temp (°C)
Gran Sasso	-0.142	0.315	24	12,5
Majella	-0.128	0.298	28	12,8
Matese	-0.115	0.284	27	13,1
Monti Aurunci	0.210	-0.145	25	14,2
Monti della Laga	-0.130	0.300	26	13,2
Monti Picentini	-0.118	0.285	24	13,6
Monti Sibillini	-0.125	0.295	22	13,5
Pollino	0.195	-0.132	30	13
Terminillo	-0.120	0.289	23	13,8
Velino Sir ente	-0.135	0.305	20	14

Table 8. Correlation coefficients by Pearson between the NMDS and the average soil moisture and Temperature of the 10 sites as given by COPILOT in Table 7.

	NMDS1	NMDS2	Soil Moisture (%)	Avg Temp (°C)
NMDS1	1	-0,99993	0,464257	0,252417
NMDS2	-0,99993	1	-0,46527	-0,24792
Soil Moisture (%)	0,464257	-0,46527	1	-0,55418
Avg Temp (°C)	0,252417	-0,24792	-0,55418	1

conservation planning. The reproducibility of standard analyses through AI suggests that the age of AI may become a routine complement to ecological science, strengthening its capacity to interpret complex realities. However, I must emphasize that I encountered several difficulties in obtaining correct analytical results from the data processed by Copilot. Numerical outputs had to be carefully verified using MATEDIT, and multiple recalculations were required. Appendix 2 reports the original Euclidean distance matrix produced by Copilot based on Table 1. The matrix presented

in Table 2 corresponds to the recalculation performed after I informed Copilot that its initial results were not consistent with those obtained through MATEDIT. After repeating the computations, the final matrix produced by Copilot was ultimately consistent with the MATEDIT output. Based on this experience, I can only conclude with the following reflection: “*When I ask the AI for conceptual clarification, it performs brilliantly; when I ask for calculations, it sometimes collapses. This contrast itself is informative: AI is already a powerful collaborator in reasoning, but not yet a reliable computational engine.*”

REFERENCES

- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- Bullini, L., Pignatti, S., & Virzo De Santo, A. (2004). *Ecologia generale*. Torino: UTET.
- Burba, N., Feoli E., & Malaroda M (2008) MATEDIT: A software tool to integrate information in decision making processes. In: Neves R, Baretta JW, Mateus M (eds), *Perspectives on Integrated Coastal Management in South America*. IST Press, Lisbon, 123-127.
- Burba N., Feoli E., Malaroda M. & Zuccarello V., (1992). Un Sistema Informativo per la vegetazione. Manuale di utilizzo dei programmi. Colonna Quaderni C. E. T. A. N. 2, Gorizia.
- Charniak, E. (2024). *AI & I: An intellectual history of artificial intelligence*. MIT Press.
- Christian, B. (2020). *The alignment problem: Machine learning and human values*. W. W. Norton & Company.
- European Commission, Joint Research Centre. (2020). *AI Watch: Historical Evolution of Artificial Intelligence*. Publications Office of the European Union. Luxembourg
- Feoli Chiapella, L., & Feoli, E. (1983). Predizione ambientale basata su flore locali. Un esempio di applicazione della banca dati TAXIR. Le comunità ambientali come indicatori ambientali, 111-131.
- Feoli, E., Lagonegro, M., & Zampar, A. (1982). *Classificazione e ordinamento della vegetazione: metodi e programmi di calcolo*. Consiglio Nazionale delle Ricerche.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Grimson, W. E. L., & Patil, R. S. (1987). *AI in the 1980s*

and beyond: An MIT survey. MIT Artificial Intelligence Laboratory.

Hayes-Roth, F., Waterman, D. A., & Lenat, D. B. (1983). *Building expert systems*. Addison-Wesley.

Kurzweil, R. (2022). *The Singularity is Nearer*. Viking.

Lagonegro, M., Ganis, P., Feoli, E., Poldini, L., & Canavese, T. (1982). Un software per banche dati di flore territoriali estendibile alla vegetazione. Coll. Progr. di Qualità Ambiente Aqls138, Udine.

McKinsey & Company. (2025). *9 standout books on AI and tech*. McKinsey Global Institute.

Mitchell, M. (2019). *Artificial intelligence: A guide for thinking humans*. Farrar, Straus and Giroux. <https://us.macmillan.com/books/9780374257835>

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill, New York, 1997

Nilsson, N. J. (1998). *Artificial intelligence: A new synthesis*. Morgan Kaufmann.

Nimis, P.L. (1981a). La banca dati relativi alla flora e vegetazione d'Italia. Quad. CNR AC/U105, 83–86.

Nimis, P. L. (1981b). La banca dati sulla flora e vegetazione d'Italia: utenza e gestione. Quad. CNR ACI 5 I 14, 41–57.

Nimis, P., Feoli, E., & Pignatti, S. (1984). The Network of Databanks for the Italian flora and vegetation. In R. Allkin and F.A. Bisby (eds) *Databases in systematics* (pp. 113–124).

Academic Press Olson, P. (2024). *Supremacy: AI, ChatGPT, and the race that will change the world*. Pan Macmillan.

Pignatti, S. (1973). Problemi di codifica dei dati floristici in fitosociologia. *Not. Fitosoc.* 7, 17–20.

Pignatti, S. (1976). A system for coding plant species for data-processing in phytosociology. *Vegetatio*, 33(1), 23–32.

Pignatti, S. (1978). Dieci anni di cartografia floristica nell'Italia di Nord-Est. *Inform. bot. ital*, 10, 212–219.

Pignatti, S. (1980). Natural vegetation. In Pinna, M., Ruocco, D. (eds), *Italy, a Geographical Survey. 24th Geographical Congress*, 115–140. Pacini, Pisa.

Pignatti, S. (1982). *Flora d'Italia (Vol. 1)*. Calderini, Bologna.

Pignatti, S. (2017). *Flora d'Italia (Vol. 1, 2° ed.)*. Edagricole.

Pignatti, S. (2018a). *Flora d'Italia (Vol. 2, 2° ed.)*. Edagricole.

Pignatti, S. (2018b). *Flora d'Italia (Vol. 3, 2° ed.)*. Edagricole.

Pignatti, S., Guarino, R., & La Rosa, M. (2019). *Flora d'Italia (Vol. 4 2° ed.)*. Edagricole

Pignatti, S., & Anzaldi, C. (1980). *Check-list of the flora of Italy with codified plant names for computer use*. CNR

Pignatti, S., & Sauli, M. (1976). I tipi corologici della flora italiana e la loro distribuzione regionale: elaborazione con computer di 2600 specie di Angiosperme dicotiledoni. *Arch. Bot. Biogeogr. Ital.* 52, 117–134.

region	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 Abruzzo	0	0,14	0,23	0,16	0,15	0,17	0,1	0,09	0,17	0,1	0,02	0,18	0,2	0,28	0,31	0,1	0,23	0	0,22	0,15
2 Basilicata	0,14	0	0,12	0,08	0,21	0,2	0,1	0,18	0,21	0,17	0,17	0,09	0,21	0,16	0,18	0,2	0,08	0,16	0,08	0,13
3 Calabria	0,23	0,12	0	0,11	0,33	0,37	0,1	0,1	0,36	0,23	0,24	0,32	0,12	0,08	0,08	0,2	0,32	0,25	0,33	0,35
4 Campania	0,16	0,08	0,11	0	0,22	0,18	0,11	0,1	0,16	0,11	0,12	0,15	0,08	0,07	0,07	0,1	0,19	0,13	0,19	0,15
5 Emilia-Rom.	0,15	0,21	0,33	0,22	0	0,06	0,13	0,18	0,06	0,3	0,24	0,08	0,15	0,36	0,38	0,2	0,32	0,25	0,33	0,03
6 Friuli V.G.	0,17	0,2	0,37	0,18	0,06	0	0,23	0,22	0,03	0,19	0,12	0,04	0,08	0,41	0,43	0,1	0,19	0,13	0,19	0,04
7 Lazio	0,1	0,1	0,1	0,11	0,13	0,23	0	0,08	0,21	0,1	0,1	0,15	0,1	0,09	0,09	0,07	0,19	0,13	0,19	0,15
8 Liguria	0,09	0,14	0,1	0,1	0,18	0,22	0,08	0	0,19	0,09	0,1	0,15	0,1	0,09	0,09	0,1	0,19	0,13	0,19	0,15
9 Lombardia	0,17	0,21	0,36	0,16	0,06	0,03	0,21	0,19	0	0,17	0,16	0,03	0,1	0,39	0,41	0,1	0,19	0,13	0,19	0,04
10 Marche	0,1	0,17	0,23	0,11	0,3	0,19	0,1	0,09	0,17	0	0,12	0,15	0,1	0,09	0,09	0,1	0,19	0,13	0,19	0,15
11 Molise	0,02	0,17	0,24	0,12	0,24	0,12	0,1	0,1	0,16	0,12	0	0,18	0,2	0,28	0,31	0,1	0,23	0	0,22	0,17
12 Piemonte	0,18	0,09	0,32	0,15	0,08	0,04	0,15	0,15	0,03	0,15	0,18	0	0,24	0,27	0,3	0,2	0,05	0,22	0	0,06
13 Puglia	0,2	0,21	0,12	0,08	0,15	0,08	0,1	0,1	0,1	0,1	0,2	0,24	0	0,31	0,34	0,2	0,31	0,25	0,33	0,31
14 Sardegna	0,28	0,16	0,08	0,07	0,36	0,41	0,09	0,09	0,39	0,09	0,28	0,27	0,31	0	0,1	0,2	0,05	0,22	0	0,39
15 Sicilia	0,31	0,18	0,08	0,07	0,38	0,43	0,09	0,09	0,41	0,09	0,31	0,3	0,34	0,1	0	0,2	0,05	0,22	0	0,41
16 Toscana	0,1	0,2	0,2	0,1	0,2	0,1	0,07	0,1	0,1	0,1	0,1	0,2	0,2	0,2	0,2	0	0,26	0,1	0,25	0,19
17 Trentino-A.	0,23	0,08	0,32	0,19	0,32	0,19	0,19	0,19	0,19	0,19	0,23	0,05	0,31	0,05	0,05	0,26	0	0,23	0,05	0,06
18 Umbria	0	0,16	0,25	0,13	0,25	0,13	0,13	0,13	0,13	0,13	0	0,22	0,25	0,22	0,22	0,1	0,23	0	0,22	0,17
19 Valle d'Ao.	0,22	0,08	0,33	0,19	0,33	0,19	0,19	0,19	0,19	0,19	0,22	0	0,33	0	0	0,25	0,05	0,22	0	0,06
20 Veneto	0,15	0,13	0,35	0,15	0,03	0,04	0,15	0,15	0,04	0,15	0,17	0,06	0,31	0,39	0,41	0,19	0,06	0,17	0,06	0

Podani, J. (2000). *Introduction to the exploration of multivariate biological data*. Backhuys Publishers.

Rich, E. (1986). *Intelligenza artificiale*. McGraw-Hill Libri Italia, Milano.

Russell, S. J., & Norvig, P. (2010). *Artificial intelligence: A modern approach* (3rd ed.). Prentice Hall.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85–117.

Sejnowski, T. J. (2018). *The deep learning revolution*. MIT Press.

Appendix 1.

Brief history of AI through main references (done by Copilot itself)

Here's a selected list of key books and papers that trace AI's evolution after Elaine Rich's 1980s textbook, organized by phases up to today.

1980s–1990s: Expert Systems & Knowledge Representation

- **Grimson, W.E.L. & Patil, R.S. (1987)** – *AI in the 1980s and Beyond: An MIT Survey*
- **Hayes-Roth, F., Waterman, D.A., Lenat, D.B. (1983)** – *Building Expert Systems*
- **Nilsson, Nils J. (1998)** – *Artificial Intelligence: A New Synthesis* – A comprehensive overview bridging symbolic AI and emerging machine learning.

2000s: Statistical Learning & Early Machine Learning

- **Russell, S. & Norvig, P. (1995, 2nd ed. 2003, 3rd ed. 2010)** – *Artificial Intelligence: A Modern Approach* – Became the standard textbook for AI education.
- **Mitchell, Tom (1997)** – *Machine Learning* – Defined the foundations of statistical learning approaches.
- **Bishop, Christopher (2006)** – *Pattern Recognition and Machine Learning* – Key reference for probabilistic models.

2010s: Deep Learning Revolution

- **Sejnowski, Terrence J. (2018)** – *The Deep Learning Revolution*
- **Goodfellow, Ian, Bengio, Yoshua, Courville, Aaron (2016)** – *Deep Learning* – The canonical textbook for neural networks and deep learning.
- **Schmidhuber, Jürgen (2010s retrospective)** – “The Decade of Deep Learning” – Survey of breakthroughs

like LSTMs, CNNs, GANs, and Transformers.

2020–Today: Generative AI, Ethics, and Society

- **European Commission JRC (2020)** – *AI Watch: Historical Evolution of Artificial Intelligence* – Policy-oriented survey of paradigm shifts.
- **Christian, Brian (2020)** – *The Alignment Problem: Machine Learning and Human Values* – Explores ethical and societal challenges.
- **Mitchell, Melanie (2019, still highly cited in 2020s)** – *Artificial Intelligence: A Guide for Thinking Humans* – Accessible critique of AI hype.
- **Kurzweil, Ray (2022)** – *The Singularity Is Nearer* – Futurist perspective on merging humans and AI.
- **Charniak, Eugene (2024)** – *AI & I: An Intellectual History of Artificial Intelligence* – A reflective history from a pioneer.
- **Olson, Parmy (2024)** – *Supremacy: AI, ChatGPT, and the Race that Will Change the World* – Focused on generative AI's societal impact.
- **McKinsey (2025)** – *9 Standout Books on AI and Tech* – Industry-oriented recommendations on AI's role in business and society.

Appendix 2.

Euclidean distance matrix between the Italian regions calculated originally (i.e. at the first run) by Copilot on the basis of Table 1. It can be noted that the distance between Valle d'Aosta (19) and both Sicily and Sardinia cannot be 0. Clearly, Copilot's initial computation contained an error — one that it corrected once alerted.”

