



Article

From Weeds to Feeds: Exploring the Potential of Wild Plants in Horticulture from a Centuries-Long Journey to an AI-Driven Future

Diego Rivera ^{1,*} , Diego-José Rivera-Obón ² , José-Antonio Palazón ³ and Concepción Obón ⁴

- ¹ Departamento de Biología Vegetal, Facultad de Biología, Campus de Espinardo, Universidad de Murcia, 30100 Murcia, Spain
- ² Faculté Jean Monnet, Université Paris-Saclay, 54 Boulevard Desgranges, 92230 Sceaux, France; dieguitojos@gmail.com
- ³ Departamento de Ecología e Hidrología, Facultad de Biología, Campus de Espinardo, Universidad de Murcia, 30100 Murcia, Spain; palazon@um.es
- ⁴ CIAGRO, EPSO, Universidad Miguel Hernández de Elche, 03312 Orihuela, Spain; cobon@umh.es
- * Correspondence: drivera@um.es; Tel.: +34-868884994

Abstract: Given the increasing food needs of humanity and the challenges cultivated species face in adapting to the climatic uncertainties we experience, it is urgent to cultivate new species. A highly relevant repertoire for this purpose is offered by the array of edible wild plants. We analyzed data from Murcia (Spain), involving 61 species and 59 informants, and the Global Database of Wild Food Plants, which includes 15,000 species, 500 localities, and nearly 700 references. Using local consensus, global distribution, and GBIF occurrence data, we built simple unimodal or bimodal models to explore their limitations. Our study highlights that approximately 15,000 wild or feral plant species are consumed as food, underlining the urgent need to support existing crops with new species due to current food crises and climate irregularities. We examined wild plant diversity from a horticultural perspective, considering their relationships with weeds and invasive species. Partial criteria, such as local consensus or global use, were found insufficient for selecting candidate species. We propose developing a specific artificial intelligence to integrate various factors—ecological, nutritional, toxicological, agronomic, biogeographical, ethnobotanical, economic, and physiological—to accurately model a species' potential for domestication and cultivation. We propose the necessary tools and a protocol for developing this AI-based model.

Keywords: wild food plants; gathered food; local food; nutraceuticals



Citation: Rivera, D.; Rivera-Obón, D.-J.; Palazón, J.-A.; Obón, C. From Weeds to Feeds: Exploring the Potential of Wild Plants in Horticulture from a Centuries-Long Journey to an AI-Driven Future. *Horticulturae* **2024**, *10*, 1021. <https://doi.org/10.3390/horticulturae10101021>

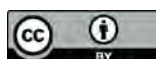
Academic Editors: Boris Duralija and Dario Kremer

Received: 29 July 2024

Revised: 19 September 2024

Accepted: 20 September 2024

Published: 25 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Despite the global abundance of food, it is estimated that over 840 million people in the developing world suffer from chronic undernourishment. Many more experience deficiencies in protein and essential micronutrients such as iodine, vitamins, and iron [1]. The issue of food security, which gained prominence in the 1970s, continues to be a major global concern. Initially, the focus was on ensuring a stable supply of staple foods in the international market. However, global food availability does not equate to food security for individual countries, particularly for famine-stricken nations that often lack the foreign currency to access food from the world market.

Wild food plants, which historically formed the basis of the diet for hunter-gatherer societies, still play a significant role as a dietary supplement in agricultural communities. Their relevance in these societies depends largely on how well dietary needs are met through conventional farming. The diversity of wild food plants far exceeds that of major crops, making them a potentially crucial resource in future food crises. While some are only consumed during extreme events like famines, they provide critical support in times of need and contribute to enhancing global food security. Rural communities, particularly

those in areas with seasonal food shortages, continue to consume wild plants, preserving valuable knowledge that could prove vital during food shortages. In some cases, wild plants may even evolve into new crops or be rediscovered as hidden gems, such as the “quelites” of Mexico. However, the toxicity of some wild plants poses challenges, as detoxification processes are often labor-intensive. Although domestication has reduced toxicity in many plants, it has also diminished their levels of beneficial compounds like antioxidants and nutrients.

Famine foods, or emergency foods, have received limited attention from scientists and international organizations. These plants are typically consumed when conventional food sources are unavailable, and though they are often unpalatable or require extensive processing, they play a crucial role in survival [2]. Certain famine foods, particularly those that endure severe shortages, are resilient to factors that affect the yields of preferred crops. The consumption of famine foods often reflects the severity of a food shortage. The classification of a plant as a famine food is influenced by biological, biochemical, and cultural factors [3]. Recent analyses suggest that many wild famine food species could be cultivated in catastrophic circumstances to replace staple crops [4].

Interestingly, even foods that are ritually tabooed can be consumed during times of extreme scarcity [2]. Cultural rituals and myths play a critical role in passing down knowledge about these plants, as observation alone may not suffice to preserve such information. For instance, Jews consume bitter herbs during Passover to remember their ancestors’ slavery, integrating these foods into traditional dishes. Disruptions to oral traditions, however, can result in the loss of knowledge about these low-preference foods, especially as their use becomes rarer [2].

In some regions, wild plants serve as “seasonal hunger foods”, consumed during periods of low food availability. The following are some examples. In India, for example, Handa [5] identified several wild plants, referred to as “mini food banks”, which have both nutritional and medicinal value. These plants, common in higher Himalayan households, contribute to food security and healthcare during severe winters that can last up to nine months. Across India, tribal groups use over 3900 wild species as food, most as seasonal hunger foods, and at least 250 species have been identified as having potential for development as alternative crops [6]. Similarly, the Sherpa in Tibet, China, collect wild edible plants almost year-round, except for January and February, to mitigate seasonal food shortages [7]. However, in Brazil, Gomes et al. [8] have argued against the seasonality hypothesis, demonstrating that wild fruits and reproductive plant parts are available year-round.

It is important to note that wild food plants are frequently consumed in mixtures of several species, sometimes numbering over 30 [9]. They are often used as seasoning or flavor enhancers in other foods [10,11]. In the Middle East, numerous wild plants are consumed in mixtures, either raw or prepared, providing a supplementary vegetarian diet for Bedouins and villagers. Many of these plants are also sold in local markets [12]. In Türkiye, for instance, villagers from Aksaray identified over 100 edible species out of 300 locally named plants. Similarly, in Mugla, around 140 uncultivated plants are used, often mixed with several species, for food from a documented 360 useful plants [13].

Optimal foraging theory predicts animal behavior based on the assumption that individuals strive to maximize their genetic fitness through efficient foraging. This theory posits that foraging behavior is heritable, evolves over time, and is constrained by factors like genetic linkages and environmental conditions. Animals are assumed to have access to information and make decisions that optimize their fitness, though these decisions may be influenced by constraints. This theory has been applied as a useful means to understand foraging by social insects [14]. One might be inclined to apply Optimal foraging theory as a model to explain the interactions between human communities and the wild plants they foraged, particularly to better understand regional preferences [15]. Nevertheless, optimal foraging theory can only be expected to be useful when its assumptions, mathematical development, and testing are appropriate for the studies to which it is applied [16].

Proposing the cultivation of species traditionally consumed as wild foods requires a solid foundation of knowledge and clear criteria. For instance, research by Gomes et al. has shown that the wild food plants used in Brazil are not chosen at random but are selectively consumed by different populations [17–19]. Not all wild species are used equally, and their use depends on a variety of factors, including the structural or circumstantial needs of the population and the degree to which these species are relied upon in daily diets. This study will explore these factors in detail.

In addition, the plant's palatability, nutritional content, health benefits, and potential toxicity are key considerations when assessing its suitability for cultivation. Many of these species are considered healthy by local populations, and factors such as ease of cultivation, seed abundance, and short germination periods also play an important role in determining their potential as future crops [20,21]. However, not all wild species exhibit these traits. Understanding the optimal bioclimatic and ecological conditions for each species, as well as its current distribution, is crucial. The fact that a species is consumed across distant localities suggests its potential for global use but also raises concerns about its potential to become invasive.

In addressing the challenge of developing new crops capable of withstanding the pressures of drastic climate change by utilizing wild species, we approach the issue through several sections and assumptions: (1) The most promising wild species for food are those already being harvested for this purpose. (2) Among wild species with extensive geographical distribution, those collected and consumed throughout their entire range are more relevant than those used in only a few localities. (3) To make globally significant decisions, it is essential to start with an exhaustive catalog of wild species harvested worldwide. (4) Undertaking a field study of such breadth from scratch is impractical. (5) Consequently, published information is crucial as a starting point, not only regarding species and their uses. (6) Local studies are fundamental, providing essential primary information. (7) At the local level, consensus data from informants offer insights into the popularity of different species. (8) While the local popularity of a species may suggest its potential for starting local cultivation, at a global scale, additional criteria are required. (9) Although verified usage data are important, they represent only a small fraction of the prior information necessary for a global-scale analysis. (10) Given the numerous factors involved, employing Artificial Intelligence tools, or even designing a specific tool to integrate various data sources, would be invaluable for the approximately 15,000 wild plant species currently consumed as food by humans. In a field involving thousands of species and millions of data points, the use of specific AI tools and a multidisciplinary team of experts becomes indispensable.

In this study, we will analyze examples of local knowledge and global distribution patterns, and propose a framework for an AI tool that can help select candidate species for cultivation under varying conditions. This research is vital for global food security and sustainable agriculture. By identifying and cultivating wild plants with high nutritional value, resilience, and adaptability, we can diversify diets, reduce dependence on traditional crops, and promote sustainable farming practices, ultimately contributing to a more secure and resilient global food system.

2. Materials and Methods

2.1. Research Design and Approach

This article is based on three primary areas of work: First, fieldwork conducted by teams from the University of Murcia, Miguel Hernández University, the University of Alicante, and the Botanical Institute of the University of Castilla La Mancha, which began in the 1980s and continues to the present day. Second, the comprehensive database on the global use of wild plant species as food (Global Database of Wild Food Plants) at the Universidad de Murcia. Third, specific models based on local consensus and global occurrences for delimiting a set of the species potentially relevant as a candidate for developing a novel crop or recovering those formerly abandoned.

Finally, we have analyzed the limitations of existing partial models and propose a comprehensive methodology to overcome these restrictions. This methodology includes tools and conventions for integrating multimodal data and explores the potential of using AI to evaluate thousands of wild food species as promising crops for future food production.

2.2. *Gathering Ethnobotanical Knowledge in the Field*

Our ethnobotanical research has resulted in various publications concerning the National Parks of Cabañeros and Tablas de Daimiel, the Serranía de Cuenca [22], and the provinces of Albacete, Alicante [23], and Murcia. The work involved open and semi-structured interviews with hundreds of informants, focusing on the collection of wild food plants, with prior informed consent obtained verbally from the informants. In general, the informants were rural inhabitants, especially farmers, livestock breeders, and housewives aged forty years or older, who were recognized in their locality as having a good knowledge of local plants and their traditional uses. Control sheets were prepared for all the species recorded, following the usual protocols of collection, pressing, drying, and labeling of plant material in the different zones, in at least one locality per zone. Voucher specimens were deposited in the herbaria ALBA (Botanical Institute, Castilla-La Mancha University, Albacete), UMH (Miguel Hernández University, Orihuela), and MUB (University of Murcia, Murcia). Details such as sheet number or herbarium code are specified in the different articles published and cited here. More information on these herbaria can be found at The New York Botanical Index Herbariorum. Over one hundred of these species have been the subject of phytochemical and pharmacological studies within the framework of collaborations and common projects. The findings from these studies have already been published [24,25]. Additionally, we have collaborated with other groups in studying wild food plants collected in the Alps [26,27], Basilicata [28], Tuscany, and Emilia Romagna (Italy) [29], Crete (Greece), Lebanon, and Syria. Overall, we have observed the simultaneous existence of common patterns in preferred species alongside local or regional patterns with preferences for certain species or specific types of plant-based foods.

For the specific purpose of the present study, we selected among these a catalog generated in 2004 for the Huerta de Murcia (Murcia, Spain) [30], interviewing 59 informants in 14 districts (Figure 1), with the aforementioned characteristics of living in rural areas, even though they are close to the city of Murcia, and of engaging in activities related to the use of the natural and agricultural environment, which reported 61 species. Our objective was to develop a framework aimed at establishing a hierarchical background to assess the relative potential of these species as viable food crops. The Huerta de Murcia is an extensive agricultural area located in the region of Murcia, southeastern Spain. Covering approximately 12,000 hectares, it is characterized by its fertile land primarily used for intensive horticulture, including the cultivation of vegetables, fruits, and flowers. This region benefits from a network of irrigation channels originally developed during the Arab period, which supports its agricultural productivity. The Huerta de Murcia is situated along the Segura River, contributing to its rich agricultural tradition and economic significance in the region [30].

2.3. *Building a Global Database of Wild Food Plants*

The Global Database of Wild Food Plants, at present, covers 58 regions in Africa, 56 in East and Central Asia, 24 in West Asia, 47 in the Pacific region, 31 in Europe in general, 77 in Mediterranean Europe, and 84 in the Americas (Figure 2), based on a literature review involving the analysis of 633 articles, books, and theses (Supplementary Table S1). The 633 articles were meticulously analyzed by one of the authors, D. Rivera, over a span of more than fifteen years, as this work was carried out alongside other research activities. The authors reviewed all relevant articles and books, focusing on wild food plants, gathered food plants, edible wild plants, and general ethnobotanical studies that included specific chapters on these topics. A comprehensive table was created using Excel, now comprising 377 columns (representing different zones) and 14,813 rows (representing

taxa). Standardization tools were applied to ensure consistency in scientific names, and the barycenter for each zone was calculated to georeference the data accurately. This allows us to understand the geographical distribution of use as food of nearly 15,000 plant species along the 377 areas (Figure 2). The basic dataset is the occurrence or not of each species as consumed in the different localities and regions analyzed, whose barycenters are represented in Figure 2.



Figure 1. Geographical Distribution of zones where our laboratory studied Wild Food Plants in the Huerta de Murcia (Murcia, Spain). 1. Algezares, 2. Alquerías, 3. Beniján, 4. Cabezo de Torres, 5. La Alberca, 6. La Arboleja, 7. Llano de Brujas, 8. Los Garres, 9. Monteagudo, 10. Puebla de Soto, 11. Puente Tocinos, 12. Rincón de Beniscornia, 13. Rincón de Seca, 14. Torreagüera. Image by Diego Rivera with a base map from Google Earth.

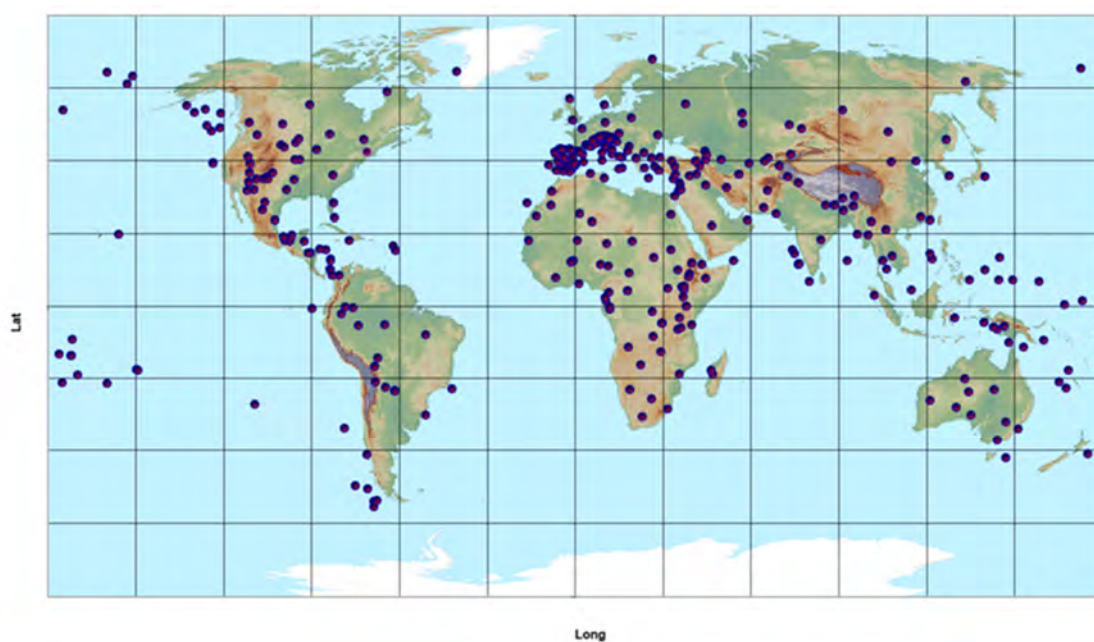


Figure 2. Geographical distribution of information sources for the Global Database of Wild Food Plants. Image by Diego Rivera.

2.4. Conceptual Analysis

We have approached the analysis of the concept of “wild edible plant” by breaking it down into two parts. First, we explore the various existing definitions of wild plants, focusing on the horticultural perspective. Following the steps of Levadoux [31] for wild grapevines, we generalize the categories to apply to all plant species. In developing this concept, we identify two points that need resolution: the categories of “weed” [32] and “invasive species”, which may sometimes overlap with certain wild species of interest for human consumption. After clarifying these aspects, we assess the plant’s significance as a food source, considering those who consume it regularly, occasionally, or only under exceptional circumstances. This assessment can influence the overall model for selecting species of interest.

2.5. Specific Models

Simple models that relate the number of species to the degree of local consensus or the number of species to the number of localities where their consumption has been recorded in the Global Database have been conducted using the graphing and regression tools available in Excel. The Lorenz curve [33–35] is a graphical representation used in economics and sociology to illustrate the distribution of income or wealth within a population. It is a tool used to visualize inequality. The area of concentration comprised between the line of perfect equality and the Lorenz curve indicates the degree of inequality. The larger the area, the greater the inequality. Implications: Here, we use the empirical approximation to the Lorenz curve to assess and address the incertitude of information and patterns of use and knowledge inequality between species. These allow comparison at different scales. For example, a region with a Lorenz curve far from the line of equality could be considered under-explored and requires further studies to reduce incertitude or inequality.

2.6. Proposal for a Specialized Model Combining Multimodal Data

We searched for tools and references providing a comprehensive starting point for developing a robust AI system to select novel crop candidates from wild plant species using the assistance of Google Scholar, Scopus, and PubMed as the main search tools. AI tools such as ChatGPT, Gemini, Perplexity, and Mistral [36–39] were used to analyze relevant tools and the main steps in the protocol.

3. Results and Discussion

3.1. Literature Background—Concept of Wild Plants and Their Potential Use as a Novel Crop

3.1.1. Steps in the Concept of Wild Plants

The term “wild plant” is polysemic. Ecologically, it refers to species that thrive in their natural habitats, playing crucial roles in biodiversity and ecosystem functioning. These plants are well-adapted to local environmental conditions, such as climate and soil, and contribute significantly to maintaining ecological balance by thriving in undisturbed ecosystems [40].

From a phylogenetic perspective, wild plants are species that have not been altered by human intervention, remaining subject to natural selection and retaining their original genetic makeup. This view emphasizes the absence of domestication-associated genetic changes [41,42]. However, as we will discuss, some ‘wild’ plants originate from previously domesticated populations that have become feral, occupying natural habitats and thus are not truly phylogenetically wild.

In anthropology, a wild plant refers to a species not intentionally cultivated or managed by human societies, with emphasis on the absence of deliberate cultivation [43]. These plants are primarily seen as objects of gathering rather than farming, and human perception of them is shaped by intentional management practices [44]. Plants lacking domestication traits—such as large fruits, uniformity, or specific cultivation needs—are more likely to be viewed as wild by foragers. This perception is closely linked to cultural, ecological, and

culinary contexts, as well as the knowledge and practices of the collectors in foraging and utilizing these plants.

The perception of a plant as ‘wild’ by collectors, particularly in the context of food use, is shaped by several factors. Plants are often regarded as wild based on their edibility and traditional knowledge passed down through generations. Species that are not commonly cultivated but known for their edibility and traditional uses are frequently considered wild edibles. Plants gathered through foraging from natural habitats—such as forests or meadows—are typically labeled as wild. Additionally, plants that are not part of regular agriculture but are valued in traditional cuisines are often perceived as wild ingredients due to their cultural significance.

From a horticultural perspective, a “wild plant” is commonly described as a plant species that grows spontaneously in natural habitats [44], without direct human cultivation or intentional management. These plants are not subjected to deliberate cultivation practices, unlike crops or cultivated species. They thrive in diverse ecosystems, adapting to various environmental conditions. Horticulture recognizes wild plants as potential crops, with increasing interest stemming from evidence of continued human consumption. The focus shifts to features indicating preadaptation to cultivation, acknowledging their potential as cultivated species. Curiously, a weed would be a ‘wild’ plant whose natural habitat is cultivated fields. This paradox is because the process of agricultural development has occupied the natural habitats of numerous species, particularly in the Mediterranean and Asia, hundreds of which, especially annuals and bulbous plants, have adapted to traditional agricultural practices. In addition to the above, we should add abandoned crops that are able to propagate autonomously in cultivated fields. Both are promising sources of future crops and presently are cryptocrops [32].

Wild plants consumed by humans are often referred to using terms that emphasize specific aspects of their use. For example, “gathered food plants” highlights the act of foraging [25,28], while “wild edible plants” underscores their potential as a food source [45]. The term “wild food plants” conveys a similar concept and will be the preferred term used throughout this paper.

Wild food plants’ availability often follows seasonal patterns [46]. Collectors may view plants as wild if they are part of the natural biodiversity of an area and exhibit seasonal variations in growth and availability. Ethical foragers consider sustainability in their practices. Plants collected sustainably from natural environments, without causing harm to ecosystems, are often seen as wild. Sustainable foraging ensures the long-term viability of both the plant species and the ecosystems they inhabit.

The term “wild plant” thus encompasses diverse perspectives. In summary, each perspective provides a unique lens through which wild plants are perceived, adding layers to their significance in the context of human interaction and cultivation practices.

3.1.2. Horticultural Perspective of the “Wild Plant”

Focusing on the horticultural approach, the concept of the “wild plant” is more complex than its apparent simplicity would lead one to expect. Here we present a globalized version of the concepts expressed by Levadoux [31] for wild grapevines or “lambrusques” (*V. sylvestris* complex) compared with their domesticated relative (*Vitis vinifera*), which has been expanded to encompass all wild plants:

We designate as ‘wild’ any plant that appears to grow in a state of nature, thus from an ecological perspective. Consequently, there can be several types of wild plants:

- Post-cultivated wild plants, which merely extend a previously domesticated line but have been left uncultivated thereafter.
- Sub-spontaneous wild plants that originate in uncultivated soil from seeds of cultivated plants.
- Spontaneous wild plants that represent a natural element in the local flora. These spontaneous wild plants can, at least theoretically, have a triple origin:

1. They may derive from sub-spontaneous wild plants that have found favorable conditions in the natural environment for a return to a wild state. We will refer to them as colonial wild plants.
2. They may descend from ancestors that have never passed through the cultivated stage. We will refer to them as autochthonous, or indigenous, wild plants.
3. Hybrid mixed wild plants. They may result from the hybridization of indigenous wild plants with either of the forms, and these are referred to as hybrid wild plants.

Let us represent the categories of wild plants’ diversity of mutually exclusive types using set theory notation (Table 1).

Table 1. Categories of wild plants and their relationships with domesticated cultivated ¹.

Basic Codes	Basic Categories	Notes
A	Post-cultivated wild plants	Remains of perennial crops in abandoned fields occupied by natural vegetation
B	Sub-spontaneous wild plants	Natural habitats suitable for the species and close to crop fields
C	Spontaneous wild plants	Natural habitats
D	Colonial wild plants (subset of spontaneous)	Natural habitats
E	Autochthonous or indigenous wild plants (subset of spontaneous)	Natural habitats
F	Hybrid mixed wild plants	Natural habitats
G ¹	Domesticated cultivated plants	Cropland
Coded relationships	Relationships between Categories ²	
$C = D \cup E$	Spontaneous wild plants are the union of colonial and autochthonous wild plants	
$D \cap E = \emptyset$	Colonial and autochthonous wild plants are mutually exclusive	
$A \cap B = \emptyset$	Post-cultivated and sub-spontaneous wild plants are mutually exclusive	
$A \cap C = \emptyset$	Post-cultivated and spontaneous wild plants are mutually exclusive	
$B \cap C = \emptyset$	Sub-spontaneous wild plants and spontaneous wild plants are mutually exclusive.	
$A \cup B \subseteq G^1$	All post-cultural and Sub-spontaneous wild plants are domesticated plants, despite their wild appearance.	
$F = (G^1 \cap E) \cup (E \cap D)$	Relationship of hybrid wild plants with other categories	

¹ To represent the fact that hybrid wild plants (F) are the result of specific crosses, we introduce an additional set, G, to denote the cultivated plants involved in the hybridization. The expression captures the idea that hybrid wild plants are the result of the intersection of cultivated and autochthonous wild plants or the intersection of autochthonous wild plants and post-cultivated, sub-spontaneous, and colonial wild plants or even pre-existing hybrids. This representation allows for a clear delineation of the contributing categories to the hybrid wild plants based on the specified crosses. ² These expressions capture the relationships between the distinct categories of lambrusques based on Levadoux’s [31] description expanded to all crop plants.

3.1.3. “Weeds” Versus “Invasive Species”

Weeds occupy a unique position between wild sub-spontaneous and wild post-cultivated plants, often thriving in cultivated fields (Table 2). In agriculture and gardening, weeds are typically defined as unwanted plants that grow in areas not intentionally cultivated. They compete with crops for resources such as sunlight, water, and nutrients, potentially harming growth and productivity. Weeds often spread rapidly and adapt to various conditions, with some having been previously cultivated.

Sub-spontaneous and colonial wild plants can become invasive species. Invasive plants, introduced either accidentally or intentionally, spread rapidly and displace native species, disrupting ecosystem balance (Table 2). These plants, which may have high reproductive rates and adaptability and lack natural predators or pathogens, initially thrive in human-altered habitats but can later invade natural areas, causing significant ecological damage [47–49].

Table 2. Comparing the concepts of “weed” and “invasive species” ¹.

Theme	Features	Weed ¹	Invasive ¹
Reproductive Strategy	Exhibit a high reproductive capacity (R-strategists), allowing them to establish and spread rapidly in various environments.	Yes	Yes
Ability to Compete	Can outcompete native or desirable plants for resources such as light, water, and nutrients.	Yes	Yes
Adaptability	Tendency to be highly adaptable to different environmental conditions, enabling them to thrive in a variety of habitats	Yes	Yes
Habitat Occupancy	Primarily associated with disturbed or cultivated areas, competing with crops or desirable plants in human-altered landscapes	Yes	Not
Ecological Impact	Often considered nuisances in agricultural settings, impacting crop yields and quality and sometimes damage infrastructure	Yes	Not
Human Perception	Generally perceived as unwanted plants in cultivated areas	Yes	Not
Management Approach	Control measures often include herbicide application, cultivation practices, or manual removal, including collection as food for humans or livestock	Yes	Not
Ecological Impact	Can have broader ecological impacts, leading to the decline or displacement of native species and disruption of ecosystem functions. They may alter habitat structures, nutrient cycles, and community compositions, affecting the overall biodiversity.	Not	Yes
Human Perception	These are also generally perceived as unwanted plants, but in natural environments. The preservation of native biodiversity is a fundamental goal of conservation biology. Invasive species can contribute to the decline of native species, making their management essential for maintaining healthy ecosystems	Not	Yes
Management Approach	Management may involve a more comprehensive ecological approach, considering the restoration of native habitats and the prevention of further spread.	Not	Yes

¹ It is important to note that the same species may be considered native, weedy, or invasive depending on the continent, country, or territory.

In summary, while weeds and invasive species share similar reproductive and competitive traits, differences in their habitat preferences, ecological impacts, and human perceptions influence their management and perception (Table 2). Evolutionarily, the distinctions between weeds and invasive species may lack inherent biological significance; these terms are anthropological constructs reflecting human attitudes toward plants in relation to cultivated crops, natural ecosystems, or indigenous species. The globalization of trade and travel has exacerbated the spread of invasive species, complicating the task of differentiating between natural ecological processes and human-mediated introductions.

3.1.4. Weeds as Novel Crops

Let us explore the potential advantages of using a weed, especially one that is widely gathered and consumed as food, for the development of a novel crop.

There are several key advantages to considering weeds as potential crops for human consumption:

First, weeds are highly adaptable to a range of environmental conditions, which makes them suitable for cultivation across diverse regions and climates. Their inherent adaptability can be harnessed to create resilient crop varieties capable of thriving under varying conditions.

Second, weeds are typically robust and resilient, thriving even in suboptimal environments. This resilience can be utilized to develop crops with improved tolerance to environmental stressors such as poor soil quality, pests, or extreme weather conditions.

Third, the drought-resistant qualities of certain weeds could be particularly valuable in developing crops that are better suited to regions experiencing water scarcity, contributing to more resilient agricultural systems in the face of climate change.

Another advantage lies in the local knowledge surrounding certain weeds that are already gathered and consumed by communities. This familiarity provides valuable insights into their edibility, nutritional benefits, and potential uses, facilitating the transition from wild plant to cultivated crop.

Weeds also possess significant genetic diversity, which has enabled them to flourish in a variety of ecosystems. This genetic richness presents opportunities to introduce beneficial traits into cultivated crops, such as enhanced pest resistance or the ability to thrive in specific soil types.

Nutritional value is another important factor. Many weeds consumed as food offer nutritional benefits that are currently underutilized. By identifying and enhancing these nutritional properties, we can develop crops that are not only resilient but also contribute to improved dietary health.

Cultural acceptance is a further consideration. The fact that certain weeds are already part of local diets indicates that they are culturally accepted, which can ease the process of introducing cultivated varieties into these communities.

Weeds are also known for their ability to grow with minimal human intervention. If these traits can be transferred to cultivated crops, it could lead to agricultural systems with reduced input requirements, including a lower dependence on fertilizers and pesticides, thereby promoting more sustainable farming practices.

In conclusion, leveraging the positive attributes of weeds, especially those already used as food, offers an opportunity to develop new crops with enhanced adaptability, resilience, nutritional benefits, and cultural acceptance. This approach aligns with the goals of promoting agrobiodiversity and sustainable agriculture by unlocking the potential of traditionally overlooked plant species.

While the potential advantages of using weeds for crop development are significant, several challenges and considerations must be carefully addressed:

One major challenge is the inherently competitive nature of weeds. Their success in natural environments often stems from their ability to outcompete other plants for resources such as nutrients, water, and light. Managing their growth in a cultivated setting to prevent them from overtaking other crops is essential.

Weeds may also exhibit unpredictable traits, including undesirable characteristics such as aggressiveness, invasiveness, or allelopathy—the release of chemicals that inhibit the growth of neighboring plants. These traits could pose significant challenges in a controlled agricultural environment.

Another consideration is the genetic heterogeneity of many weed species. This genetic variability, while contributing to their resilience, can make it difficult to develop uniform and stable crop varieties. The process of stabilizing desirable traits in such plants requires careful and often complex breeding efforts.

Weeds can also serve as hosts for pests and diseases, creating risks for nearby cultivated crops. The introduction of a weed species into an agricultural system may inadvertently increase the presence of these harmful agents, which could threaten overall crop health. Conversely, certain weed species found in proximity to cultivated crops may attract pests away from the crops, potentially preventing or mitigating pest establishment within the cultivated areas.

In addition, cultural perceptions of weeds vary significantly. While some communities have a history of consuming certain weed species, others may view them negatively, either due to their association with pests or because of local taboos. Overcoming these cultural barriers and fostering broader acceptance of weeds as crops could prove difficult.

Toxicity is another critical concern. Not all weed species are safe for human consumption, and some may even be poisonous. Rigorous evaluation is necessary to ensure that any selected weed is not only edible but also safe for widespread consumption.

The potential ecological impact of cultivating weeds also requires close scrutiny. If a particular weed species is invasive, introducing it into agricultural systems could result in unintended environmental consequences, including the displacement of native plant species and disruption of local ecosystems.

Crossbreeding between weeds and cultivated crops presents an additional risk. Such genetic mixing could compromise the purity of cultivated varieties, leading to unintended and potentially harmful consequences in terms of crop yield or quality. This does not apply to non-pollinated plants that reproduce asexually, nor to plants that produce fruits or seeds through mechanisms such as parthenocarpy or apomixis, which do not require pollination.

Effective weed management is essential in preventing these plants from becoming problematic in a controlled agricultural environment. Developing cultivation practices that allow for the growth of the desired crop while containing the weed's spread is a crucial aspect of successful integration.

Finally, legal and regulatory obstacles must be considered. Some weed species may be subject to regulation due to their classification as invasive or harmful, and their introduction into agricultural systems could face legal restrictions that must be carefully navigated.

In conclusion, while weeds offer several promising advantages for crop development, their integration into agricultural systems requires thoughtful management and mitigation of potential risks. A thorough evaluation of each species, along with its ecological, genetic, and cultural characteristics, is necessary to ensure responsible and successful utilization. In this context, the development of robust AI models trained on a wide range of variables will be crucial for assessing and managing these factors effectively.

3.1.5. Indigenous Wild Plants Versus Weeds as Potential Novel Crops

A comparison between indigenous wild food plants and long-established weeds as potential novel crops offers valuable insights, particularly when examining their utilization as food sources.

Indigenous wild plants contribute significantly to biodiversity by offering distinctive flavors and unique nutrients. However, their propagation poses challenges that may limit their availability for large-scale cultivation. While these plants are naturally adapted to local environmental conditions, transitioning them to cultivated fields or orchards may present further difficulties. Some species may resist domestication, complicating their integration into agricultural systems. Despite these challenges, indigenous wild plants are renowned for their high nutritional density. Yet, cultivating them under different conditions may impact their nutritional content, requiring a careful balance of cultivation practices to maintain these benefits.

In addition to their nutritional value, indigenous wild plants hold cultural significance, with local knowledge being key to their sustainable use. Cultivating these plants may necessitate collaborative efforts to preserve cultural heritage while adapting to agricultural settings. Their potential to promote environmental sustainability is also noteworthy. Although the challenges of cultivation remain, with careful management, these plants could contribute to ecosystem health and sustainability.

In contrast, weeds offer a different set of advantages. Their abundance and resilience make them readily available for both foraging and potential cultivation, enhancing food security in various regions. Weeds are characterized by rapid growth, which can be beneficial for quick crop establishment, though this same adaptability could pose challenges if their growth is left unmanaged. Many weeds also exhibit natural resistance to pests and diseases, an attribute that favors their use in cultivation. However, careful oversight is necessary to prevent these resilient plants from becoming ecologically disruptive or displacing native species.

Culinary potential is another consideration when evaluating weeds for cultivation. While weeds can introduce unconventional flavors that enrich culinary diversity, their taste profiles—often including bitterness or other strong flavors—require innovative culinary approaches to make them palatable for a wider audience. Despite their invasive

reputation, certain weeds exhibit considerable potential for cultivation due to their adaptability to a range of environments and minimal input requirements. This could make them viable candidates for low-maintenance agricultural systems, provided their spread is carefully managed.

Both wild food plants and weeds share several key considerations when evaluating their potential for crop development. A rigorous nutritional and phytochemical assessment is essential for both categories, ensuring that their health benefits are maintained throughout cultivation. Toxicity evaluation is equally crucial, as some species may contain harmful substances. Addressing these concerns is vital before promoting any wild or weedy species for wider consumption. Engaging local communities is another fundamental aspect of sustainable cultivation, as it fosters cooperation, respects traditional knowledge, and ensures alignment with local practices and needs.

Culinary creativity will also play a key role in integrating both indigenous wild plants and weeds into mainstream diets. By adapting traditional recipes to incorporate these species and taking advantage of their unique flavors, they can be more easily accepted by consumers.

In summary, while cultivating indigenous wild food plants presents challenges, including propagation difficulties and the need to preserve cultural heritage, collaborative efforts and adaptive cultivation strategies could facilitate their successful integration into agricultural systems. On the other hand, the abundance, resilience, and rapid growth of weeds make them valuable candidates for novel crops, provided their cultivation is managed responsibly to prevent ecological disruption. Interestingly, many of these weeds were once cultivated crops that have since been forgotten, which has led to their classification as “cryptocrops”—silent, secondary sources of food that coexist with traditional crops in agricultural fields [32].

3.2. Occurrences, Extension of Use, and Local Consensus for Selecting Wild Food Plant Species as Potential Crops

The Local Consensus Model for the Huerta de Murcia (Spain)

The estimation of the degree of consensus among informants regarding specific species within a locality or region has been proposed as a method to differentiate between more important and less relevant species. One of the most straightforward methods for estimating local consensus on plant usage is the Relative Frequency of Citation (RFC). This approach calculates the proportion of informants who report using a particular species, denoted as species X. The RFC for species X is determined by dividing the number of informants who mention its use by the total number of informants surveyed. Despite its simplicity, the RFC is a highly effective tool, offering a clear and accessible measure of consensus. It is particularly advantageous in overcoming the statistical complexities often associated with more intricate indices, providing a reliable alternative for capturing local knowledge and usage patterns [50–52].

This differentiation reflected by RFC values is based on subjacent criteria such as culinary but also medicinal uses, especially in plants that can be simultaneously consumed as spices and teas such as oregano, sage, or thyme.

Using the Relative Frequency of Citation (RFC) for each species, we aim to represent the distribution of these values within the study. Rather than observing a uniform distribution, a typical pattern emerges in which most species have low RFC values, while only a few species exhibit high RFC values. This distribution forms a gradient, highlighting that the majority of species are less frequently cited, with only a small number reaching the highest citation frequencies (Figure 3).

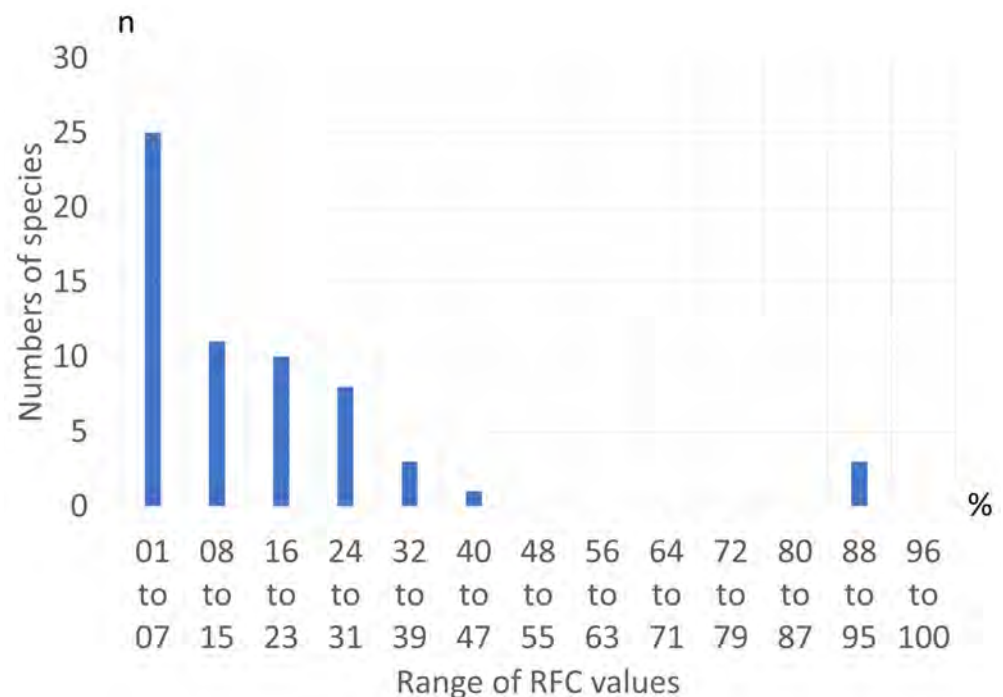


Figure 3. Distribution of Relative Frequency of Citation or percentage of informants citing each species in the Huerta de Murcia, expressed in terms of numbers of species. Graphics by Diego Rivera.

Of the 60 species reported, only three are cited by more than 50% of the informants. Notably, each of these three species achieves an RFC value exceeding 88. Interestingly, no species falls within the intermediate RFC range of 48 to 88, highlighting the stark contrast between the high consensus for these three species and the significantly lower values for the remaining species. This pattern mirrors the phenomenon of inequality of wealth distribution among human populations, often illustrated in economics through the Lorenz curve [33,53].

In all fields where the Lorenz curve is applied, the population is typically plotted along the x -axis, while wealth, abundance, or availability is represented on the y -axis [54,55]. As a starting point, the population or the variable on the x -axis is ordered from the lowest to the highest value of the variable plotted on the y -axis. In our case, we aim to relate species to their popularity or consensus in terms of citation frequency. To achieve this, we would generate a table with the number of species in one column and their corresponding RFC values in another, ordering the table from the lowest to the highest RFC values.

From this table, we calculate the cumulative pattern of species popularity or consensus, similar to how wealth distribution is represented across populations in economics. However, unlike economic models that assess the wealth of individuals, in this analysis, the species themselves constitute our population, and we evaluate the patterns of knowledge associated with each species within the studied group.

In the graphical representation, we distinguish between data related to the informants and the data on the number of records. The latter arises from the fact that a single plant may have multiple uses mentioned by the same informant. Consequently, the number of records for a given plant is always equal to or greater than the number of informants who mention it.

In this analysis, the variable X represents the cumulative percentage of species, ordered from the smallest to the largest percentage, while the variable Y corresponds to the cumulative percentage of consensus and popularity, measured through the cumulative RFC and the number of records associated with those species (Figure 4). Focusing on data from the Huerta de Murcia, our goal is to develop a model that effectively captures the relationship between these two variables.

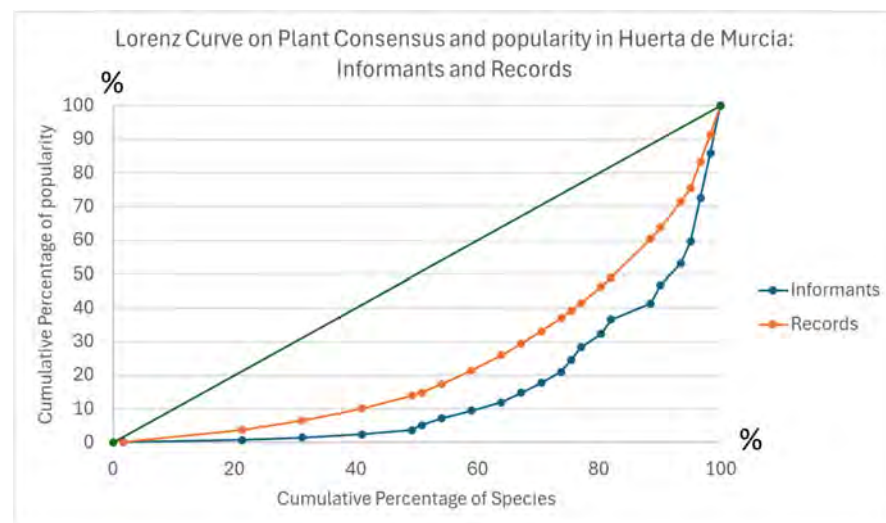


Figure 4. Relationships between the number of species and the number of informants or records in the Huerta de Murcia, expressed in terms of cumulative percentages. Graphics by Diego José Rivera-Obón.

In an ideal scenario of equal distribution, each cumulative percentage of species would match the corresponding cumulative percentage of records or informants, forming a 45-degree line from (0, 0) to (100, 100). However, the more the Lorenz curve deviates and bows below this line of equality, the greater the disparity in the distribution of records and informants across species (Figure 4).

Analyzing the model from a horticultural perspective, we can infer several important points about selecting species for cultivation development. The Lorenz curve suggests a significant inequality in the distribution of popularity among species. A high percentage of species account for a disproportionately small percentage of records and informants. Specifically, the first 50% of species account for only 4% of informants and 14% of the records. Conversely, at the other extreme, only two species cover 28% of the accumulated records, and four species cover 28% of the accumulated informants. This disparity highlights the need for a strategic approach in selecting candidate species for cultivation, focusing on those with higher popularity and broader acceptance (Figure 4).

However, the model overlooks cultural, sociological, and environmental factors, such as local traditions, culinary habits, socio-economic conditions, soil type, climate, and ecological interactions, which are crucial for accurately determining the feasibility of cultivating certain species.

The Lorenz curve model provides a useful tool for understanding the relationship between the number of informants and records and species consumption. It suggests focusing on species with optimal support values: *Sonchus tenerrimus* L. (weed), *Beta vulgaris* L. subsp. *maritima* (L.) Arcang. [= *Beta maritima* L.] (weed and autochthonous wild), *Sonchus oleraceus* L. (weed), *Cynara cardunculus* L. subsp. *cardunculus* [incl. *C. scolymus* L.] (sub-spontaneous), *Oxalis pes-caprae* L. (weed and invasive), *Foeniculum vulgare* Mill. (autochthonous wild and occasionally cultivated), *Beta vulgaris* L. subsp. *vulgaris* (sub-spontaneous or feral, intermediates with *Beta vulgaris* L. subsp. *maritima* (L.) Arcang. (*B. maritima* L.) are frequently met with), *Helminthotheca echioides* (L.) Holub [= *Picris echioides* L.] (weed), and *Crepis vesicaria* L. ssp. *taraxacifolia* (Thuill.) Thell. (weed).

3.3. Analysis at Different Scales: A Comparison of Murcia and World Data

3.3.1. The Global Occurrence Patterns of Potential Crops

The global pattern of wild species consumed as food helps identify widely consumed species suitable for cultivation and those locally adapted to unique conditions that may be valuable for smaller areas.

Starting from the Global Database of Wild Food Plants, we analyzed data from various scales: Huerta de Murcia (61 species and 656 records). These 61 wild plant species are commonly found in Huerta Murcia and were selected based on their mention during interviews. The remaining species were identified through the literature review described previously. Spain (482 species and 3582 records), the Mediterranean (773 species and 7966 records), and globally (14,813 species and 42,555 records) (Figure 5). We used cumulative percentages of species and localities or countries where the plant is consumed to conduct our analysis.

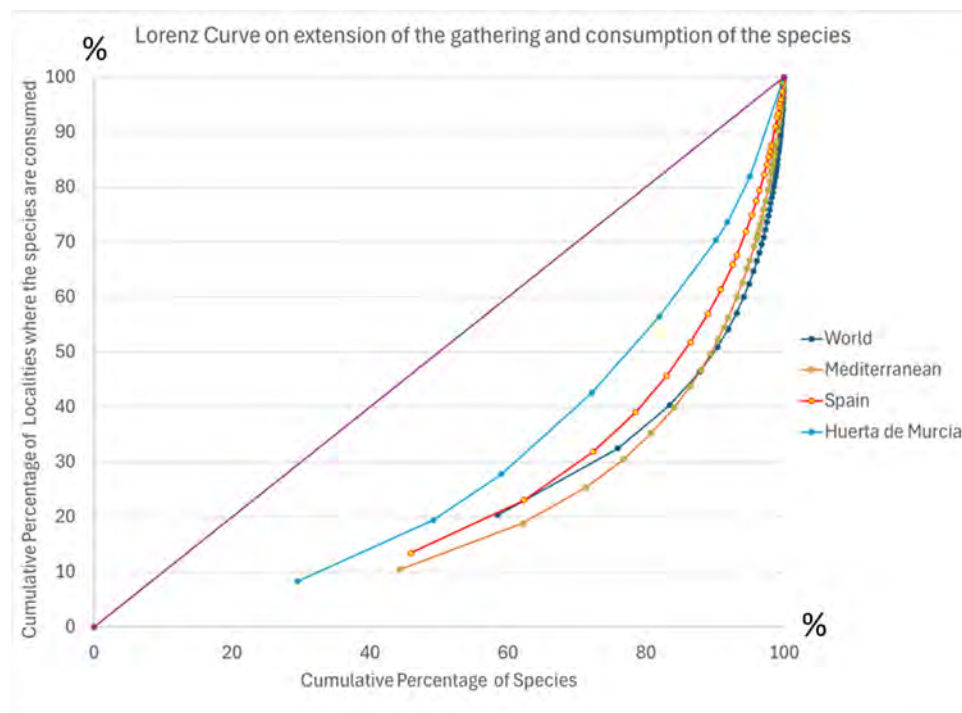


Figure 5. Relationships between species and localities registered in the Global Database of Wild Food Plants at different geographical levels (Huerta de Murcia, Spain, Mediterranean, World). Graphics by Diego José Rivera-Obón and Diego Rivera.

Species used exclusively in a single locality mark the lower limit of each curve (Figure 5) and account for 25% of the species in Huerta de Murcia and are found in only 8% of its localities. In contrast, at a global scale, 58.5% of the species are consumed in only one locality, with such cases occurring in just 20% of the localities. In Spain and the Mediterranean region, the values are approximately 45%, with only 10% of the localities presenting exclusive species (Figure 5). This simply indicates that most food plants are extremely local. Specifically, at a global scale, 58% (8662 taxa) were consumed in a single locality, 17% (2580 taxa) in two localities, and 7.5% (1115 taxa) in three localities. In contrast, only 5 taxa among the 14,813 recorded, namely, *Portulaca oleracea* L., *Taraxacum officinale* Wigg. [= *Leontodon taraxacum* L.], *Nasturtium officinale* R.Br. [= *Rorippa nasturtium-aquaticum* (L.) Hayek], *Chenopodium album* L., and *Sonchus oleraceus* L., were consumed as food in over one hundred localities. This provides vital information for selecting candidate species, although the distribution of the species per se and the distribution of its use as food do not always coincide.

The practical implications of this simple model allow us to differentiate Generalist Species, which are consumed in many localities (>40 localities in our model) and are likely generalists, adaptable to various environments and culturally accepted across different regions, from Specialist Species, which are consumed in few localities (<5 localities in our model) and likely have specific ecological niches or cultural significances, making them less widespread but potentially more important locally.

Several Points of Incertitude must be considered for this global model:

The model does not incorporate environmental factors such as soil type, climate, and ecological interactions, which are crucial for understanding species distribution and suitability for cultivation. Cultural preferences and traditional uses are not explicitly considered in the model. These can significantly impact which species are collected and consumed and may lead to over or underestimation of certain species' importance. The quality and consistency of data on species collection and consumption may vary. Any biases or inaccuracies in the data collection process can introduce uncertainty into the model. The model's applicability might be limited by the geographical scope of the data. Thus, the aim of reaching the widest possible coverage (Figure 2).

Our analysis reveals that very few species are widely consumed, while most are limited to specific localities. The model's limitations and the uncertainties related to ecological, sociological, and data quality factors must be carefully considered. Further research and field validation are essential to refine the model and ensure its practical applicability in agronomic decision-making.

3.3.2. Integrating Diverse Sources of Evidence

Comparing the occurrences of plant species in a general distribution database like GBIF [56] with the localities where they are consumed as food, as recorded in our Global Database of Wild Food Plants, can help discern cultural, sociological, and other factors that limit their consumption. We present an example of the wild food plants in the Huerta de Murcia. Figure 6 compares the number of global occurrences of these 61 plant species as food plants with their overall geographical distribution based on GBIF [56] data. This analysis highlights species that are widely recognized as food plants and those with a broad geographical distribution, providing insights into their cultural and practical importance, as well as their potential availability.

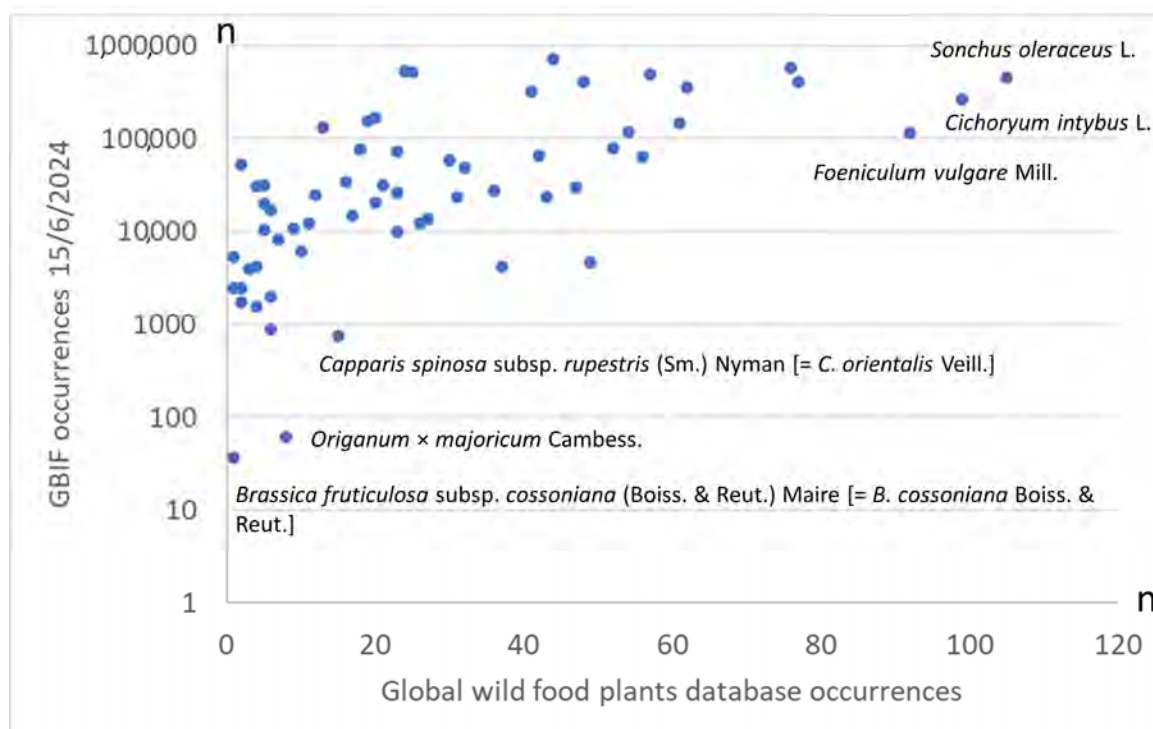


Figure 6. Relationships between global occurrences of wild food plants from Huerta de Murcia recorded in GBIF and global occurrences of their use as food. Graphics by Diego Rivera.

Although acknowledging that the GBIF database frequently suffers from data gaps across various countries and that such data may not qualify as “primary data”, it remains valuable for modeling purposes. It is important to account for limitations such as missing

information and inaccuracies in georeferencing. For example, some terrestrial locations may be incorrectly placed in the middle of the ocean, while others may correspond to atypical settings like greenhouses in botanical gardens. Recognizing these issues helps in effectively integrating GBIF data into the model while considering its constraints.

The analysis in Figure 6 reveals patterns of recognition and disparity between global occurrences as food plants and their overall geographical distribution. Species with high recognition as food plants and broad geographical distribution (*Sonchus oleraceus* and *Cichorium intybus* L.) are likely to be universally recognized and utilized. In contrast, species with lower recognition as food plants but extensive geographical distribution may have primary uses other than food or may not be used at all (*Papaver dubium* L.). Other categories are widely cultivated plants that are rarely found growing 'wild' (*Beta vulgaris* L. subsp. *vulgaris*, *Vicia sativa* L. and *Cynara cardunculus* L. subsp. *cardunculus* [incl. *C. scolymus* L.]). Species endemic to the Mediterranean or smaller areas within this region have lower values of both (*Capparis spinosa* subsp. *rupestris* (Sm.) Nyman [= *Capparis orientalis* Veill.], *Origanum* × *majoricum* Cambess., *Brassica fruticulosa* subsp. *cossoniana* (Boiss. & Reut.) Maire [= *Brassica cossoniana* Boiss. & Reut.]).

The integration of local support or consensus data for the wild species used in the Huerta de Murcia with the occurrences recorded in the Global Database of Wild Food Plants reveals peculiar patterns (Figure 7). Species such as *Sonchus oleraceus* and *Beta vulgaris* subsp. *maritima* exhibit high local consensus and high global occurrences, indicating their widespread recognition and use. Similarly, *Cichorium intybus* shows high global occurrence but medium local consensus, suggesting it is more widely recognized outside the specific local context.

On the other hand, some species present discrepancies. For instance, licorice, *Glycyrrhiza glabra*, has high global occurrences but low local consensus, indicating it may be more important or recognized globally than locally. Considering that licorice was widely cultivated in the area for centuries and persists locally as a weed, the fact that it has a low level of consensus as a food is because most of the informants consider it to be a medicinal product and not a food item. Conversely, species like *Sonchus tenerrimus* display high local consensus but low global occurrences, signifying their local significance but lesser global use.

This analysis aids in understanding the disparity between local and global importance.

3.4. Developing a General Model to Integrate Multimodal Data

3.4.1. Preliminary Considerations

Given the limitations, discrepancies, and uncertainties of the previous partial models, it is necessary to develop a multimodal model capable of integrating data from a wide range of sources. However, generalist artificial intelligence is not able to consider certain issues that a researcher may notice. For this reason, we recommend moving away from the use of generalist AI tools and instead advocate for the development of a specialized tool. This tool should be built on the collaborative efforts of a diverse team of researchers, ensuring that it can account for a broad spectrum of environmental, biological, and social factors.

We must address critical questions regarding the development of a global multimodal-based AI, despite the significant geographical, historical, and cultural differences between continents, countries, or regions. Is it not more practical and effective to design models tailored to specific regions, countries, or continents? For example, creating models for major bioclimatic regions such as tropical, temperate, or Mediterranean might be more feasible. However, a global model remains valuable, as it can incorporate diverse constraints and conditions. Additionally, a broader scope for cultivation, especially within the context of a rapidly changing climate, is essential for the development of novel crops.

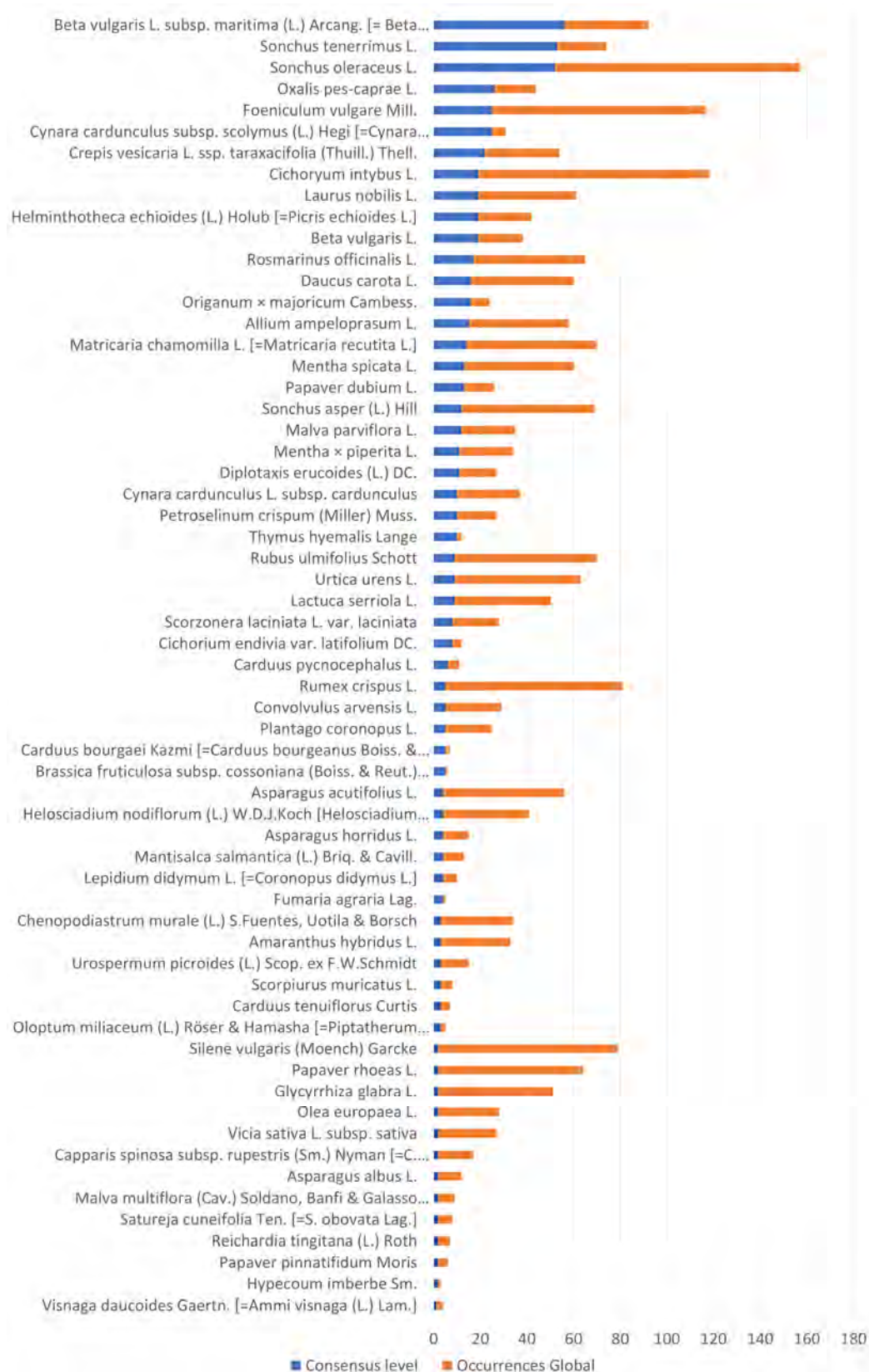


Figure 7. Comparison of local consensus levels and global occurrences of wild food plants from Huerta de Murcia. The graph illustrates the correlation between the consensus level in terms of the number of informants for wild food plants in Huerta de Murcia and their global occurrences as food species in terms of the number of zones. Graphics by Diego Rivera.

We acknowledge, however, that this is an ambitious undertaking requiring substantial resources and research personnel. Additionally, much of the necessary information is fragmented across various databases, some of which are not easily accessible. As a result, this remains, for now, a conceptual proposal.

3.4.2. Tools and Techniques Available

To develop a specific AI for selecting candidate novel crops from a large list of wild food plants, several relevant tools and technologies should be employed.

- **Machine Learning Algorithms [57]:** Supervised learning is useful for predicting outcomes based on labeled training data. Common algorithms include decision trees, random forests, support vector machines, and neural networks. Unsupervised learning helps in clustering and finding patterns in data without pre-existing labels. Algorithms like k-means clustering and hierarchical clustering can be used to group plants with similar traits.
- **Natural Language Processing (NLP) [58]:** Text mining and data extraction: Tools like NLTK [59] (it's an open-source project developed and maintained by a community of contributors), spaCy [60] (it is an open-source project developed and maintained by a community of contributors), and BERT [61] (Google AI. While Google AI, Alphabet Inc., Mountain View, CA, USA) can extract relevant information about plant characteristics, nutritional value, health benefits, and cultivation requirements from the scientific literature, databases, and online resources. Sentiment analysis and entity recognition tools can be used to understand and categorize qualitative data from various sources.
- **Geographic Information Systems (GISs):** Spatial analysis, tools like ArcGIS (Esri, Redlands, CA, USA) and QGIS (it is developed by a global community of volunteers and is not affiliated with any specific company or country) can help analyze the geographic distribution of wild plants and their environmental conditions, providing insights into optimal growing regions. Remote sensing, satellites, and drones can gather data on land use, soil types, and climate conditions to support decision-making.
- **Genomic and Bioinformatics Tools:** Gene sequencing data analysis [62]: Tools like BLAST [63] (NCBI, Bethesda, MD, USA), ClustalW (its development and maintenance have been primarily carried out at the University of Cambridge in the United Kingdom), and Bioconductor (initially developed at the Fred Hutchinson Cancer Research Center in Seattle, WA, USA, it has since become a collaborative effort) can analyze genetic information to identify plants with desirable traits for breeding and cultivation. Genome-wide association studies (GWAS) (is not associated with a specific company or country) [64] are used to find correlations between genetic variants and traits of interest in plants.
- **Big Data Platforms:** Data storage and processing in platforms like Hadoop (the Apache Software Foundation, which oversees Hadoop's and Spark's development and governance, is a global collaboration of individuals and organizations), Spark, and Google Cloud (Google LLC, Mountain View, CA, USA) to manage and process large datasets efficiently. Data integration [65] with tools like KNIME (it's a collaborative project involving researchers and developers from various institutions worldwide. The KNIME AG company, based in Switzerland, provides commercial support and additional features for enterprise users) and Alteryx (Alteryx, Inc. Irvine, CA, USA) can integrate data from multiple sources, including genomic, phenotypic, and environmental data.
- **Decision Support Systems (DSSs) [66]:** multi-criteria decision analysis (MCDA): Tools like DEXi (Dexi.io San Francisco, CA, USA) and PROMETHEE (it is not associated with a specific company or country) can evaluate multiple criteria to rank and select candidate species based on their suitability for cultivation. Expert systems: AI-driven systems that use rule-based approaches to simulate expert-level decision-making in selecting crop candidates.

- **AI and Machine Learning Platforms:** TensorFlow (Google Brain, a research division of Google AI, which is part of Alphabet Inc., the parent company of Google, Mountain View, CA, USA) [67] and PyTorch (Facebook AI Research (FAIR), Menlo Park, CA, USA) [68]: widely used frameworks for developing and deploying machine learning models. AutoML Tools [69,70]: tools like Google AutoML (Google Cloud, Alphabet Inc., Mountain View, CA, USA) and H2O.ai (H2O.ai, Mountain View, California, Estados Unidos) [71] automate the process of model selection, hyperparameter tuning, and feature engineering.
- **Data Visualization Tools [72]:** Visualization libraries: libraries like Matplotlib (it is not associated with a specific company or country) [73], Seaborn (it is an open-source project developed and maintained by a community of contributors) [74], and Plotly (Plotly, Inc., Montreal, QC, Canada) [75] in Python (it was created by Guido van Rossum in the late 1980s and released to the public in 1991) can create insightful visualizations to interpret complex data. Interactive dashboards: tools like Tableau (Salesforce, Seattle, WA, USA) and Power BI (Microsoft Power Platform, Microsoft Corporation, Redmond, WA, USA) [76] can create interactive dashboards to help stakeholders understand the data and results.
- **Collaborative Platforms and Research Collaboration Tools [77]:** Platforms like GitHub (GitHub, San Francisco, CA, USA. It was acquired by Microsoft in 2018) [78] and Jupyter Notebooks (Jupyter Notebooks don't have a specific physical location) [79] facilitate collaboration among researchers by enabling code sharing and collaborative data analysis.

3.4.3. Critical Points

Considering the repertory of available tools, developing a mathematical model to integrate multimodal data and leveraging AI to evaluate the potential of various wild plant species, whether autochthonous, weeds, or invasive, as promising crops for future food production involves several critical points and key steps.

- **Multimodal Data Integration:** The model must build a coherent data system integrating ecological (environmental conditions where the species thrive, such as soil type, climate, and geographic distribution), genetic (genomic information to understand the genetic diversity, adaptation mechanisms, and potential for breeding), agronomic (growth rates, yield potential, resistance to pests and diseases, and watering requirements) [80], socioeconomic (local and global market demand, cultural importance, and economic viability), ethnobotanical (historical and traditional uses of the species, including medicinal, nutritional, and cultural significance), nutritional, toxicological, and pharmacological data (active substances present, their relative abundance and their physiological and pharmacological properties in humans and livestock) [81].
- **AI Techniques:** Machine learning supervised and unsupervised learning algorithms to identify patterns and make predictions based on integrated data. Deep learning, neural networks for complex data analysis, such as image recognition of plant species or genomic data interpretation [82]. Natural language processing (NLP) [83] to analyze textual data from the scientific literature, reports, and ethnobotanical records. Geospatial analysis, GIS, and remote sensing data to map species distribution and environmental conditions.
- **Model Development:** Feature selection, identifying the most relevant variables from the multimodal data that influence the potential of species as crops. Model training using historical data to train AI models to predict the suitability and potential productivity of plant species. Cross-validation and testing with independent datasets to ensure model accuracy and reliability.

3.4.4. Key Steps

- **Data collection and preprocessing:** Gather data from various sources, including ecological databases, genetic repositories, agronomic trials, and socioeconomic surveys.

Clean and preprocess the data to manage missing values, normalize scales, and encode categorical variables.

- **Feature Engineering:** derive new features that capture interactions between diverse types of data, such as genotype–environment interactions or socio-economic factors influencing plant use. Use domain knowledge to guide feature selection and ensure meaningful variables are included.
 - **Model Selection:** Start with basic machine learning models (e.g., decision trees [84], random forests [85]) to establish baselines. Progress to more complex models (e.g., neural networks [86], ensemble methods [87]) for improved accuracy and robustness. Implement geospatial models to incorporate environmental data effectively.
 - **Training and Validation:** Split the data into training and validation sets. Train the model on the training set, using cross-validation techniques to fine-tune hyperparameters. Validate the model on the validation set, iterating to improve performance.
1. **Interpretability and Explainability:** Use techniques like SHAP values [88] or LIME [89] to interpret the model’s predictions and understand the contribution of distinctive features. Ensure the model is transparent and its predictions are explainable to stakeholders.
 2. **Deployment, Monitoring, and Updating:** Deployment involves implementing the AI model in a practical and accessible format, such as a software application or online platform so that stakeholders—such as researchers, practitioners, or decision-makers—can easily interact with and utilize the model. The goal is to make the AI tool user-friendly, ensuring that it meets the needs of its intended audience and integrates smoothly into their workflows. Monitoring involves evaluating its accuracy, efficiency, and overall effectiveness as it processes new data. Continuous monitoring helps to identify any issues or discrepancies that may arise and ensures that the model remains reliable and relevant. Updating involves regularly refining the model based on performance evaluations and new information to maintain its relevance and effectiveness.
 3. **Case Studies and Pilot Projects:** Conduct case studies or pilot projects to validate the model using real-world scenarios and gather feedback from agronomists, farmers, and other stakeholders to refine the model and its application.

In summary, developing a mathematical model to integrate multimodal data and using AI to evaluate the potential of wild plant species as crops involves collecting diverse data, using advanced AI techniques for analysis, and ensuring model interpretability and accuracy. By following the outlined steps and leveraging sophisticated AI models, it is possible to identify promising species for future food production, considering ecological, genetic, agronomic, and socioeconomic factors.

However, one of the primary challenges in developing a specific AI lies in accessing the necessary data. Some, like geographical distribution, can be obtained online through platforms such as GBIF [56], though with limitations. Standardizing scientific names and including exhaustive lists of synonyms is crucial for data search tools like web scraping, and this nomenclature can be accessed through resources such as POWO [90].

However, a significant portion of the required information must be gathered through extensive bibliographic reviews. The locations where species are collected and consumed can be identified by reviewing thousands of sources (articles, books, and other references). For example, Google Scholar [91] lists 14,400 studies on “wild edible plants” (702 in Web of Science [92] and 215 in PubMed [93]), 8660 on “wild food plants” (427 and 117, respectively), and only 442 on “gathered food plants” (15 and 3). Unfortunately, only a fraction of these studies include data on local consensus.

Key factors limiting the plant’s development in its bioclimatic origin, such as temperature range, precipitation, and atmospheric humidity, can be derived from specific studies or inferred through models based on geographic distribution. However, these data generally pertain to wild populations, and cultivated varieties may respond differently, necessitating the modeling of potential cultivation areas. Edaphic factors, which may be as simple as

the distinction between calcareous and siliceous soils, extend to more complex aspects of chemical, physical, and biological soil diversity. Geobotanical, phytosociological, and ecophysiological studies provide crucial insights into the conditions under which plants develop, though their geographic coverage is inconsistent, leaving some species with little available information.

The toxicity of plants under study can be assessed through numerous publications or databases such as Toxic Plants-Phyto Toxins (TPPT) by searching species' scientific names and synonyms alongside terms like "toxic". There are 21,900 records on "toxic plants" in Google Scholar [91], 16,799 in PubMed [93], and 6649 in Web of Science [92].

The nutritional value of wild plants is known for a limited number of species, often within the context of phytochemical studies, which constrains the availability of these data. A total of 794,000 "Phytochemistry" papers are registered in Google Scholar [91], 19,943 in PubMed [93], and 9175 in Web of Science [92].

The primary limitation of previously published works on the cultivation of wild edible plants is their lack of an explicit, homogeneous selection criterion, and they often remain at the proposal stage without providing experimental data [94–96]. In some cases, the proposals appear promising, but the cited data sources are included merely for the sake of inventory, without specifying their concrete utility [97]. In contrast, there have been successful reports of the domestication and cultivation of wild edible species, such as *Silene vulgaris* [98,99], *Scolymus hispanicus* [100,101], or *Asparagus acutifolius* [102]. Particularly promising is the growing interest in wild edible species that are wild relatives of domesticated and cultivated plants [103,104], which could even lead to the development of new organisms through hybridization between the two groups [105].

A sensible approach would be to structure the research in phases: conduct an initial screening based on elementary data such as geographic distribution and regions where the species are collected; employ existing models to assess potential cultivation areas; and for the primary set of selected species, pursue more detailed research into nutritional value, toxicity, and other agronomically relevant factors such as resistance to pathogens. Finally, an experimental phase could be undertaken with the limited set of the most promising species identified through these previous filters.

4. Conclusions

The terms "weed" and "invasive species" are not purely scientific classifications but rather reflect human judgments and perceptions shaped by the ways these plants interact with cultivated crops, natural ecosystems, or indigenous species. These labels often carry subjective connotations, as they are rooted in how people view the plant's impact—whether it competes with valuable agricultural species, disrupts ecosystem balance, or threatens native biodiversity. As such, a plant deemed "invasive" or a "weed" in one context may be viewed differently in another, depending on geographic, cultural, ecological, and economic considerations.

Distinguishing natural ecological responses from human-mediated introductions is challenging. However, leveraging the positive attributes of weeds, particularly those consumed as food, presents an opportunity to develop novel crops with enhanced adaptability, resilience, nutritional value, and cultural acceptance. This approach supports agrobiodiversity and sustainable agriculture by utilizing traditionally overlooked plants. While beneficial, using weeds in crop development necessitates careful management to address potential challenges and risks.

The Lorenz curve model provides insights into the inequalities in the relationships between consumers, records, and species at local and various geographical levels. Our analysis reveals that only a few species are widely consumed, while most are confined to specific localities and even there, are consumed by a few individuals. Globally, most food plants are extremely localized, with only 5 taxa among 14,813 recorded being consumed in over 100 localities. This model distinguishes between Generalist Species, consumed in many localities (more than 40 in our global model), and Specialist Species, consumed in

few localities (fewer than 5 in our global model). Generalist Species are adaptable and culturally accepted across regions, whereas Specialist Species have specific ecological niches or cultural significance. These limitations and uncertainties must be considered, and further research is needed to refine the model for agronomic decision-making.

The abundance and resilience of weeds position them as promising candidates for the development of novel crops, provided they are managed with careful attention to their potential impacts. Due to their ability to thrive in a wide range of environmental conditions and resist pests, weeds often require fewer resources to cultivate, making them attractive for sustainable agriculture. Furthermore, many weeds were historically cultivated as crop species, but over time, they became abandoned or forgotten as agricultural practices evolved and preferences shifted. Revisiting these species offers the opportunity to tap into a diverse genetic pool that could contribute to food security and agricultural innovation, particularly in the face of climate change and increasing global demand for resilient crops. However, careful management and research are essential to prevent the risks associated with their invasive nature and ensure they can be re-integrated into agricultural systems without negative ecological consequences.

The local consensus model based on RFC, as well as global distribution models, fail to incorporate crucial environmental factors such as soil type, climate, and ecological interactions, which are key to understanding species distribution and cultivation potential. Cultural preferences and traditional uses, which significantly influence species collection and consumption, are also not explicitly factored in, potentially leading to over- or underestimation of certain species' significance. The model's limitations, including uncertainties related to ecological, sociological, and data quality factors, must be addressed. Further research and extensive field validation are necessary to refine these models for accurate agronomic decision-making, highlighting the need for more complex multimodal approaches. Therefore, we propose the development of a specific artificial intelligence assistant that can integrate numerous factors: ecological, nutritional, toxicological, agronomic, biogeographical, ethnobotanical, economic, physiological, and other characteristics that define a species' potential as a new crop. We outline the appropriate tools and propose the steps for developing this AI assistant in the future.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/horticulturae10101021/s1>, Table S1: references analyzed for the Global Database of Wild Food Plants. Book SB1: basic data used for graphics in the figures.

Author Contributions: Conceptualization, D.R. and D.-J.R.-O.; methodology, D.R. and D.-J.R.-O.; software, D.-J.R.-O.; validation, C.O. and J.-A.P.; formal analysis, C.O.; investigation, D.R. and C.O.; resources, C.O.; data curation, D.R.; writing—original draft preparation, D.R.; writing—review and editing, C.O.; visualization, D.-J.R.-O.; supervision, D.R.; project administration, D.R.; funding acquisition, D.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Ministerio de Educación y Universidades (Spain), Estudio de indicadores de procesos de coevolución entre las poblaciones humanas y las especies vegetales recolectadas como alimento (gfps) en el Mediterráneo, CGL2008-04635/BOS.

Data Availability Statement: The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Acknowledgments: We gratefully acknowledge the contributions of Alonso Verde, Arturo Valdés, and José Fajardo for their assistance in Castilla La Mancha. We also extend our gratitude to Carmen Nicolás, Francisco Méndez, and Francisco Alcaraz for their fieldwork in Murcia. Furthermore, we thank Latifa Attieh for her support in Lebanon.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Smith, L.C.; El Obeid, A.E.; Jensen, H.H. The geography and causes of food insecurity in developing countries. *Agric. Econ.* **2000**, *22*, 199–215. [\[CrossRef\]](#)
- Minnis, P.E. Famine Foods of the North American Desert Borderlands in Historical Context. In *Ethnobotany a Reader*; Minnis, P.E., Ed.; University of Oklahoma Press: Norman, OK, USA, 2000; pp. 214–239.
- Begossi, A. Food Taboos—A Scientific Reason? In *Plants for Food and Medicine*; Prendergast, H.D.V., Etkin, N., Harris, D., Houghton, P., Eds.; Royal Botanic Gardens: Kew, UK, 1998; pp. 41–46.
- Winstead, D.J.; Jacobson, M.G. Food resilience in a dark catastrophe: A new way of looking at tropical wild edible plants. *Ambio* **2022**, *51*, 1949–1962. [\[CrossRef\]](#) [\[PubMed\]](#)
- Handa, S.S. The Integration of food and medicine in India. In *Plants for Food and Medicine*; Prendergast, H.D.V., Etkin, N., Harris, D., Houghton, P., Eds.; Royal Botanic Gardens: Kew, UK, 1998; pp. 57–68.
- Sinha, R.; Sinha, S. *Ethnobiology*; Surhaby Publications: Jaipur, India, 2001; pp. 1–335.
- Ding, X.Y.; Zhang, Y.; Wang, L.; Zhuang, H.F.; Chen, W.Y.; Wang, Y.H. Collection calendar: The diversity and local knowledge of wild edible plants used by Chenthang Sherpa people to treat seasonal food shortages in Tibet, China. *J. Ethnobiol. Ethnomed.* **2021**, *17*, 40. [\[CrossRef\]](#)
- Gomes, L.C.A.; Medeiros, P.M.; Prata, A. Patterns of use of wild food plants by Brazilian local communities: Systematic review and meta-analysis. *J. Ethnobiol. Ethnomed.* **2023**, *19*, 47. [\[CrossRef\]](#) [\[PubMed\]](#)
- Guarrera, P.M.; Savo, V. Wild food plants used in traditional vegetable mixtures in Italy. *J. Ethnopharmacol.* **2016**, *185*, 202–234. [\[CrossRef\]](#)
- Motti, R. Wild Plants Used as Herbs and Spices in Italy: An Ethnobotanical Review. *Plants* **2021**, *10*, 563. [\[CrossRef\]](#)
- Fongnzossie, E.F.; Nyangono, C.F.B.; Biwole, A.B.; Ebai, P.N.B.; Ndifongwa, N.B.; Motove, J.; Dibong, S.D. Wild edible plants and mushrooms of the Bamenda Highlands in Cameroon: Ethnobotanical assessment and potentials for enhancing food security. *J. Ethnobiol. Ethnomed.* **2020**, *16*, 12. [\[CrossRef\]](#)
- Zohary, M. *Geobotanical Foundations of the Middle East*; Gustav Fisher Verlag: Stuttgart, Germany, 1973; Volumes 1 and 2.
- Ertuğ, F. An ethnobotanical study in Central Anatolia (Turkey). *Econ. Bot.* **2000**, *54*, 155–182. Available online: <https://www.jstor.org/stable/4256288> (accessed on 14 September 2024).
- Pyke, G.H.; Starr, C.K. Optimal foraging theory. In *Encyclopedia of Social Insects*; Starr, C.K., Ed.; Springer International Publishing: Berlin/Heidelberg, Germany, 2021; pp. 677–685. [\[CrossRef\]](#)
- O’Connell, J.F.; Hawkes, K. Alyawara plant use and optimal foraging theory. In *Hunter-Gatherer Foraging Strategies: Ethnographic and Archaeological Analyses*; Winterhalder, B., Smith, E., Eds.; University of Chicago Press: Chicago, IL, USA, 1981; pp. 99–125. Available online: <https://archive.org/details/huntergathererfo0000unse/mode/2up> (accessed on 15 July 2024).
- Pyke, G.H. Optimal foraging theory: A critical review. *Annu. Rev. Ecol. Syst.* **1984**, *15*, 523–575. [\[CrossRef\]](#)
- Gomes, L.; de Medeiros, P.; Prata, A. Wild food plants of Brazil: A theoretical approach to non-random selection. *J. Ethnobiol. Ethnomed.* **2023**, *19*, 28. [\[CrossRef\]](#)
- Haq, S.M.; Khoja, A.A.; Waheed, M.; Siddiqui, M.H.; Alamri, S.; Alfagham, A.T.; Al-Humaid, L.; Bussmann, R.W. Food ethnobotany of forest resource in the high-altitude Himalaya Mountains: Enhancing the food sovereignty of ethnic groups. *For. Policy Econ.* **2024**, *164*, 103247. [\[CrossRef\]](#)
- Ceccanti, C.; Landi, M.; Benvenuti, S.; Pardossi, A.; Guidi, L. Mediterranean Wild Edible Plants: Weeds or “New Functional Crops”? *Molecules* **2018**, *23*, 2299. [\[CrossRef\]](#) [\[PubMed\]](#)
- Guarrera, P.M.; Savo, V. Perceived health properties of wild and cultivated food plants in local and popular traditions of Italy: A review. *J. Ethnopharmacol.* **2013**, *146*, 659–680. [\[CrossRef\]](#) [\[PubMed\]](#)
- Abbet, C.; Mayor, R.; Roguet, D.; Spichiger, R.; Hamburger, M.; Potterat, O. Ethnobotanical survey on wild alpine food plants in Lower and Central Valais (Switzerland). *J. Ethnopharmacol.* **2014**, *151*, 624–634. [\[CrossRef\]](#) [\[PubMed\]](#)
- Rivera, D.; Obón, C.; Inocencio, C.; Heinrich, M.; Verde, A.; Fajardo, J.; Palazón, J.A. Gathered Food Plants in the Mountains of Castilla-La Mancha (Spain): Ethnobotany and Multivariate Analysis. *Econ. Bot.* **2007**, *61*, 269–289. [\[CrossRef\]](#)
- Obón, C.; Martínez, R.; Giner, J.F.; Rivera, D. Las plantas comestibles recolectadas en la provincia de Alicante, estudio comparativo entre la Marina Alta y el Bajo Segura. In *Salut, Alimentació I Cultura Popular al País Valencià*; Guillem, X., García, G., Eds.; CEIC Alfons el Vell: Gandía, Spain, 2009; pp. 279–294.
- Rivera, D.; Verde, A.; Fajardo, J.; Inocencio, C.; Obón, C.; Heinrich, M. *Guía Etnobotánica de los Alimentos Locales Recolectados en la Provincia de Albacete*; Instituto de Estudios Albacetenses “Don Juan Manuel”: Albacete, Spain, 2006; pp. 1–470.
- Rivera, D.; Cobon Heinrich, M.; Inocencio, C.; Verde, A.; Fajardo, J. Gathered Mediterranean Food Plants—Ethnobotanical Investigations and Historical Development. In *Local Mediterranean Food Plants and Nutraceuticals*; Heinrich, M., Muller, W., Galli, C., Eds.; Karger: Basel, Switzerland, 2006; pp. 18–74.
- Rivera, D.; Alcaraz, F.; Obón, C. Wild and Cultivated Plants Used as Food and Medicine by the Cimbrian Ethnic Minority in the Alps. *Acta Hort.* **2012**, *955*, 31–39. [\[CrossRef\]](#)
- Obón, C.; Rivera, D.; Alcaraz, F. Wild and Cultivated Plants Used as Food and Medicine by the Mòcheni Ethnic Minority in the Alps. *Acta Hort.* **2012**, *955*, 113–118. [\[CrossRef\]](#)
- Leonti, M.; Nebel, S.; Rivera, D.; Heinrich, M. Wild Gathered Food Plants in the European Mediterranean: A Comparative Analysis. *Econ. Bot.* **2006**, *60*, 130–142. [\[CrossRef\]](#)

29. Baldi, A.; Bruschi, P.; Campeggi, S.; Egea, T.; Rivera, D.; Obón, C.; Lenzi, A. The Renaissance of Wild Food Plants: Insights from Tuscany (Italy). *Foods* **2022**, *11*, 300. [\[CrossRef\]](#)
30. Obón, C.; Nicolás, C.; Rivera, D. Estudio de las plantas comestibles silvestres del municipio de Murcia. In *Actas del III Congreso de la Naturaleza de la Región de Murcia*; García, P., Ed.; ANSE-CEMACAM: Murcia, Spain, 2007; pp. 97–105.
31. Levadoux, L. *Les populations sauvages et cultivées des Vitis vinifera L*; Institut national de la recherche agronomique: Paris, France, 1956; Volume 1, pp. 59–118.
32. Rivera, D.; Verde, A.; Fajardo, J.; Obón, C.; Inocencio, C.; Valdés, A. Modelos etnobiológicos como alternativa al control de malas hierbas con agricultura biológica, los criptocultivos. In *La Malherbología en los Nuevos Sistemas de Producción Agraria*; Mansilla, J.A., Artigao Monreal, J.A., Eds.; Sociedad Española de Malherbología: Albacete, Spain, 2007; pp. 149–154.
33. Gastwirth, J.L. Estimation of the Lorenz Curve and Gini Index. *Rev. Econ. Stat.* **1972**, *54*, 306–316. [\[CrossRef\]](#)
34. Davies, J.; Hoy, M. Making inequality comparisons when Lorenz curves intersect. *Am. Econ. Rev.* **1995**, *85*, 980–986.
35. Dagum, C. The generation and distribution of income, the Lorenz curve and the Gini ratio. *Économie Appliquée* **1980**, *33*, 327–367. [\[CrossRef\]](#)
36. OpenAI. ChatGPT, GPT-4, AI Assistant. Available online: <https://www.openai.com/chatgpt> (accessed on 15 July 2024).
37. Gemini. Bard, AI Assistant. Available online: <https://gemini.google.com/app> (accessed on 15 July 2024).
38. Perplexity, AI Assistant. Available online: <https://www.perplexity.ai/> (accessed on 15 July 2024).
39. Mistral, AI Assistant. Available online: <https://chat.mistral.ai/chat> (accessed on 15 July 2024).
40. Jones, E.T.; McLain, R.J.; Weigand, J. *Nontimber Forest Products in The United States*; University Press of Kansas: Lawrence, KS, USA, 2021; pp. 1–445.
41. Milla, R.; Bastida, J.M.; Turcotte, M.M.; Jones, G.; Violle, C.; Osborne, C.P.; Chacón-Labela, J.; Sosinski, Ê.E., Jr.; Kattge, J.; Laughlin, D.C.; et al. Phylogenetic patterns and phenotypic profiles of the species of plants and mammals farmed for food. *Nat. Ecol. Evol.* **2018**, *2*, 1808–1817. [\[CrossRef\]](#)
42. Heywood, V.H. *Use and Potential of Wild Plants in Farm Households*; FAO: Rome, Italy, 1999; pp. 1–113.
43. Etkin, N.L. The cult of the wild. In *Eating on the Wild Side: The Pharmacologic, Ecologic, and Social Implications of Using Noncultigens*; Etkin, N.L., Ed.; The University of Arizona Press: Tucson, AZ, USA, 1994; pp. 1–21.
44. Harlan, J.R. Genetic Resources in Wild Relatives of Crops. *Crop Sci.* **1976**, *16*, 329–333. [\[CrossRef\]](#)
45. Sõukand, R.; Kalle, R. *Changes in the Use of Wild Food Plants in Estonia 18th–21st Century*; Springer Nature: Berlin/Heidelberg, Germany, 2016.
46. Peters, C.R.; O'Brien, E.M.; Box, E.O. Plant types and seasonality of wild-plant foods, Tanzania to southwestern Africa: Resources for models of the natural environment. *J. Hum. Evol.* **1984**, *13*, 397–414. [\[CrossRef\]](#)
47. Head, L. The social dimensions of invasive plants. *Nat. Plants* **2017**, *3*, 1–7. [\[CrossRef\]](#)
48. Gioria, M.; Hulme, P.E.; Richardson, D.M.; Pyšek, P. Why are invasive plants successful? *Annu. Rev. Plant Biol.* **2023**, *74*, 635–670. [\[CrossRef\]](#)
49. Müller-Schärer, H.; Schaffner, U.; Steinger, T. Evolution in invasive plants: Implications for biological control. *Trends Ecol. Evol.* **2004**, *19*, 417–422. [\[CrossRef\]](#)
50. Leonti, M. The relevance of quantitative ethnobotanical indices for ethnopharmacology and ethnobotany. *J. Ethnopharmacol.* **2022**, *288*, 115008. [\[CrossRef\]](#)
51. Tardío, J.; Pardo-de-Santayana, M. Cultural importance indices: A comparative analysis based on the useful wild plants of southern cantabria (northern Spain). *Econ. Bot.* **2008**, *62*, 24–39. [\[CrossRef\]](#)
52. Medeiros, M.F.T.; Silva, O.S.; Albuquerque, U.P. Quantification in ethnobotanical research: An overview of indices used from 1995 to 2009. *Sitientibus Série Ciências Biológicas* **2011**, *11*, 211–230. [\[CrossRef\]](#)
53. Piketty, T. *Capital in the Twenty-First Century*; Harvard University Press: Cambridge, MA, USA, 2014.
54. Chotikapanich, D. *Modeling Income Distributions and Lorenz Curves*; Springer Nature: Berlin/Heidelberg, Germany, 2008.
55. Duan, C.; Chen, B. Analysis of global energy consumption inequality by using Lorenz curve. *Energy Procedia* **2018**, *152*, 750–755. [\[CrossRef\]](#)
56. GBIF—Global Biodiversity Information Facility. Free and Open Access to Biodiversity Data. Available online: <https://www.gbif.org/> (accessed on 11 September 2024).
57. Pandey, R.; Kumar Khatri, S.; Kumar Singh, N.; Verma, P. *Artificial Intelligence and Machine Learning for EDGE Computing*; Elsevier—Academic Press: Amsterdam, The Netherlands, 2022. [\[CrossRef\]](#)
58. Lauriola, I.; Lavelli, A.; Aiolfi, F. An introduction to deep learning in natural language processing: Models, techniques, and tools. *Neurocomputing* **2022**, *470*, 443–456. [\[CrossRef\]](#)
59. Bird, S. NLTK: The Natural Language Toolkit. In *Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions*, Sydney, Australia, 17–18 July 2006; Curran, J., Ed.; Association for Computational Linguistics: Sydney, Australia, 2006; pp. 69–72. Available online: <https://aclanthology.org/P06-4018.pdf> (accessed on 14 September 2024).
60. Vasiliev, Y. *Natural Language Processing with Python and Spacy: A Practical Introduction*; No Starch Press: San Francisco, CA, USA, 2020; pp. 1–216.

61. Yilmaz, Z.A.; Wang, S.; Yang, W.; Zhang, H.; Lin, J. Applying BERT to document retrieval with birch. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP): System Demonstrations, Hong Kong, China, 3–7 November 2019; Padó, S., Huang, R., Eds.; Association for Computational Linguistics: Hong Kong, China, 2019; pp. 19–24. Available online: <https://aclanthology.org/D19-3004.pdf> (accessed on 14 September 2024).
62. Singh, V.; Kumar, A. *Advances in Bioinformatics*; Springer: Singapore, 2021. [CrossRef]
63. McGinnis, S.; Madden, T.L. BLAST: At the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res.* **2004**, *32* (Suppl. S2), W20–W25. [CrossRef]
64. Uffelmann, E.; Huang, Q.Q.; Munung, N.S.; De Vries, J.; Okada, Y.; Martin, A.R.; Martin, H.; Lappalainen, T.; Posthuma, D. Genome-wide association studies. *Nat. Rev. Methods Primers* **2021**, *1*, 59. [CrossRef]
65. Kazi, L.; Cherkashin, E.; Ristevski, B. IX International Conference, Applied Internet and Information Technologies, AIIT2019, Proceedings; University of Novi Sad, Technical faculty “Mihajlo Pupin”: Zrenjanin, Republic of Serbia, 2019; Available online: https://eprints.uklo.edu.mk/id/eprint/8749/1/Proceedings_AIIT2019.pdf (accessed on 14 September 2024).
66. Zhai, Z.; Martínez, J.F.; Beltran, V.; Martínez, N.L. Decision support systems for agriculture 4.0: Survey and challenges. *Comput. Electron. Agric.* **2020**, *170*, 105256. [CrossRef]
67. Pang, B.; Nijkamp, E.; Wu, Y.N. Deep learning with TensorFlow: A review. *J. Educ. Behav. Stat.* **2020**, *45*, 227–248. [CrossRef]
68. Raschka, S.; Liu, Y.H.; Mirjalili, V. *Machine Learning with PyTorch and Scikit-Learn: Develop Machine Learning and Deep Learning Models with Python*; Packt Publishing Ltd.: Birmingham, UK, 2022; pp. 1–216.
69. Karmaker, S.K.; Hassan, M.M.; Smith, M.J.; Xu, L.; Zhai, C.; Veeramachaneni, K. AutoML to date and beyond: Challenges and opportunities. *ACM Comput. Surv. (CSUR)* **2021**, *54*, 1–36. [CrossRef]
70. AutoML.org. Freiburg-Hannover-Tübingen. AutoML. Available online: <https://www.automl.org/automl/> (accessed on 25 July 2024).
71. H2O. H2O Danube. Available online: <https://h2o.ai/> (accessed on 25 July 2024).
72. Islam, M.; Jin, S. An Overview of Data Visualization. In Proceedings of the 2019 International Conference on Information Science and Communications Technologies (ICISCT), Tashkent, Uzbekistan, 4–6 November 2019; pp. 1–7. [CrossRef]
73. Matplotlib. Matplotlib: Visualization with Python. Available online: <https://matplotlib.org/> (accessed on 25 July 2024).
74. Waskom, M. Seaborn: Statistical Data Visualization. Available online: <https://seaborn.pydata.org/> (accessed on 25 July 2024).
75. Plotly. Plotly.py. Available online: <https://github.com/plotly/plotly.py> (accessed on 25 July 2024).
76. Gonçalves, C.T.; Gonçalves, M.J.A.; Campante, M.I. Developing Integrated Performance Dashboards Visualisations Using Power BI as a Platform. *Information* **2023**, *14*, 614. [CrossRef]
77. Romano, P.; Giugno, R.; Pulvirenti, A. Tools and collaborative environments for bioinformatics research. *Brief. Bioinform.* **2011**, *12*, 549–561. [CrossRef] [PubMed]
78. GitHub. Let’s Build from Here. The World’s Leading AI-Powered Developer Platform. Available online: <https://github.com/> (accessed on 25 July 2024).
79. Jupyter. Jupyter Notebooks. Free Software, Open Standards, and Web Services for Interactive Computing Across all Programming Languages. Available online: <https://jupyter.org/> (accessed on 25 July 2024).
80. Li, L.; Liu, L.; Peng, Y.; Su, Y.; Hu, Y.; Zou, R. Integration of multimodal data for large-scale rapid agricultural land evaluation using machine learning and deep learning approaches. *Geoderma* **2023**, *439*, 116696. [CrossRef]
81. Xie, L.; Draizen, E.J.; Bourne, P. Harnessing big data for systems pharmacology. *Annu. Rev. Pharmacol. Toxicol.* **2017**, *57*, 245–262. [CrossRef]
82. Choi, R.Y.; Coyner, A.S.; Kalpathy-Cramer, J.; Chiang, M.F.; Campbell, J. Introduction to machine learning, neural networks, and deep learning. *Transl. Vis. Sci. Technol.* **2020**, *9*, 14. [CrossRef] [PubMed]
83. Min, B.; Ross, H.; Sulem, E.; Veyseh, A.P.B.; Nguyen, T.H.; Sainz, O.; Aguirre, E.; Heinz, I.; Roth, D. Recent advances in natural language processing via large pre-trained language models: A survey. *ACM Comput. Surv.* **2023**, *56*, 1–40. [CrossRef]
84. Blockeel, H.; Devos, L.; Frénay, B.; Nanfack, G.; Nijssen, S. Decision trees: From efficient prediction to responsible AI. *Front. Artif. Intell.* **2023**, *6*, 1124553. [CrossRef] [PubMed]
85. Antoniadis, A.; Lambert-Lacroix, S.; Poggi, J.M. Random forests for global sensitivity analysis: A selective review. *Reliab. Eng. Syst. Saf.* **2021**, *206*, 107312. [CrossRef]
86. Gawlikowski, J.; Tassi, C.R.N.; Ali, M.; Lee, J.; Humt, M.; Feng, J.; Kruspe, A.; Triebel, R.; Jung, P.; Roscher, R.; et al. A survey of uncertainty in deep neural networks. *Artif. Intell. Rev.* **2023**, *56* (Suppl. S1), 1513–1589. [CrossRef]
87. Dong, X.; Yu, Z.; Cao, W.; Shi, Y.; Ma, Q. A survey on ensemble learning. *Front. Comput. Sci.* **2020**, *14*, 241–258. [CrossRef]
88. Marcílio, W.E.; Eler, D.M. From explanations to feature selection: Assessing SHAP values as feature selection mechanism. In Proceedings of the 2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Porto de Galinhas, Brazil, 7–10 November 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 340–347. [CrossRef]
89. Visani, G.; Bagli, E.; Chesani, F.; Poluzzi, A.; Capuzzo, D. Statistical stability indices for LIME: Obtaining reliable explanations for machine learning models. *J. Oper. Res. Soc.* **2022**, *73*, 91–101. [CrossRef]
90. POWO. Plant of the World Online. Available online: <https://powo.science.kew.org/> (accessed on 18 September 2024).
91. Google Scholar. Google Scholar. Available online: <https://scholar.google.com/> (accessed on 18 September 2024).

92. Web of Science. Web of Science. Available online: <https://www.webofscience.com/wos/woscc/basic-search> (accessed on 18 September 2024).
93. PubMed. National Library of Medicine. Available online: <https://pubmed.ncbi.nlm.nih.gov/> (accessed on 18 September 2024).
94. Molina, M.; Pardo-de-Santayana, M.; Tardío, J. Natural production and cultivation of Mediterranean wild edibles. In *Mediterranean Wild Edible Plants*; Sánchez-Mata, M.C., Tardío, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2016; pp. 81–107. Available online: https://link.springer.com/chapter/10.1007/978-1-4939-3329-7_5 (accessed on 14 September 2024).
95. Mayer-Chissick, U.; Lev, E. Wild edible plants in Israel tradition versus cultivation. In *Medicinal and Aromatic Plants of the Middle-East, Medicinal and Aromatic Plants of the World 2*; Yaniv, Z., Dudai, N., Eds.; Springer: Dordrecht, The Netherlands, 2014; pp. 9–26. [\[CrossRef\]](#)
96. Singh, N.; Pandey, R.; Chandraker, S.K.; Pandey, S.; Malik, S.; Patel, D. Use of Wild Edible Plants Can Meet the Needs of Future Generation. In *Agro-Biodiversity and Agri-Ecosystem Management*; Kumar, P., Tomar, R.S., Bhat, J.A., Dobriyal, M., Rani, M., Eds.; Springer: Singapore, 2022; pp. 341–366. [\[CrossRef\]](#)
97. Bacchetta, L.; Visioli, F.; Cappelli, G.; Caruso, E.; Martin, G.; Nemeth, E.; Bacchetta, G.; Bedini, G.; Wezel, A.; van Asseldonk, T.; et al. A Manifesto for the Valorization of Wild Edible Plants. *J. Ethnopharmacol.* **2016**, *191*, 180–187. [\[CrossRef\]](#) [\[PubMed\]](#)
98. Arreola, J.; Franco, J.A.; Martínez-Sánchez, J.J. Fertilization strategies for *Silene vulgaris* (Caryophyllaceae) production, a wild species with alimentary use. *HortScience* **2004**, *39*, 796D. Available online: <https://journals.ashs.org/hortsci/view/journals/hortsci/39/4/article-p796D.xml> (accessed on 25 July 2024). [\[CrossRef\]](#)
99. Arreola, J.; Franco, J.A.; Vicente, M.J.; Martinez-Sanchez, J.J. Effect of nursery irrigation regimes on vegetative growth and root development of *Silene vulgaris* after transplantation into semi-arid conditions. *J. Hortic. Sci. Biotechnol.* **2006**, *81*, 583–592. [\[CrossRef\]](#)
100. Paschoalinotto, B.H.; Polyzos, C.; Roupahel, A.; Dias, M.I.B.; Petropoulos, S.A. Domestication of wild edible species: The response of *Scolymus hispanicus* plants to different fertigation regimes. *Horticulturae* **2023**, *9*, 103. [\[CrossRef\]](#)
101. Papadimitriou, D.M.; Daliakopoulos, I.N.; Kontaxakis, E.; Sabathianakis, M.; Manios, T.; Savvas, D. Effect of moderate salinity on Golden Thistle (*Scolymus hispanicus* L.) grown in a soilless cropping system. *Sci. Hortic.* **2022**, *303*, 111182. [\[CrossRef\]](#)
102. Paoletti, A.; Benincasa, P.; Famiani, F.; Rosati, A. Spear yield and quality of wild asparagus (*Asparagus acutifolius* L.) as an understory crop in two olive systems. *Agrofor. Syst.* **2023**, *97*, 1361–1373. [\[CrossRef\]](#)
103. Ford-Lloyd, B.V.; Schmidt, M.; Armstrong, S.J.; Barazani, O.Z.; Engels, J.; Hadas, R.; Hammer, K.; Kell, S.P.; Kang, D.; Khoshbakht, K.; et al. Crop wild relatives—Undervalued, underutilized and under threat? *BioScience* **2011**, *61*, 559–565. [\[CrossRef\]](#)
104. Nair, K.P. Utilizing crop wild relatives to combat global warming. *Adv. Agron.* **2019**, *153*, 175–258. [\[CrossRef\]](#)
105. Bohra, A.; Kilian, B.; Sivasankar, S.; Caccamo, M.; Mba, C.; McCouch, S.R.; Varshney, R.K. Reap the crop wild relatives for breeding future crops. *Trends Biotechnol.* **2022**, *40*, 412–431. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.