Journal of Information Economics

Homepage: https://www.anserpress.org/journal/jie

Different perspectives on open data in agriculture: a case study on olive oil

Antonia Ferrer Sapena [a], Ezgi Erdogan [b], Eduardo Jiménez Fernández [c], Elena Sánchez Arnau [a], Enrique A. Sánchez Pérez [a,*], Christian Vidal Cabo [d]

[a] *Instituto Universitario de Matemática Pura y Aplicada, Universitat Politècnica de València, Valencia, Spain*

[b] *Department of Mathematics, Marmara University, Istanbul, Turkey*

[c] *Departamento de Teoría e Historia Económica, Facultad Ciencias Económicas y Empresariales, Granada, Spain*

[d] *Documentación, Universidad Católica de Valencia, Valencia, Spain*

ABSTRACT

Achieving food security through improved agricultural technology is one of the greatest challenges of our time. Indeed, it is one of the elements explicitly mentioned in the Sustainable Development Goals (SDGs). One factor that can lead to improved production and distribution systems is the availability of data of all kinds on the sector. In this paper we study what is happening in this respect in two Mediterranean olive oil producing countries, and show how available data can facilitate the coordination and communication of technical developments and improve the competitiveness of producers. As main conclusions, the paper shows that information systems can be different depending on the approach of governments and their ideological underpinnings of what open data should be. The final target is who are the final beneficiaries of open information, also in the case of technical data such as agricultural data.

KEYWORDS

Open data; Olive oil; Agriculture; Production; Spain; Turkey

* Corresponding author: Enrique A. Sánchez Pérez
E-mail address: easancpe@mat.upv.es

ISSN 2972-3671
doi: 10.58567/jie01040004
This is an open-access article distributed under a CC BY license
(Creative Commons Attribution 4.0 International License)

Received 15 November 2023; Accepted 26 December 2023; Available online 15 January 2024; Version of Record 15 January 2024

1. Introduction

Agriculture today is a knowledge-intensive industry. Over the past two decades the industry has valued data for what it is, a tool to generate, share and exploit knowledge to improve yields, reduce losses and increase returns on investments. Today's agriculture is a more than five trillion dollars annual business (source: FAO). Extensive use of Big Data (Coble et al, 2018; Misra et al, 2020; Osinga et al, 2022) and Artificial Intelligence techniques (Santos et al, 2020; Tantalaki et al, 2019) is now becoming widespread. In particular, Open Data can aid more efficient and effective decision making along the agricultural value chain, fostering innovation through new services and applications and driving organizational change through transparency. As the world's population will grow to around 10 billion by 2050, the global agricultural system will be under great pressure to provide enough food to meet demand. Water resources are drying up, while fertile land is being degraded and genetic resources are disappearing. Unsustainable agricultural practices and other anthropogenic pressures are fueling this degeneration process.

Thus, climate change, the increase in extreme weather events and the (climate change-induced) spread of pests and diseases make agricultural production even more uncertain. The world food market situation does nothing to alleviate this situation that the changing physical environment is forcing. Price volatility has a disruptive effect on production systems and food security, especially for vulnerable populations.

In this overall context, the emergence of data-driven technology to improve all aspects of production, competitiveness, appropriateness and sustainability is an important factor to be taken into account. In this paper we are interested in analyzing how Open Data public policy (and also some private initiatives) can be integrated into the new global agriculture. In particular, we want to analyze how different countries are doing it, based on their own criteria. This is not only a technological issue, but also involves economic, social and ethical aspects (Carbonell, 2016).

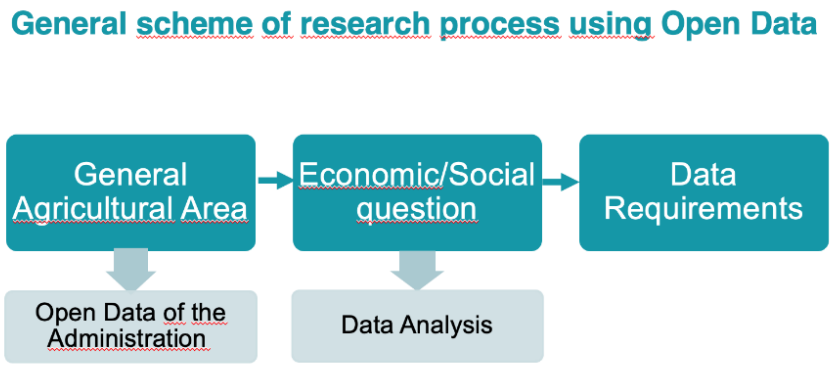

Figure 1. Data-driven socioeconomic research procedure in agriculture.

Our main argument is that its use strongly depends on the socio-economic-cultural situation of each country, as the policy of universalization of Open Data in public administration is not always necessarily good for citizens and data consumers (Fig.1). Let us explain this now. According to Fairbairn and Kish (2023), although in some countries Open Data is connected to the notion of government transparency, the mainstream conceptualization and use of Open Data is not. According to these authors, Open Data in public administrations is an apolitical initiative, a priori without social content, and its main objective is to limit the action of governments by tightly controlling them, and to serve large companies to improve their performance in order to increase their profits. Thus, at least in the USA, it is entirely consistent with a neoliberal ideology, and no positive social impact is expected from its use. However, in Europe there is a different way of understanding it, more related to governmental transparency and citizen control of state expenditures, as well as a source of social projects to help men in agriculture in their work

and the government in designing new strategies to improve competitiveness and other factors. This, which is the usual situation when comparing European countries with the USA due to the traditional difference in economic systems, in this case more pronounced due to the Common Agricultural Policy, which was one of the main pillars in the construction of the EU, and is still very important.

However, the European statistical institutes generally do a very comprehensive job of opening up all kinds of statistical data on all related topics (social, economic, technical, resources, etc.). All this abstract reasoning must be set in a concrete context to make sense in order to be able to provide objective quantitative information. In this paper we focus on data related to olive oil production, which is traditionally centered on Mediterranean countries. We will use the cases of Turkey and Spain to show how diverse the potential use of Open Data can be, and how this depends very much on the cultural, economic, political and social environment.

Thus, the objectives of this paper are to analyze the different responses that are currently given to the requirement to make agricultural data obtained by public administrations freely available, in order to see how diverse they can be, as well as to provide some clues from the data collected as to what might be the objectives of governments to support (or not) the opening of data, focusing attention on who might reap the benefits of such a policy.

1.1. Open data in agriculture

For data users, trust in the use and exchange of data depends on several factors. Knowing the source is fundamental. Trust starts with knowing the origin, and it is necessary to trust the source. If the source is known to be trustworthy, the conclusions we draw from it will be trustworthy. In addition, we must verify the timeliness and quality of the data. Data are not useful if they are out of date. Reliable data must accurately and precisely reflect what they indicate.

Let us now point out what are the main axes for the analysis of the quality of open data in public administrations. The first axis with respect to the Open Data philosophy is that all data related to public affairs should be open. On the other hand, according to the International Open Data Charter, open data should be timely and understandable, accessible and reusable, comparable and interoperable, and should contribute to improved governance and citizen participation. But the issues surrounding the use of open data are not limited to these technical requirements. Other factors are just as important and relate to the way in which data are generated and used. There is a long list of ethical requirements in the use of Open Data. This link between openness and accountability in relation to data can best be approached by focusing attention on the issues that today appear daily in the public arena. Firstly, ethical issues: who owns the data, who can use it for profit (power, culture and technical capacity). Democratic governments are obliged to fight to allow all citizens (and not only big companies) to profit from the use of data. Thus, there is a need to monitor data breaches that may affect vulnerable communities. On the other hand, it is a constant situation around the world that key datasets remain absent or inaccessible, as this would not be the case if the ideology of Open Government were shared by democratic institutions. In summary, we can say that governments have a key role to play in setting the framework for openness and supporting the infrastructure necessary for a sustainable open data commons for agricultural research and practice.

Another relevant aspect of open data is that it must have basic technical requirements in order to be useful for the community (Addison et al, 2015). They must be sustainable, i.e. they must be preserved over time. Furthermore, their availability to all citizens must be ensured, and their availability must be real, i.e. the data must be discoverable, as they are only useful if they are easy to find. Documentation on its use and the technical support necessary for effective use must be made available to all interested parties. Users must be able to access the data medium if necessary, to ensure the veracity of the information and its availability, and they must be able to inform those responsible if they detect any errors. Therefore, all stakeholders need to be connected. Many elements are involved.

It must be taken into account that the data and metadata being collected are complex because they are generated by thousands of stakeholders from multiple sectors, using many types, formats and ontologies.

This brings us to the need to consider what the standards are for Open data in agriculture. In general, they are very varied: vocabularies, taxonomies, measurement protocols, data models and equipment interfaces. There are several standards with low interoperability, as they have been developed to serve specific subfields. Thus, GODAN Action has mapped agri-food standards noting, among many other things, that 16% of standards are not online, only 56% are machine-readable and only 21% are clearly available under open licenses. GODAN and the FAO-sponsored Agricultural Information Management Standards (AIMS) initiative have developed the VEST Registry to make standards more open and useful by cataloguing ontologies in use in different agricultural subfields.

But where does data in agriculture come from? Agribusinesses collect, analyze and use data to inform changes in the services or products they intend to market to customers. Data can be aggregated from a variety of sources, whether internal, from government agencies or from their own customers. Larger agricultural companies are increasingly interested in open data and are working on global issues such as conservation and food security. This means that there are many issues and discussions currently underway regarding the extent to which industry, government and research institutions should have access to data on farming practices, land use and chemical use by individual farmers. In this context, some very relevant initiatives have emerged to try to ensure data collection and management in a cooperative and legally secure environment. The most relevant example is probably the Open Ag Data Alliance, a platform that defines itself as follows. "The Open Ag Data Alliance is an open project designed to bring interoperability, security, and privacy to agricultural data." With regard to its operability, it is summarized in the following maxim, which the platform has as its declaration of principles. "The central guiding principle of OADA is that each farmer owns data generated or entered by the farmer, their employees, or by machines performing activities on their farm."

In addition to the companies, it should be taken into account that there are other actors, both public and private, who also provide and use agricultural data. These are researchers at universities, study centers, research institutes, as well as the organizations themselves that are related to agricultural production in some way. All this gives us an idea of the complexity of agricultural data management, and makes it necessary to work to improve standards, data collection processes and the accessibility of all this information.

Let us finish this section by reporting a relevant international project related to open data in agriculture as a sample of how international efforts are directed to create a data network for general users. There are several initiatives to meet the information needs of agriculture, which follow more or less the same general aim. Let us describe the Open Data Charter. Its objective is to facilitate its use by the sector, helps empower farmers, optimizes farms, stimulates the rural sector and enhances the agricultural value chain. The Open Data Charter identifies the 14 key data categories for the sector that governments should make an effort to publish. In its presentation, this organization exposes that "The Open Data Charter is a collaboration between over 170 governments and organizations working to open up data based on a shared set of principles. We promote policies and practices that enable governments and CSOs to collect, share, and use well-governed data, to respond effectively and accountably to the following focus areas: anti-corruption, climate action and pay equity." This platform has a specific section devoted to agriculture, that is called the Agriculture Open Data Package (AgPack). It is possible to find in it a big collection of different documents and data bases regarding fundamental aspects on this general topic. The categories that can be found in it are shown in Figure 2.

This general scheme contains all the elements that are assumed to be relevant for the competitiveness of agricultural production. Any other international platform trying to cover this topic must consider all these aspects. In this sense, countries that record these data and place them in a public repository for use by the general public have a technical advantage over others. These ideas lead to the general matter of searching for a good strategy in

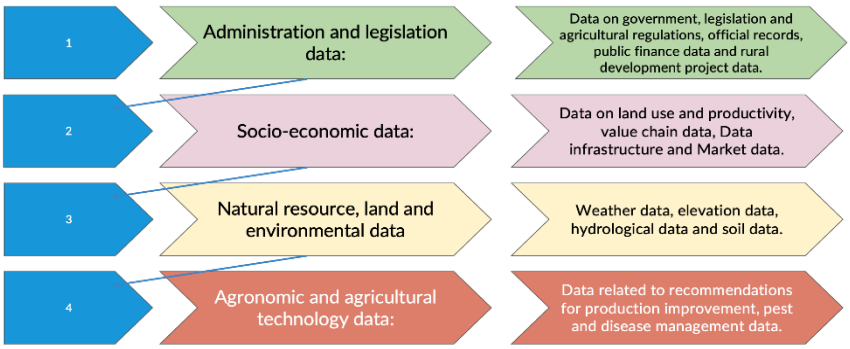

Figure 2. Categorization of data related to data analysis in agriculture.

the new data-based technological development for agriculture, mainly devoted to politicians who have to design agricultural-supporting government policies and government open data policy. The following questions (and their answers obtained with the help of data) should be the starting point for developing such a strategy: What is the role of agriculture in development? What is the current agricultural policy and how will it boost agricultural development? What data can be published to support development? What are the current government policies regarding data, and can they be used to support data publication? What international policy frameworks can be used? Let us fix some answers to these questions in the context of the olives oil business.

2. Production of olive oil in the world

Nowadays, competitiveness of the producers of olive oil is a global mater. Most producing countries have to compete in shared markets, often constrained with similar requirements, problems and situations. Two aspects stand out in this matter. The first one is the strong roots of this crop in the Mediterranean area, based on the traditional consumption of this product in these countries (Fig.3). The second one is that besides, there is a growing interest of the global markets in olive oil due to the increasing consume in other countries without this tradition. The problem so is how to face the challenge of surviving in such a global market improving the technology but preserving the traditional elements of this production in order to maintain the quality and the consumer approval. Our idea is that detailed open data on the production and consumption of olive oil can play a central role in this emerging global market.

Figure 3. Olive tree in its natural growing area.

Let us present first a global view of the production of olive oil. (We use data before 2020 in order to reduce the distortion produced in the global markets in the subsequent years by the COVID-19 pandemic.) Olives find their natural environment in the mediterranean climate. This is the reason why the main production of this product is located in Spain, Italy, Greece, Turkey, Tunisia, Syria, Morocco, Egypt, France and Portugal (Table 1). Among them, the main production area is in Spain, with a significant difference, and the three leading producers (Spain, Tunisia, Italy) account for more than half of the world production area (Fig.4).

Table 1. Olive grove areas in different countries (thousand hectares).

	2017	2018	2019
Spain	2.555	2.579	2.602
Tunisia	948	1.534	1.607
Italy	1.142	1.142	1.139
Morocco	1.021	1.045	1.073
Greece	793	963	903
Turkey	846	864	879
Syria	692	693	693
Algeria	433	431	432
Portugal	358	361	360
Other	909	874	889
World	9.697	10.486	10.578

Source: FAO, 2022 (Access date: 28.10.2022).

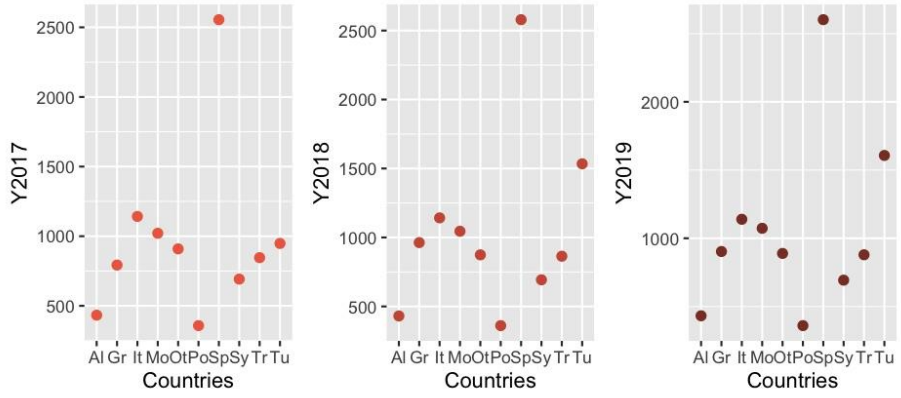

Figure 4. Olive growing areas by country.

Notes: The first two letters of the names indicate the corresponding country on the X-axis of the figure.

Olive production has increased since 2010, but has not been regular due to the natural circumstances affecting agricultural production. Spain, Italy, Greece, Turkey and Morocco, in that order, are the most productive countries (source: FAO). However, the production rate per cultivated area is not regular. The world average yield of olive fruit was 20.432 kg/hectare in the decade of 2010, being Spain with 27.083 kg/hectare the country with the highest rate, followed by Italy (22.458), Greece (28.216) and Turkey (20.432). (Table 2).

The second part of the equation of economic viability of the olive oil production is the demand of the market. Regarding habits of consumption of olive oil, the main trends are conditioned by deep cultural reasons. Olive oil is a fundamental actor in the so-called mediterranean diet, that essentially comes from the traditional mediterranean cook. Since this cooking is becoming popular all over the world, as well as being preserved in the mediterranean countries by cultural reasons, the natural trend in the world olive oil consumption is to increase. Spain, Italy and

Table 2. Olive oil production (thousand tons).

World olive oil production (thousand tons)			
	2017/18	2018/19	2019/20
Spain	1.262,2	1.789,9	1.125,3
Italy	428,9	173,6	366,0
Greece	346,0	185,0	275,0
Turkey	263,0	193,5	225,0
Morocco	140,0	200,0	145,0
Tunisia	325,0	140,0	350,0
Syria	100,0	104,0	120,0
Portugal	134,8	100,3	140,5
Algeria	82,5	97,0	125,5
Other	296,6	278,7	334,7
World	3.379,0	3.262,0	3.207,0

Source: IOOC (International Olive and Olive Oil Council), 2021 (Access date: 28.10.2022).

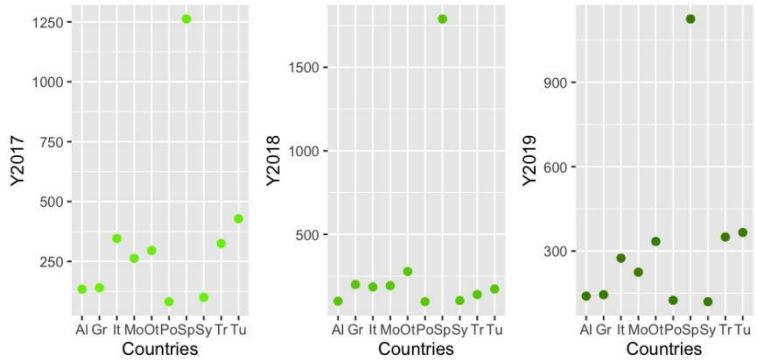

Figure 5. Production of olive oil by countries.

Notes: The first two letters of the names indicate the corresponding country on the X-axis of the figure.

Greece focus their export in its natural market the European Union and the UK, in which a stable high demand is reported.

Then, the data from the International Olive Council (IOC) reflects that world olive oil consumption increased by 5.8 % to attain the value of 3.23 million tons in 2019/20. Out of the European Union, which market is in a sense saturated, the demand for olive oil is shifting to USA (a big consumer), Japan, Brazil, Australia and recently China.

The productivity of olive trees depends to a large extent on the method of agricultural production (Fig.4). In countries where cultivation technology is developed, three main methods are used: traditional (dryland agriculture), irrigated cultivation area (intensive olive groves) and high-yield fields (super-intensive olive groves). However, in many cases the use of advanced technology is not only a question of economic investment, but also a cultural issue, or the result of general agricultural policies set by the states or (in the case of Spain) the European Union. Thus, traditional olive growing requires few water resources and technological aids, but implies the need for a large amount of human resources, which necessarily degrades the economic situation of the workers' collectives, since intensive and super-intensive production is more competitive. The reader can find in Table 2 the production of olive oil by countries (Fig.5.). This makes a strong difference in terms of the price of production. In general, in countries where labour force is expensive, highly technified production allows a large reduction in production costs, which incentivizes the introduction of high level technified methods (intensive and super-intensive production), relegating traditional rainfed olive grove production to the most depressed areas. But very often these regions, due to the properties of the territory or to the lack of investors, are not able to improve their competitivity through the

introduction of new technological tools (Fig.6).

Figure 6. Small size olives. The productivity of olive trees depends on the production method.

This leads to a difficult situation for poor regions that have to compete in the same market with highly developed production companies, since, paradoxically, production costs increase in poorer areas, which does not allow these regions to take advantage of the reduction prices to remain competitive. This is what is happening, for example, in the south of Spain, with a long tradition in the production of olive oil, where the (economically depressed) territories with dryland olive groves have no chance of competing with the new companies that use super-intensive production methods.

Two main solutions have been proposed to enable all production areas to remain competitive in local and international markets. The first is to try to give added value (e.g. quality, specificity or traditional flavor) to oil produced in special environments. The other is to try to obtain special regulations and protection for this type of agriculture, as is the case, for example, in the countries of the European Union. (Common Agricultural Policy (CAP) of the European Union.) In general, the second option leads to a non-competitive scenario, in which the basis of the activity does not lie in the market, but in the regulatory framework.

In any case, the general trend is to converge towards a global scenario in which the use of high-level technology seems to play a key role. This is the point at which the use of open public administration data, as part of a complex management environment, could make the difference between different productive systems.

3. A case study. Two ways of understanding open agricultural information: Turkey and Spain

There are several recent papers analyzing some questions related to the production of olive oil in some mediterranean countries. We will describe the situation of open data in two production countries, that are in a sense representative of different ways of managing open data, as well as the analysis that scientific make of this fact. The first reference paper is the one of Pehlivanoğlu (2021). This paper studies the factors affecting competition in olive oil exports, understood through a revealed comparative advantage (RCA) index. The determinants that are studied are price, consumption in third countries and production. The hypothesis of this article is based on an article that studied this phenomenon but of other products, coconut and something else, so perhaps these hypotheses are not entirely correct. On the other hand, there is another recent article that criticises the index used, giving the reasons why it is not very reliable and proposes an alternative indicator (Danna-Buitrago, 2021). The first point to note is the different nature of the reports/data published in the two countries. In the case of the Turkish government report on the subject, there are some issues that are not compatible with the way the Spanish data are presented. The fragmentation of output used in the Spanish report of the National Institute of Statistics uses different categories

than those used in the Turkish report. It is convenient to distinguish whether we are talking about extra virgin olive oil production or what is equivalent to an acidity below 1 degree, if we are talking about other types of oils or table olives. The information provided in the Turkish report does not distinguish all this and may cause confusion, although the initial tables may serve perfectly well to define a new vulnerability index in olive oil production areas.

3.1. The specific case of Turkey

The information on olive oil production in Turkey that is provided by the public administration is mainly presented in a deep and comprehensive report (Okzan, 2021). Turkey is one of the main olive oil productor in the worlds. In 2020, a total of 8.87 million decares of olive groves were recorded in Turkey. According to TurkStat (Turkish Statistical Institute) data the total olive grove area increased by 11.1 % between 2011 and 2020. For 2020, 73.7 % of the total olive grove area is for oil olives and 26.3% for table olives. The areas allocated for oil olives increased by 13.4 % in ten years. When the number of olive trees in Turkey is analyzed, there are 159.38 million fruit-bearing olive trees in 2020 and 68.3 % of the fruit-bearing olive trees are reserved for oil and 31.7 % for table olives. In terms of the proportion of trees devoted to each type of production, 61 % of the amount of olive production realised in 2020 in Turkey corresponds to oil olives and 39 % to table olives. Between 2011 and 2020, there are fluctuations in the production of grain olives due to climatic conditions and periodicity. All these statistics effectively place Turkey at the top of world production.

The Turkish report ends by providing the reader with a lot of general information on import and export data. All this information can be successfully used, for example, by analysts and politicians who want to design new agricultural policies in the country, and also for research on the economy of Turkish regions in relation to macroeconomic issues. Even large companies can benefit from this comprehensive overview report, as they can obtain valuable information on the export viability of a large agricultural company. However, this data is not useful for individual owners who wish to improve their productivity, or analyse the possibility of changing their production procedure to obtain a more competitive product for export.

Olive oil consumption in 2019 was over 5 kg per capita in producing countries in the European Union such as Greece, Spain, Italy and Portugal, while the EU average consumption was 2.9 kg. In non-IOC countries, Syria stood out with 4 kg, while the average overall consumption of these countries remained at 0.4 kg. In Turkey, per capita consumption approaches two kilograms. Thus, Turkish olive oil consumption has increased by about 4\% compared to 10 years ago. However, consumer prices of olive oil are high compared to other vegetable oils, putting pressure on domestic olive oil consumption. Domestic olive oil consumption per capita is quite low compared to the consumption of producing countries in Europe.

Probably due to cultural reasons, the Turkish population consume less olive oil than other mediterranean countries, so the role that plays such product in the habits of consume of the people is different than the one that plays in the southern countries of the EU. Consequently, also taking into account the volume of production, the public effort in providing data to improve competitiveness is not as high as in the case of Spain, as we explain below.

3.2. The case of Spain

In Spain, the updating of agricultural technology for olive growing has revolutionised the sector in recent years. While marketing is becoming less and less profitable for traditional producers, the overproduction of intensive farms using all the technological advances continues to impose itself on traditional olive growing. EU policies facilitate the granting of subsidies, often without clear criteria, which, given the circumstances, mean that many conventional producers turn the profitability of production into a secondary value, relying on these subsidies (from the CAP, Common Agricultural Policy) to make the business profitable in the end. The change of model towards

intensive production has consequences at other levels, such as social (depopulation of certain areas due to lack of resources) and ecological (intensive and super-intensive cultivation requires large quantities of water, a scarce resource in the productive areas of southern Spain).

Even so, olive oil production is today one of the fundamental sources of wealth in many areas of the Iberian Peninsula, especially in the south and east, where entire regions are dedicated to olive monoculture. However, certain forms of production such as rainfed production in mountain areas are becoming less and less viable, sometimes leading to the abandonment of olive groves, with serious consequences such as the loss of jobs in the countryside, the depopulation of large areas and the deterioration of soils (Rodríguez and Balmón, 2011).

In the case of Spain, in contrast to the relative absence of open statistical data in the Turkish case, the different public institutes that handle information on the Spanish territory provide a lot of material, which allows both the work of scientists and the preparation of technical reports to advise policy makers, as well as the use by companies and individuals to evaluate the profitability of their initiatives in an informed manner. The open data available and made available to the public by different institutions and associations, which we will see below, deal with the following topics: production data and farming patterns, geographical distribution and farm management data, population data and environmental data. There are several institutions and associations that provides this information. For example, for topics related to water resources and soil, can be found in the platform of the local government of Andalucia (south of Spain) and in the web page of the Instituto de Estadística y Cartografía de Andalucía. Topics as soil losses in Andalusia, according to erosion levels, as well as data on the hydraulic infrastructures and water resources can be found there and in the platform of the Spanish Instituto Nacional de Estadística (INE). Data on the production of olive oil are provided also by some associations not directly connected to the government, as the association of Cooperativas Agroalimentarias de Andalucía and the Spanish Association of Olive Municipalities. Consumer trends can be found in the platform of the Spanish Ministerio de agricultura pesca y alimentación; and, in general, more public and private associations put open data available to public. The Sistema de Información Multiterritorial de Andalucía (SIMA), that is provided by the Instituto de Estadística y Cartografía de Andalucía, is a very helpful tool and becomes the main source used to obtain data, as it collects open and reusable data from other sources.

In summary, the relevance of olive oil production in Spain for the welfare of large regions of the country (mainly low-income regions), together with the decision of the state to support Open Data policies, results in a developed information system.

4. Discussion

As we mentioned in the Introduction, open data policies from the governments are more connected with the ideological roots of the economical design of their political actions. As pretended, Open Data is an apolitical trend but with deep roots in the stated ideology regarding economics. In the country where this movement started (USA), there is no discussion about the role of open data of the public administration. As Fairbairn and Kish (2023) explain, open data is an instrument to control the state (with the final aim of limit their areas of competence) as well as a tool for making easier the development of big companies with the help of this information. This way of understanding is not at all followed by all the countries, even those that clearly work under liberal economic rules. As we have seen, Turkey and Spain represent different schemes of use of open data in agriculture. Turkey represents a conservative policy regarding agriculture data. While in Turkey the way data is exposed is merely informative and is not designed for its exploitation for any agent (no general citizens, no individual farmers, no big agriculture companies), Spain have developed a big platform to facilitate the access to all this information. The way this is done make clear that the general aim is to offer an easy instrument for individuals, scientists and companies to get these data, that can potentially revert in a direct benefit to society.

Let us explain a concrete example of such application. The abundance of data allow to precisely compute economical indices of social interest which could help to the design of social policies, as the notion of vulnerability that can be quantified defining an index that is inversely proportional to the concept of sustainable agriculture (Jiménez-Fernández et al, 2023). Gómez-Limón Rodríguez and Sánchez Fernández (2010) base sustainability on three components: 1. Environmental sustainability (those processes compatible with the maintenance of natural ecosystems ecosystems capable of guaranteeing agricultural continuity.) 2. Economic sustainability (the economic viability of farms and a positive contribution of the sector as a whole to regional/national income.) and 3. Socio-cultural sustainability (ensures food sufficiency and equity in the distribution of the income generated by agriculture, as well as contributing to the viability of rural communities.) Translating to the vulnerability case, the associated index helps to detect which municipalities of the olive oil production areas are in danger of suffer bad life conditions to the inhabitants and help them to improve the situation. In order to define it, several indicators are needed, as well as data to compute them (Jiménez-Fernández et al, 2023). For example, environmental and geographic data are necessary: erosion, distance to the main town of the region, slope and irrigation. All of them can be found in the platform of the Spanish Instituto Nacional de Estadística, the Instituto de Estadística y Cartografía de Andalucía and the Consejería de Agricultura, Ganadería, Pesca y Desarrollo Sostenible.

Economical regional data are also needed. For example, ratio of proprietaries of the grown fields that are older than 65, or the average income of the inhabitants of a given municipality; these can be found in the platform of the Ministerio de Hacienda y Función Pública. Agencia Tributaria. Finally, social data as ratio of internet connection or ratio of inhabitants living in urban centers are necessary, and can be found in the web page of the Ministerio de Asuntos Económicos y Transformación digital. The possibility of acceding to all these data does not means that the situation is 100 % satisfactory. Far to this, often happens that data posted in the information repositories fail with the basics of the FAIR principles. Often a long and cost data curation process is needed. Summing up, accessibility to open data is difficult even if public administration has a clear policy of data openness, but this policy strongly depends on the state assumptions on what economic model has to be followed, although this is sometimes not explicitly exposed by the governments. Up to some point, this issue can be fixed in a simple question. Who are the beneficiaries of the open data policy in each country?

Often, in addition to public institutions, smallholders produce data on their farms and can share them. They can use the information produced from these data, or from external sources. This information can help them to know when and where to plant and harvest, what the market prices are, or much other information of direct relevance to their businesses. But are smallholders really the ones who benefit? Who are the beneficiaries? It is clear that, in rural areas and in most countries, there is a lack of digital skills, access to technology, the internet, or word processing and spreadsheet management skills. Traditionally, these skills are lacking in rural agricultural areas and among women and vulnerable communities.

On the other hand, large companies also benefit from this information, but generally in a much more efficient way, and under a much more aggressive competitive model. So the question is, who benefits the most: does society get enough benefit to justify public spending on the welfare of citizens, or should the state rather provide this data for the use of large companies, so that they can contribute to the development of regions through their own achievements? As we have argued throughout this paper, the answer to this question is deeply ideological, and is never clearly stated by public authorities. As a consequence, open data policy around the world necessarily constitutes a patchwork of different practices, paradoxically contradictory in some cases, and in general very unorganized, despite the efforts of certain international initiatives and bodies. But, because of the very different ideological approaches that underlie them, this may necessarily be the case.

5. Conclusions

Open data in agriculture could help improve productivity and competitiveness in a global market. However, the principles of open data do not exactly coincide with the notion of social development. Born in the context of a technological revolution, mainstream theoretical thinking about open data assumes that it is part of an apolitical action, undertaken primarily to aid (general economic) development as well as to serve as an effective check on government spending. Thus, from this point of view, the open data movement is predominantly an economic tool. These data can be used by large companies to design better methods of exploitation, as well as by governments and civil associations to improve individual production conditions and social development.

Consequently, open data is not necessarily a vector for social development, but provides an advanced framework to help increase the productivity and living conditions of agricultural workers and farmers by helping to provide accurate data on the economic and social situation of the people involved. For open data to become an agent of social development, government regulations on how to use it are necessary, as well as control of the exploitation rules of large companies. Open data have to be accompanied by other complementary policies and economic resources for the collective of agricultural workers so that the data can be used by all citizens on equal terms. It is not necessarily true that open data per se helps this goal; on the contrary, without proper regulation and institutional, scientific and citizen projects to use it as such a vector of social goodness. Large companies can benefit from open data much better than individuals, as this often requires high-level technological knowledge and tools.

Therefore, open data is neither bad nor good. It is its use that should make the difference between it becoming a free tool for mass exploitation or a social benefit. Governments or international institutions are the actors that should control what kind of data should be open, and try to articulate the right tools to make Open Data a factor of community benefit.

Funding Statement

This research was funded by Ministerio de Ciencia e Innovación (Spain), grant number PID2019-105708RB "Stable methodologies to evaluate and measure quality, interoperability, blockchain, and reuse of opendata in the agricultural field: DATAUSE" funded by MCIN/AEI/https://doi.org/10.13039/501100011033.

Acknowledgments

We thank Observatori Valencià de Dades Obertes y Transparència (Generalitat Valenciana, Universitat Politècnica de València) by the technical support provided for carrying this research. We also thank Marmara University (Istanbul) for facilitating the collaboration. Finally, we also thank Zeynep Özcan, head of the EU Harmonization department of the Turkish Ministry of Agriculture and Forestry, for providing us with information on open data in Turkey, and Zerrin Özcan, author of the rigorous and extensive report on olive oil production in Turkey.

Conflict of interest

All the authors claim that the manuscript is completely original. The authors also declare no conflict of interest.

Author contributions

Conceptualization: Ezgi Erdogan, Antonia Ferrer, E. Jiménez and Enrique A. Sánchez. Data curation: Elena Sánchez and Christian Vidal. Formal analysis, funding acquisition: Antonia Ferrer. Investigation: Ezgi Erdogan, Elena Sánchez and Christian Vidal. Methodology: Antonia Ferrer and Enrique A. Sánchez. Resources, software, supervision,

validation, visualization: Ezgi Erdogan, Elena Sánchez, Christian Vidal. Writing–original draft, and writing–review & editing: Antonia Ferrer and Enrique A. Sánchez.

References

Carbonell, I. (2016). The ethics of big data in big agriculture. *Internet Policy Review* 5(1) https://doi.org/10.14763/2016.1.405

Coble, K. H., Mishra, A. K., Ferrell, S., and Griffin, T. (2018). Big data in agriculture: A challenge for the future. *Applied Economic Perspectives and Policy* 40(1), 79-96 https://doi.org/10.1093/aepp/ppx056

Danna-Buitrago, J. P., and Stellian, R. (2022). A New Class of Revealed Comparative Advantage Indexes. *Open Economies Review* 33(3), 477-503 https://doi.org/10.1007/s11079-021-09636-4

Fairbairn, M., and Kish, Z. (2023). Setting data free: The politics of open data for food and agriculture. *New Media & Society* 25(8), 1935-1959.
https://www.fao.org/faostat/en/#home. https://www.fao.org/food-systems/es/

FAO, CIHEAM, UfM, PRIMA. (2023). Sustainable agrifood systems in the Mediterranean. https://www.fao.org/documents/card/en/c/cc8220en.
https://godan.info/godan-action. https://godan.info/research-resource

Gómez-Limón, J. A. and Granado-Díaz, R. (2022). La Viabilidad de las Explotaciones Comerciales de Olivar en España. *Studies of Applied Economics* 40, 2, 1-15 https://doi.org/10.25115/eea.v40i2.8400

Gómez-Limón Rodríguez, J. A. and Sánchez Fernández, G. (2010). Empirical evaluation of agricultural sustainability using composite indicators. *Ecological Economics* 69, 1062–1075 https://doi.org/10.1016/j.ecolecon.2009.11.027

Addison, C., Jellema, A., and Meijninger, W. (2015). Open data and smallholder food and nutritional security. CTA Working Paper 15/01

Jiménez-Fernández, E., Ferrer Sapena, A., Sánchez-Pérez, E. A., and Petrosyan, L. (2022). Vulnerability of Olive Municipalities in Southern Spain. Available at SSRN 4254741

Kamilaris, A., Kartakoullis, A., and Prenafeta-Boldú, F. X. (2017). A review on the practice of big data analysis in agriculture. *Computers and Electronics in Agriculture* 143, 23-37
https://doi.org/10.1016/j.compag.2017.09.037

Misra, N. N., Dixit, Y., Al-Mallahi, A., Bhullar, M. S., Upadhyay, R., and Martynenko, A. (2020). IoT, big data, and artificial intelligence in agriculture and food industry. *IEEE Internet of things Journal* 9(9), 6305-6324. https://opendatacharter.net/. https://opendatacharter.net/resources/

Osinga, S. A., Paudel, D., Mouzakitis, S. A., and Athanasiadis, I. N. (2022). Big data in griculture: Between opportunity and solution. *Agricultural Systems* 195, 103298
https://doi.org/10.1016/j.agsy.2021.103298

Özkan, Z. Ürün Raporu Zeytinyagi. (2021). Tarımsal Ekonomi ve Politika Geliştirme Enstitüsü Müdürlüğü

Pehlivanoğlu, F., Erarslan, C., and Demir, S. (2021). Factors affecting competition in olive oil exports: Panel data analysis of selected countries. Case study. A*gricultural Economics* 67(12), 511-518
https://doi.org/10.17221/494/2020-AGRICECON

Rodríguez, J. A. G. L., and Balmón, M. A. (2011). Evaluación de la sostenibilidad de las explotaciones de olivar en Andalucía. *Editorial Analistas Económicos de Andalucía,* Málaga, 2011

Santos, L., Santos, F. N., Oliveira, P. M., and Shinde, P. (2020). Deep learning applications in agriculture: A short review. I*n Robot 2019: Fourth Iberian Robotics Conference: Advances in Robotics*, Volume 1 (pp. 139-151). Springer International Publishing

Tantalaki, N., Souravlas, S., and Roumeliotis, M. (2019). Data-driven decision making in precision agriculture: The rise of big data in agricultural systems. *Journal of agricultural & food information* 20(4), 344-380
https://doi.org/10.1080/10496505.2019.1638264