*data*

MDPI

*Article*

# Perception and Reuse of Open Data in the Spanish University Teaching and Research Community

**Christian Vidal-Cabo [1,*], Enrique Alfonso Sánchez-Pérez [2] and Antonia Ferrer-Sapena [2]**

[1] Servicio de Publicaciones, Universidad Católica de Valencia San Vicente Mártir, C/de Quevedo, 46001 Valencia, Spain
[2] Instituto Universitario de Matemática Pura y Aplicada, Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain; easancpe@mat.upv.es (E.A.S.-P.); anfersa@upv.es (A.F.-S.)
* Correspondence: christian.vidal@ucv.es

**Abstract:** Introduction. Open Government is a form of public policy based on the pillars of collaboration and citizen participation, transparency and the right of access to public information. With the help of information and communication technologies, governments and administrations carry out open data initiatives, making reusable datasets available to all citizens. The academic community, highly qualified personnel, can become potential reusers of this data, which would lead to its use for scientific research, generating knowledge, and for teaching, improving the training of university students and promoting the reuse of open data in the future. Method. This study was developed using a quantitative research methodology (survey), which was distributed by email in one context block and six technical blocks, with a total of 30 questions. The data collection period was between 15 March and 10 May 2021. Analysis. The data obtained through this quantitative methodology were processed, normalised, and analysed. Results. A total of 783 responses were obtained, from 34 Spanish provinces. The researchers come from 47 Spanish universities and 21 research centres, and 19 research areas of the State Research Agency are represented. In addition, a platform was developed with the data for the purpose of visualising the results of the survey. Conclusions. The sample thus obtained is representative and the conclusions can be extrapolated to the rest of the Spanish university teaching staff. In terms of gender, the study is balanced between men and women (41.76% W vs. 56.58% M). In general, researchers responding to the survey know what open data is (79.31%) but only 50.57% reuse open data. The main conclusion is that open government data prove to be useful sources of information for science, especially in areas such as Social Sciences, Industrial Production, Engineering and Engineering for Society, Information and Communication Technologies, Economics and Environmental Sciences.

**Keywords:** open data; scientific community; reuse of public service information; open government



**Citation:** Vidal-Cabo, C.; Sánchez-Pérez, E.A.; Ferrer-Sapena, A. Perception and Reuse of Open Data in the Spanish University Teaching and Research Community. *Data* **2024**, *9*, 117. https://doi.org/10.3390/data9100117

Academic Editor: Kesheng (John) Wu

Received: 20 June 2024
Revised: 21 September 2024
Accepted: 27 September 2024
Published: 11 October 2024



## 1. Introduction

Currently, one of the main vectors of innovation in research, teaching and other classic academic activities is the use of open data, in any of the multiple ways in which this term is understood. Essentially, there are two meanings in which the expression "open data" can be situated in relation to academic work [1]. The first one is open research data, which concerns the sharing of the scientific data that research groups obtain as a result of their experimental work [2]. The second is the data that administrations make available to the general public, normally related to the activities of the State, economic and social data, and any other type of data that may be of interest to the citizen in general; the scientific community is currently analyzing the real potential of public data spaces to foster innovative progress and digital transformation [3–5]. In this article we present the results of an experimental study carried out in Spanish universities on the knowledge and habits of professors and researchers related to the use of open data from public administration,

both for their professional and research or teaching activities. Our aim in conducting this research is to provide a realistic picture through a case study of how knowledgeable a higher education teacher is concerning open data concepts and tools. Although the use and reuse of open data is a hot topic and the focus of many researchers [6–8], there is not much research on the actual situation in academic institutions.

It is clear that the use of Open Data in academia affects the advancement of science and technology transfer, but also teaching, and therefore the training of new professionals [9,10]. Indeed, the advent of the open data culture has become a fundamental feature of today's understanding of the world. Almost every aspect of human activity involves the use of data, directly or indirectly. In particular, scientific research and higher education are areas where the new philosophy of data-driven knowledge is particularly relevant. However, as some researchers have pointed out, it is not always being incorporated into education in an optimal way [11]. As is often the case when a new paradigm disrupts the state of the art in a well-established environment, datafication is triggering fragmented responses, ambiguity and, at worst, damage, in many areas of higher education [12,13] and scientific research.

In order to facilitate the interpretation of the quantitative results that will be given in the paper, let us start in this introduction by explaining some key points for understanding the context. The main contextual elements to be taken into account are related to the definition of the framework in which open governmental data is used in the academic environment, and how this is reflected in the university structure of a European country, such as Spain. In the following, we will devote some explanations to these aspects.

### *1.1. Open Government and Open Government Data*

"Open Government" is an innovative philosophy of policy and application of avant-garde work methodologies in public administration that has spread in recent years in democratic countries [14–16]. It is based on transparency, accountability, citizen participation and collaboration, and it is supported and possible because of information and communications technologies and the Internet as established in all areas since the 2000s [17–19]. In these open government spaces, work is done by and for citizens, improving their trust and contributing value to society [20]. This concept gained worldwide visibility when, in January 2009, Barack Obama was elected the 44th president of the United States of America and signed the Memorandum on Transparency and Open Government [21–23].

Open government policies cannot be understood without the release of certain datasets generated by administrations and governments. Indeed, data is an essential component that must be successfully managed to achieve the aspiration of Open Government, requiring advanced technical tools at considerable expense, often becoming a significant investment for public administrations. This circumstance gives rise to what we know as Open Data [16,24]. These data can be defined as government or administration data, freely available in data portals, without restrictions, at no cost and in formats that facilitate their treatment and reuse [25–28].

Data opening is justified for several reasons. First is transparency. For a democratic society to function well, citizens and other stakeholders must be able to monitor open government initiatives and their legitimacy. This means that the citizenry can access certain datasets and, in addition, use, reuse and distribute them. The success of transparency translates into a capacity for social control [25]. However, the success of the use of open data lies to a large extent in the possibility of reusing this data. Otherwise, this means a huge effort of public administration without social or economic return, which generally leads to the failure of public initiatives [29]. This is why there are many proposals for new ways to facilitate the reuse of open data, promoting training [11] and creativity to come up with innovative reuse proposals.

Another element of interest to highlight is that data opening allows the liberation of social and commercial value. Governments are one of the largest producers and collectors of data in many areas [30]. All data, of any kind, such as geospatial, education, environmen-

tal, transport, planning or budgetary, have a social and commercial value and, therefore, can be used for purposes other than those originally intended. With the availability of different datasets, society can carry out new developments and innovate from them creating new services and research that provide added value [25]. Thus, these data from open government initiatives generate economic and social value, becoming a focus of interest for the public sector and business-people [31]. Among benefits are the improvement in operational capacity, government efficiency and cost reduction in administration, democracy, in addition to providing innovative practices, producing data-driven value and improving, as already noted, transparency, participation and citizen collaboration [32–37].

However, despite the growing interest in data, there is a lack of experimental approaches to actually measuring the impact of open government data based on surveys and field studies [38], which are necessary to justify and consolidate the open data policies [29].

Open data can also play a crucial role in training the next generation of researchers. Access to general socially and scientifically relevant data, together with open research data, is vital to enhance their professional development. In today's world, practitioners and scientists alike rely on internet platforms to access a wide range of data from almost any source, directly from their desktops [39]. It is known that research data (and in general, by extension, open data) appear to be contextual and highly dependent on different communities and professional traditions [39,40]. Not only the data, but also the metadata involved, depend on the habits and customs of the data generators, be it a specific scientific or professional group.

The factors influencing open data publishing practices are very diverse and extremely relevant to the success of open data reuse. The training of new professionals and researchers, and previously of their teachers, is also essential, as well as a fluid dialogue between the different groups that manage and use open data [9,41,42].

### *1.2. The Situation of Open Data in Spain*

Spain is one of the countries with more open data initiatives at European and international level, and in recent years has experienced an increase in the volume of services that offer public data for reuse [43]. Numerous (international and European) entities in charge of overseeing and promoting the development of open data and reuse projects analyze these initiatives from different perspectives to produce country rankings on the level of openness, data quality, policies and impact, among other indicators. For example, worldwide, Spain has moved from 11th place in the 4th edition of the Open Data Barometer to 21st place in its latest Open Data Watch 2020 report. In the annual report presented by the [44,45] European Data Portal, the [46] Open Data Maturity assessment, in the 2021 ranking, Spain occupies the third position with a score of 95%, in a ranking led by France with 98%. In the statistics of the European Data Portal as of today (August 2024) Spain has contributed 115,835 datasets, being the fifth largest contributor behind Germany, Czechia, France and Italy. According to these studies, Spain is in a good position in terms of openness and availability of data from open government initiatives and administrations. Other research focuses on measuring the use, reuse, maturity, and quality of datasets [43,47] in assessing the economic impact of open data [48].

However, as far as we know, no specific works have been published that study the perception of citizens regarding open data in Spain, although some related research can be found in the literature today [11]. Some publications can be found on contextualized research in countries such as Mexico [49] or the United Kingdom [50]. There are also some works on specific datasets with diverse themes, such as studies on the quality and level of reuse of environmental data types [51]. In recent times, the problem of reuse of open governmental data has become an important issue, due to the fact that benefit needs to be obtained from these expensive innovative investment [4]. In Spain, where this research has been conducted, there are studies that measure quality, reuse levels, policies, and other indicators related to open data. However, there is a lack of research focusing on

the scientific community's perceptions of open data and its reuse within academic and scientific contexts.

In principle, the considerable success of open data policies in the country would suggest that knowledge and use of open data is widespread in academic institutions. In this direction, this paper aims to analyze the actual knowledge of researchers and professors working in universities and research centers regarding open data, and to find out the level of reuse and usefulness of open data for this group.

### *1.3. Open Data in Academia*

Open data can be a powerful tool in teaching practice, which makes it especially relevant for university teachers to be aware of what this new trend in information access can offer, as well as how to access this new source. Some recent classroom experiences [52] show to what extent a new way of teaching can be addressed as a consequence of the use of data in the classroom, especially when the teaching plan is designed to use innovative techniques (e.g., classroom research activities), or simply to increase the quality of the teaching task by improving motivation, social interaction and personalization, for example. Low awareness of open data policies and tools among academics are obvious barriers to the implementation and standardization of data use in education [53]. These problems must be overcome in order to meet the dual objective of updating university education on data: the use of data to improve the scientific and technological outputs resulting from academic activity [54–57], and teaching on the subject of open data itself, a fundamental tool for the future professional career of our graduates [58]. Furthermore, collaboration from the professional and academic world to improve knowledge about what kind of data they are actually using could help improve this education [58,59].

The extensive use of open data as educational resources has to be also critically considered [11]. We have recent examples of how the adoption of digital resources can lead to unsuspected and undesirable consequences in higher education. For example, after COVID-19, all universities wholeheartedly adopted digital platforms, often leading to a distortion of academic activity [60]. The context of COVID-19 led to the sudden commercialization and privatization of key aspects of higher education, with industries providing digital communication tools, introducing (based on purely commercial parameters) their instruments for online learning and the analysis of academic productivity, in all its aspects [61]. This is just one example of how external factors influence educational practice and how the increase in technological tools allows for an improvement in education. However, often the changes thus brought about do not lead to improvement. The emergence of Generative Artificial Intelligence with Chat GPT in education could be the latest of these problematic developments, although the consequences of its uncontrolled use by students have not yet been accurately assessed [62]. In our case, the general question of how appropriate and useful is the use of open government data in academia is an open one, and one way to start answering this question is to obtain an overview of the actual situation through field studies, such as the one presented here. Although, in principle, it seems a positive new tool to introduce in the academic world, sometimes the actual conditions in which actions are developed affect and undermine potentially beneficial practices.

Summarizing all the comments we have presented in this introductory part, the aim of this paper is to analyze in a real case study to what extent open government data are used by academics to improve their main activities: teaching and research. To do so, it is necessary to assess several aspects related to these general issues, such as the interest in and training of academic staff in open government data, the assessment of the quality of the data they use, and the access they have to technical tools for analysis, as well as the study of how reuse is spread among the academic community depending on the scientific area to which they belong. The results of the statistical analysis of the survey carried out to clarify these issues, in the case of the Spanish university system, will be explained in the following sections.

## 2. Material and Methods

To approach this study and gain knowledge regarding the situation of open data and its use by Spanish scientists, a quantitative research technique was used. A survey was developed and, with the feedback obtained, the research questions posed were answered, which allows a detailed analysis of the current situation. In the analytical data section, a quantitative approach was utilized to evaluate the 783 survey responses collected from Spanish scientists. Given the total population of approximately 140,000 researchers, teachers and collaborators (a very high estimate), this sample size represents a small but significant subset of the scientific community. To address possible doubts about the representativeness of the sample, we carefully assessed its stratification to ensure that it adequately reflected the general population. We took into account key factors, such as employment status, academic discipline, and institutional affiliation, to verify that all relevant groups were well represented. Although the response rate was relatively low, a detailed analysis confirmed that there was no evidence of bias in the sample and that the results could be reasonably extended to the entire population of Spanish scientists.

Indeed, considering the main grouping criteria, we observed that there was significant representation from various Spanish regions and universities, and the top five were Valencia, Madrid, Alicante, Granada, and Cádiz, each comprising more than 4% of the sample. These regions adequately reflect the structure of the Spanish academic landscape: two large regions with high university populations, two smaller regions with smaller overall population but strong academic activity and tradition (such as Granada), and one relatively small region and university, as represented by Cádiz. The most widely represented universities were the two major ones in Valencia, the Complutense University of Madrid, and the University of Granada, which are among the most important institutions in the Spanish university system in terms of the number of professors and researchers. Regarding gender, the distribution was 56.58% men and 41.76% women, falling within the usual range considered by Spanish administrative standards for gender equality in representation. Age was also well represented, with over 60% of respondents falling within the 40 to 60 age range, which reflects the aging trend among Spanish university staff. Additionally, the distribution by contract type was acceptable, with more than 55% of responses coming from Professors (Profesores Titulares), Full Professors (Catedráticos), and Associate Professors (Profesor Contratado Doctor), which are the most common positions within Spanish university faculties. Finally, the distribution by scientific fields was also considered acceptable, with all major areas of science (social sciences, engineering, information and communication, health, culture and education, formal sciences, natural and earth sciences and technologies) each comprising more than 4% of the sample.

### *2.1. Population Studied*

The Spanish University System is made up of 83 universities, 50 public and 33 private, which group together 1061 university centres between schools and faculties, 537 research university institutes, 50 doctoral schools, 54 university hospitals and 76 foundations. As for the teaching and research staff (PDI per its Spanish initials), its number stands at 125,471 people in 2021, of which 105,371 are working in public universities and 20,100 working in private universities. The number of research staff (PI per its Spanish initials) and technical research support staff is 26,408, of whom 24,886 are ascribed to public universities and 1522 to private universities. Finally, the employed research staff amounts to 19,879 contracted researchers (Ministry of Universities 2021) [63].

The survey was distributed via email using the Survey Monkey survey platform. The email addresses of the surveyed sample were drawn from the mailing directories on the university websites, with a total of 22,791 addresses extracted from the websites of 47 universities and 21 research centers. The survey distribution period was between 15 March and 10 May 2021. A total of 783 responses was obtained. For a population of 140,000 researchers and professors (estimating upwards, including support staff and collaborators), a sample size of approximately 384 responses would be sufficient to reach

a confidence level of 95% with a margin of error of ±5%. Given that the survey collected 783 responses, the sample size is adequate. This larger sample size can potentially reduce the margin of error, increasing the accuracy of the conclusions.

For the development of this study, directed or intentional sampling is used. This is a biased sampling that seeks to obtain the representativeness of the selected elements by trying to reach a specialized population. This type of sampling is opportunistic, and it is advisable to carry it out in piloting or probing studies like this.

### *2.2. The Survey*

The questionnaire has a total of 30 questions with the following distribution: 6 questions in an initial context block and 24 questions distributed in 6 technical blocks. Some had the option of open response so that respondents could provide comments, and were not mandatory, so as not to condition the researcher.

The blocks and questions were designed following the research team's previous knowledge of Open Data management in Spanish administrations and how universities are organized [20]. The blocks include questions ranging from contextual information and basic knowledge about Open Data on the part of teachers and researchers, and progressively increasing in specificity to more detailed aspects about crucial concepts (data reuse, knowledge of tools for data management, access and quality of datasets). The motivational aspects as to why researchers are interested in using open government data is also a major topic we were interested to know about. The blocks are as follows (Table 1). We will discuss the specific questions asked in each of them below.

**Table 1.** Survey Blocks [1].

| Block | Title |
|---|---|
| Block 0. | Personal and work information. |
| Block 1. | Initial context: knowledge of open data. |
| Block 2. | Openness of data |
| Block 3. | Real usefulness and motivations for reusing open data. |
| Block 4. | Access to datasets. |
| Block 5. | Use and reuse of data (I). Treatment and reuse. |
| Block 6. | Use and reuse of data (II). Specific use of open data. |

[1] Source: all tables are our own elaboration.

## 3. Results

There is representation from 34 of the 50 Spanish provinces and the two autonomous cities (Table 2). However, some of the main Spanish universities have not participated in the survey (such as the University of Barcelona) due to some restrictions imposed by these centres on the distribution of surveys among staff. This will be discussed further in the comments related to respondents' university affiliation.

In terms of gender, of the 783 responses, 56.58% were answered by men and 41.76% were answered by women. These data show gender parity in the study (Table 3), taking into account that, in general, there is still a lack of parity in the staff of Spanish universities, especially at the highest levels of the structure (Catedráticos y Profesores Titulares).

**Table 2.** Province of residence of the respondent.

| **Province of Residence** | **%** |
|---|---|
| Valencia | 31.29 |
| Madrid | 14.94 |
| Alicante | 9.96 |
| Granada | 8.56 |
| Cádiz | 4.21 |
| Navarra | 3.70 |
| Jaén | 3.45 |
| Salamanca | 3.07 |
| Barcelona | 2.94 |
| Castellón | 2.17 |
| Murcia | 2.17 |
| Zaragoza | 1.66 |
| Vizcaya | 1.28 |
| Sevilla | 1.15 |
| Las Palmas | 0.89 |
| Málaga | 0.64 |
| Asturias; Lleida | 0.51 |
| Álava; Almería; Burgos; Huelva; Melilla | 0.38 |
| Ávila; Badajoz; Ceuta; Ciudad Real; Cuenca; Gipúzkoa; Valladolid; Zamora | 0.26 |
| A Coruña; Baleares; Girona; Guadalajara; Santa Cruz de Tenerife | 0.13 |
| Total | 97.70 |
| Omitted | 1.40 |
| Discarded [1] | 0.89 |

[1] All responses that are not from Spanish provinces or autonomous cities are discarded for the analysis of provinces.

**Table 3.** Representation by gender.

| **Gender** | **%** |
|---|---|
| Female | 41.76 |
| Male | 56.58 |
| Responded | 98.34 |
| Omitted | 1.66 |

The responses by age group were as follows (Table 4):

**Table 4.** Age ranges of those surveyed.

| **Age Range** | **%** |
|---|---|
| 22–30 | 8.68 |
| 31–40 | 16.48 |
| 41–50 | 31.80 |
| 51–60 | 30.14 |
| 61–65 | 8.43 |
| Over 65 | 3.58 |
| Responses | 99.11 |
| Omitted | 0.89 |

Responses were obtained from members of 47 universities and 21 centres, foundations and research institutes. Table 5 shows the universities to which the academicians are affiliated, and Table 6 shows the responses of scientists from centres and other research institutes. Centres with fewer than 1% of responses were not included in the table:

**Table 5.** Universities to which the respondents are affiliated.

| University | % |
|---|---|
| Universitat Politècnica de València | 16.86 |
| Universidad de Granada | 9.32 |
| Universidad Complutense de Madrid | 7.28 |
| Universitat de València | 6.77 |
| Universidad de Cádiz | 4.73 |
| Universidad Carlos III de Madrid | 3.70 |
| Universidad Pública de Navarra | 3.70 |
| Universidad de Jaén | 3.58 |
| Universidad Miguel Hernández | 3.45 |
| Universidad de Salamanca | 3.32 |
| Universidad de Alicante | 3.19 |
| Universidad Autónoma de Madrid | 1.66 |
| Universidad del País Vasco | 1.66 |
| Universidad de Zaragoza | 1.53 |
| Universitat Jaume I | 1.53 |
| Universidad de Murcia | 1.28 |
| Universitat Autònoma de Barcelona | 1.28 |
| Universidad Politécnica de Madrid | 1.02 |
| Total researchers Spanish universities | 86.08 |

**Table 6.** Research centres to which the respondents are affiliated.

| Centre, Institute, Research Foundation | % |
|---|---|
| Instituto de Investigación Sanitaria La Fe | 1.40 |
| Fundación Centro de Estudios Ambientales del Mediterráneo (CEAM) | 1.28 |
| Instituto Valenciano de Investigaciones Agrarias (IVIA) | 1.15 |
| ISABIAL—Instituto de Investigación Sanitaria y Biomédica de Alicante | 1.02 |
| Instituto de Agroquímica y Tecnología de Alimentos (IATA-CSIC) | 0.89 |
| Centro de Investigación Príncipe Felipe (CIPF) | 0.77 |
| Instituto de Física Corpuscular (IFIC-CSIC) | 0.77 |
| Consejo Superior de Investigaciones Científicas (CSIC) | 0.64 |
| Hospital General Universitario de Alicante | 0.38 |
| Fundació de la Comunitat Valenciana (FCV) | 0.26 |
| Instituto Tecnológico de la Alimentación (AINIA) | 0.26 |
| Instituto Tecnológico de la Energía (ITE) | 0.26 |
| Agencia de Medios | 0.13 |
| CERN—Organización Europea para la Investigación Nuclear | 0.13 |
| CIC bioGUNE—Centro de Investigación Cooperativa en Biociencias | 0.13 |
| Fundació Docència i Recerca Mútua Terrassa | 0.13 |
| Hospital del Mar—Valencia | 0.13 |
| IES Haygón | 0.13 |
| Instituto de Investigación en Inteligencia Artificial (IIIA-CSIC) | 0.13 |
| Instituto LRKlein | 0.13 |
| Instituto Universitario Menéndez Pidal (UCM) | 0.13 |
| Total researchers from research centres and institutes | 10.25 |

Some universities are over-represented in the Spanish university map, while others hardly appear or do not appear at all (for example, the University of Granada compared to the University of Barcelona, which did not take part in the survey). This was a natural consequence of the restrictions imposed by some of the centres in terms of carrying out the survey, and the facilities offered by other centres to carry out this work. However, we do not consider this bias to be a real problem for the generality of the conclusions, since in general Spanish public universities (the most relevant part of the Spanish university system and of this survey), are quite similar in terms of structure, operation and quality.

The labor categories were established according to the professional scales of the Spanish university system. In the Spanish university, there are two main categories of

professors. The first are civil servants, who are at the top of the system, and belong to several levels, which are (ordered from the top), Catedrático (full professor) and Profesor Titular (associate professor), Profesor Agregado, Catedrático de Escuela Universitaria and Profesor Titular de Escuela Universitaria. Below this scale, there are other positions that are filled through contracts, and which may also be permanent professors. They are the following (from the top): Contratado Doctor, Ayudante Doctor and Profesor Colaborador. There are also external positions, that are contracted to cover some hours of teaching, called Profesor Asociado, Professor Visitante and Profesor Sustituto, Interino (Table 7). There are also professionals who belong to public research institutions (from the top): Profesor de Investigación and Científico Titular. The rest of the professionals listed in the following table hold similar positions to those that exist in most countries, and can be directly translated to the terms appearing in this table.

**Table 7.** Job category of respondents.

| **Job Category** | **%** |
|---|---|
| Profesor Titular de Universidad | 26.69 |
| Catedrático/a de Universidad | 19.16 |
| Profesor/a Contratado/a Doctor | 9.71 |
| Doctorando/a (PhD Student) | 7.92 |
| Profesor/a Asociado/a | 7.79 |
| Profesor/a Ayudante Doctor | 6.39 |
| Investigador/a Científico | 6.39 |
| Doctor/a | 3.32 |
| Técnico de investigación (Research Technician) | 2.81 |
| Profesor/a de Investigación | 1.28 |
| Profesor/a Sustituto/a Interino/a | 1.28 |
| Personal de Administración y Servicios [1] (administrative) | 1.15 |
| Científico/a Titular | 1.02 |
| Profesor/a Ayudante | 0.89 |
| Profesor/a visitante | 0.77 |
| Empresa [2] (Private Company worker) | 0.51 |
| Profesor/a Colaborador/a | 0.38 |
| Profesor/a Agregado/a | 0.26 |
| Profesor/a Titular de Escuela Universitaria | 0.26 |
| Catedrático/a de Escuela Universitaria | 0.13 |
| CIO (Private Company worker) | 0.13 |
| Formación Profesorado Universitario | 0.13 |
| PCI (Private Company worker) | 0.13 |
| Profesor/a Sustituto | 0.13 |
| Responses | 98.60 |
| Omitted | 1.40 |

[1] Non-teaching personnel: Librarians, documentalists, assistant technicians. [2] Company: this includes journalists.

After data standardization, the results obtained were as follows:

The distribution of the teaching staff responding to the survey follows more or less the same pattern as that of Spanish universities. The main group is made up of full professors (Catedrático de Universidad) and associate professors (Profesor Titular), who define the main group of lecturers with permanent positions. Together with the rest of the catalogued teaching positions (Contratado Doctor, Asociado, Ayudante Doctor), they make up the main structure of teaching and research positions in the universities. According to the latest official statistics available on the website of the Ministerio de Ciencia, Innovación y Universidades, 58.1% of lecturers have a permanent post. This proportion is similar to that found in the sample, so we can reasonably consider that there is no bias concerning this particular point.

The research areas were organized according to the categorization of the State Research Agency (AEI, per its Spanish initials, Table 8) of the Ministry of Science and Innovation [1,2].

Although there are some that are not very well represented, overall the distribution of responses (and thus the interest of participants) can be considered adequate. Unlike the case of Open Research Data, it is worth noting that, due to the social and economic relevance of government Open Data, social sciences (including economics and others) are represented at least at the same level as STEM fields.

**Table 8.** Research areas.

| Research Area | % |
|---|---|
| Area 1: CSO/Social Sciences | 16.60 |
| Area 11: PIN/Industrial production, civil engineering and engineering for society | 8.05 |
| Area 12: ICT/Information and Communication Technologies | 8.05 |
| Area 3: ECO/Economics | 7.79 |
| Area 19: BME/Biomedicine | 7.54 |
| Area 5: FLA/Culture: Philosophy, Literature and Art | 5.36 |
| Area 9: MTM/Mathematical Sciences | 4.98 |
| Area 7: EDU/Education Sciences | 4.85 |
| Area 16: CTM/Environmental Sciences and Technologies | 4.73 |
| Area 17: AAC/Agricultural and agri-food sciences | 4.60 |
| Area 18: BIO/Biosciences and biotechnology | 4.60 |
| Area 2: DER/Law | 4.21 |
| Area 10: FIS/Physical Sciences | 4.09 |
| Area 14: CTQ/Chemical Sciences and Technologies | 3.96 |
| Area 8: PSI/Psychology | 2.55 |
| Area 15: MAT/Materials Science and Technology | 2.17 |
| Area 6: PHA/Past Studies: History and Archaeology | 2.04 |
| Area 4: MLP/Mind, language and thought | 1.79 |
| Area 13: EYT/Energy and Transport | 1.02 |
| Responses | 98.98 |
| Omitted | 1.02 |

Research areas of the State Research Agency of the Ministry of Science and Innovation: https://www.aei.gob.es/areas-tematicas/areas-tematicas (accessed on 7 October 2020); The initials in this table correspond to standard abbreviations used in Spanish by the SRA/AEI in their categorisation of research areas.

### *3.1. Knowledge and Reuse of Open Data*

The percentage of researchers who claimed to know what open data are is relevant, 79.31% compared to 7.28% who said they did not know.

Among those who claimed to know about open data, 50.57% have reused open data in their research compared to 34.99% who have not (Table 9).

**Table 9.** Reuse of open data in research and projects.

| Response | % |
|---|---|
| Yes | 50.57 |
| No | 34.99 |
| Omitted | 14.43 |

The reasons pointed out for not knowing were varied (Table 10). The five main reasons were:

**Table 10.** Reason for lack of knowledge of open data.

| Options | Responses |
|---|---|
| Lack of knowledge of sources | 10 |
| Lack of knowledge of use and potential for reuse | 34 |
| Lack of dissemination and active publicity | 15 |
| Lack of information | 14 |
| No need | 5 |
| Total of responses | 78 |

The main reasons pointed out for not reusing the data were the following (Table 11):

**Table 11.** Main reasons for non-reuse of open data.

| Options | Responses |
|---|---|
| Lack of data of interest | 138 |
| Lack of knowledge of reuse | 94 |
| Lack of standardisation | 82 |
| Lack of data reliability | 35 |
| Lack of data updating | 34 |

These reasons were extracted from one of the open questions and the answers were grouped by similarity of content.

Regarding the attributes that open data should have, eight types were proposed and rated on a scale of 0 to 4 (0 not knowing and 4 attributes of great importance). The "Mean" data are the result of calculating the weighted mean, after transforming the categorical scale to a quantitative scale, in which the number of individuals in each group is multiplied by their level of importance (I do not know = 0, No importance = 1, Little importance = 2, Important = 3, Very important = 4), then divided by the total number of individuals (Table 12).

**Table 12.** Consideration of importance of attributes in open data [1].

| Attribute | No Importance | Little Importance | Important | Very Important | I Do Not Know |
|---|---|---|---|---|---|
| Provision | 0.83 | 4.15 | 30.71 | 62.66 | 1.65 |
| Data licencing | 1.87 | 7.28 | 31.81 | 54.47 | 4.57 |
| Metadata | 0.83 | 3.52 | 22.15 | 68.74 | 4.76 |
| Accessibility | 0.62 | 1.24 | 17.77 | 76.24 | 4.13 |
| Formats | 1.04 | 4.35 | 26.92 | 63.35 | 4.35 |
| Restrictions | 2.71 | 14.55 | 42.41 | 35.76 | 4.57 |
| Updating | 3.86 | 24.52 | 34.84 | 33.76 | 3.01 |
| Geolocation | 5.23 | 20.71 | 32.01 | 34.94 | 7.11 |

[1] The table shows the consideration of open data attributes in percentages.

The researchers valued accessibility as one of the most important attributes, i.e., the fact that the data are easy to find and obtain, as the International Open Data Charter from the Open Data Charter points out. Second, in terms of importance, they point to the quality of the metadata. The fact that they contain complete, useful, and structured metadata is of vital importance for their interpretation and reuse. Thirdly, they value the formats, i.e., open and machine-accessible formats, as a vital feature for their reuse and in order to establish automatic downloading processes in order to be able to apply artificial intelligence processes when the study requires it. These three most highly valued elements should be taken into account by the administrations in charge of opening data. They are also noted by the various organizations that promote open data: Open Data Charter, the European Union, the OECD or the Open Knowledge Foundation, among others. Thus, the data must be easily accessible and traceable, available without restrictions (without the need for payment or registration to access them) and the metadata must be complete and structured to facilitate their interpretation. It is noteworthy that the fact that they are up to date is not considered too important: 24.52% of respondents thought that this is not important, the maximum rate among all the characteristics studied.

As to the services and functionalities they deem relevant in accessing the data for reuse, the fact that the portals include the number of views of the data is the most valued factor, as can be seen in Table 13:

**Table 13.** Assessment of services and functionalities of access to open data.

| Functionalities | No Importance | Low Importance | Important | Very Important | I Do Not Know |
|---|---|---|---|---|---|
| URI | 1.47 | 4.42 | 38.95 | 33.27 | 21.89 |
| Related apps | 2.11 | 21.89 | 42.11 | 26.31 | 7.58 |
| Processing tools | 2.51 | 18.41 | 40.38 | 35.98 | 2.72 |
| Views | 1.04 | 9.81 | 40.50 | 46.35 | 2.30 |
| Contact | 1.47 | 16.00 | 50.32 | 24.63 | 7.58 |
| News | 3.78 | 27.73 | 46.01 | 18.07 | 4.41 |
| API | 2.13 | 15.57 | 28.57 | 19.40 | 34.33 |

Secondly, users considered it important that there are examples or links to tools for the processing and analysis of data, and, thirdly, they considered it relevant that there are channels of communication or contact for those reusers who have difficulties or doubts when reusing public sector information. It stands out that respondents considered an unimportant feature the fact that the portals have links to applications related to the datasets that serve as a guide or example. Perhaps this is because they are not the developers of Apps for customers with a commercial purpose, or for their use.

When analyzing the barriers to access and reuse that researchers encounter when working with data, one of the most common is the lack of data quality and standardization, followed by interoperability, access difficulties and lack of up-to-dateness (Table 14). This contradicts one of the previous answers, where no importance was given to updating.

**Table 14.** Assessment of barriers to access and reuse of open data.

| Attribute | Always | Many Times | At Times | Never | I Do Not Know |
|---|---|---|---|---|---|
| Time | 5.00 | 15.50 | 43.00 | 11.75 | 24.75 |
| Registration | 9.70 | 31.59 | 37.31 | 8.22 | 13.18 |
| Comprehension | 8.00 | 27.50 | 41.75 | 11.00 | 11.75 |
| Quality and standardisation | 12.66 | 38.46 | 29.28 | 6.45 | 13.15 |
| Payment | 2.83 | 13.88 | 37.28 | 26.22 | 19.79 |
| Interoperability | 6.89 | 31.38 | 30.61 | 8.42 | 22.70 |
| Lack of updating | 6.00 | 33.25 | 35.75 | 11.25 | 13.75 |
| Access | 5.13 | 33.85 | 42.05 | 7.69 | 11.28 |

In relation to the challenges faced by governments in terms of data openness, 69.12% considered that the most relevant is that of standardization, followed by interoperability among administrations with 57.11% and lack of training with 46.08%. The latter is one of the key elements that were pointed out by different groups in different studies [43,51]. The technical issues, while important, do not represent more than 44.61%, since the technology for starting is already fully implemented and there are different tools for free and commercial access that make its application possible. To a lesser extent, issues of financing were pointed out, with 34.07%, and active publicity by the administrations, with 26.96% (Table 15).

**Table 15.** Challenges of open data in government.

| | % |
|---|---|
| Standardisation of data | 69.12 |
| Administrative Interoperability | 57.11 |
| Lack of training | 46.08 |
| Technical issues | 44.61 |
| Funding | 34.07 |
| Active publicity | 26.96 |
| Other (specify) | 10.78 |

### *3.2. Typology and Most Used Formats*

The typology of the most used data is shown in Table 16. Science and technology data and public sector data stand out. Given the target audience of this study, it is relatively normal that these are the most used data categories due to their subject matter and interest. The researchers are mostly members of the administration, and it is well-known that one of the main reusers is the administration itself.

**Table 16.** Most relevant datasets for researchers (+100 choices).

| Dataset | No. Choices |
|---|---|
| Science and Technology | 220 |
| Public Sector | 199 |
| Education | 170 |
| Environment and climate | 160 |
| Health | 151 |
| Demography and population | 139 |
| Economy | 129 |
| Society and Welfare | 111 |

In terms of data use, the majority (67.18%) have used data for their scientific or academic work, although it should also be noted that many (68.48%) have downloaded data which remained unused (Table 17). This may have its origin, as mentioned above in other questions, lack of quality and updating as well as interoperability and metadata quality. A total of 44.96% indicated the use of statistical data. The fact that this type of data has been reported as the most widely used makes us wonder whether respondents really know what open data is since, in most statistical portals in the country data is accessible, but not in formats that can be considered open. It also stands out that 35.92% indicated that, while not having used this type of data, they do plan to do so in the future. We attribute this to the fact that more and more staff are more familiar with this type of information and know how to access this type of portal.

**Table 17.** Purpose of reuse of open data.

| Item | Percentage % |
|---|---|
| I have downloaded open datasets | 68.48 |
| I have used open data for scientific research (articles, reports, presentations) | 67.18 |
| I have used open data for statistical analyses | 44.96 |
| I plan to make more use of open data | 35.92 |
| I have used open data to consult public information of interest | 28.17 |
| I have used open data to create visuals | 21.19 |
| I have used open data to create a website | 19.38 |
| I have used open data to create location maps | 16.54 |
| I have used open data to divulge public information to individuals | 13.95 |
| I have used open data to develop a web service or API | 9.30 |
| I have never used open data | 7.75 |
| I have integrated governmental open data with an application or existing website | 5.94 |
| I have used open data to develop a mobile application | 4.39 |
| I have used governmental open data to carry out vindictive campaigns | 2.58 |
| Other (specify) | 1.29 |

As already mentioned, the projects carried out for the preparation of scientific publications, academic works, research projects and studies and reports stand out. As expected, what open data have been used the least for is in the development of applications and websites. The data can be seen in Table 18.

The technical formats that are most reused are GeoJSON (47.17%) and JSON (42.75%), which indicates a broad knowledge of the most appropriate formats as observed in the gradation in the answers noted as advanced reuse. In the same way, the PDF was pointed out as the least reusable (Table 19).

**Table 18.** Specific use of open data.

| Type of Project | No. Choices |
|---|---|
| Scientific publications and academic work | 107 |
| Research projects | 50 |
| Studies and reports | 32 |
| Research and statistical projects | 15 |
| Teaching | 12 |
| Creation of visuals | 6 |
| COVID-19 | 5 |
| Various | 4 |
| App and web development | 3 |
| Communication and data journalism | 3 |
| Consultation | 2 |

**Table 19.** Assessment of technical formats.

| Technical Format | Inadequate Format (0) | Basic Reuse (1) | Improvable Advanced Reuse (2) | Advanced Reuse (3) |
|---|---|---|---|---|
| CSV | 2.76 | 33.45 | 28.62 | 35.17 |
| XLS | 4.36 | 33.89 | 32.89 | 28.86 |
| PDF | 49.35 | 31.82 | 10.39 | 8.44 |
| XML | 6.83 | 31.22 | 38.05 | 23.90 |
| JSON | 6.87 | 15.27 | 35.11 | 42.75 |
| GeoJSON | 7.55 | 16.98 | 28.30 | 47.17 |
| RDF | 8.33 | 28.58 | 29.76 | 33.33 |
| KML | 8.79 | 25.27 | 36.27 | 29.67 |
| SQL | 10.16 | 22.66 | 33.59 | 33.59 |
| RSS | 16.83 | 37.63 | 28.71 | 16.83 |
| WMS | 13.25 | 28.92 | 34.94 | 22.89 |
| WFS | 11.69 | 29.87 | 32.47 | 25.97 |
| GML | 12.66 | 26.58 | 36.71 | 24.05 |
| SHP | 9.00 | 23.00 | 33.00 | 35.00 |

## 4. Discussion

As mentioned in the Introduction, discussions of open data in science typically refer to open research data, including whether scientists share their data, how journals handle this issue (whether they encourage or require data deposition), and the role of repositories that provide a DOI for datasets upon deposit [39,64–67]. The publication of research data is fundamental to scientific activity, as this practice allows for the reproducibility of experimentation. However, this practice is dependent on many contextual factors [42,68–70]. Using open data from government sources for scientific research presents a distinct challenge, though it shares some similarities with the publication and reuse of research data. To make data reuse a common practice among researchers (a goal that would enhance the social benefits of open data policies and applied research), it is necessary to address certain aspects of research practices as well as the technical conditions surrounding data publication. On the other hand, open data have been shown to be an important tool for university teaching, but lack of knowledge of open data management techniques is a major barrier for academics in the use of open data in education [53].

In general, communication between the entities that carry out the opening of data and the citizenry is necessary [8,11]. On the one hand, administrations must satisfy the

needs of the citizen who seeks to know the datasets that are of relevance and interest for the development of innovative research and services. In turn, they must offer data that the population may not perceive as necessary but that, by consultation or navigation through open data portals, generate ideas that give way to new developments. It is not just a matter of carrying out open data initiatives, but of taking into account what citizens need, what their perception is, whether public information sources are relevant to their studies, and the technical and content requirements that data must have in order to be reusable. If we look at the scientific literature, there are many studies that evaluate in different ways the quality of open datasets and their level of reuse based on indicators [43], the info-media sector and the value generated by the data [48], studies of specific datasets applied to a certain area [51], or the perception that citizens in general have about open data [49,50]. In 2020, a survey was carried out within the framework of the Datasea project related to research data, with the aim of finding out the habits and experiences of Spanish researchers in health sciences in relation to the management and exchange of raw research data [71]. These studies, which are interesting and can be very useful, are complemented by a series of general perspectives that allow a panoramic view of specific open data situations. In some cases, these studies involve members of the scientific community, but they are not research methodologies directly focused on the academic community (with the exception of the last one cited by [71]. This paper aims to provide such information regarding the academic community. The results obtained from the analysis of our survey can be used to support the actions of the institutions in charge of data openness, since it presents opinions, patterns of behavior and the perception of the Spanish academic community on the subject.

Let us now analyze the information provided by the survey point by point, using the outline provided by the blocks into which it was organized. We will analyze the main characteristics of the academics participating in the survey in terms of general data. Regarding gender, in this research we take into account the [72] Organic Law 3/2007, of 22 March, on the effective equality of women and men. Spain has legal measures that guarantee the representation of women in institutions, with a commitment to parity democracy, i.e., with a balanced representation of men and women. This translates into a minimum representation percentage of 40% for both sexes [73].

Among the 621 researchers who report being familiar with open data, 362 are men (58.29%) and 256 are women (41.22%), reflecting a nearly representative gender balance. When comparing with other studies, that which most closely aligns with these parity criteria in responses is another quantitative research project on research data in Spain, in which 62.4% of the answers came from men (663) and 37.6% of the answers obtained are due to women (400) [71]. In other similar research works on the use and reuse of open data, the difference is remarkable. For example, in the study carried out in 2010 by Tim Davies on the use of open data in Great Britain, who used quantitative techniques, the rate of responses to the survey showed that, for every seven responses obtained, six are from men and one from women [50]. Other studies, such as a Belmont Forum survey of a community of environmentally related open data reusers, or a study conducted in Mexico, made no reference to gender distinction in their responses [51; 49]. The EPData [74] platform offers data from the Ministry of Science, Innovation, and Universities, including a visualization that tracks the evolution of university teaching and research staff by gender between 2012 and 2018. It can be observed there that the Spanish teaching and research staff is made up of 51,331 women (41.76%) and 71,579 men (58.24%), which roughly matches the response rate obtained in this survey. Thus, we can conclude that, in terms of gender, the results obtained are representative

Among the research areas, the most chosen by far was area 1 Social Sciences (130). With a similar percentage of responses, we find area 11, which corresponds to Industrial production, civil engineering and engineering for society (63), area 12, Information and communication technologies (63), area 3, Economics (61) and area 19, Biomedicine (59). With a smaller volume, responses were obtained from the other 14 research areas. The fact that the responses came from researchers from different areas enriches the study and

allows us to obtain a multidisciplinary view of the knowledge, use and reuse of open data in science. Proportionally, there is a difference between the response rate of researchers from the social sciences and those from the classical sciences. The under-representation of scientists from classical sciences, such as biology, physics and chemistry, is noteworthy as these are disciplines where experimental verification is systematically used as part of the application of the scientific method. In our opinion, the main reason for non-reuse may be associated with the fact that these scientists are used to using data generated by themselves. Therefore, it would be of interest to conduct information campaigns on open government data for this group, as a great deal of data can undoubtedly be of great interest for research in their areas as well.

When analyzing the information on researchers' level of familiarity with open data and their areas of research, it is evident that the most selected area is Social Sciences (Area 1) with 104 respondents, representing 80% of researchers who are aware of open data. This is understandable, given the large number of datasets available on open data portals that are reusable in the context of many social science disciplines. Other areas, such as Information and Communication Technologies (Area 12), Economics (Area 3) and Biomedicine (Area 19), are also disciplines where, regardless of specific research applications, there is an awareness of these information sources by researchers, recognizing their potential for reuse. In view of the percentages reflecting the lack of awareness of open data in some fields of research, it is crucial that the entities responsible for disseminating open information re-evaluate the datasets they publish, publicizing their characteristics, emphasizing their interest and facilitating access to data of real scientific value.

The survey established six age groups, gathering a total of 249 responses of lecturers with some knowledge of Government Open Data. The age group with the highest number of responses was 41 to 50 years old, with the 51 to 60 age group following closely behind. Among the employment categories, University Professors (Profesores Titulares) contributed 209 responses, University Full Professors (Catedráticos) provided 150 responses, and Contracted Professors (Profesores Contratado Doctor) accounted for 76 responses. The responses from the 22 to 30 age group (68) aligned closely with those from doctoral students (62). The 31 to 40 age group, with 129 responses, corresponded to that of Associate Professors and Contracted Professors (Profesores Asociados y Profesores Contratado Doctor), who provided 61 and 50 responses, respectively. The responses from the 41 to 50 age group were primarily from Tenured Professors.

In comparison to other studies, the Belmont Forum study reports 1247 respondents who provided age information, with an average age of 43. The median age falls within the 41 to 45 range, and the ages span from 20 to 90 [51]. Other research, such as the analysis of open data from Great Britain [50], [48] the Asedie Report of 2021, and the study on public perception of open data in Mexico [49], do not detail the age ranges of their samples, making it challenging to establish direct comparisons. However, when comparing with studies on demographics in Spanish public universities [75], it is evident that the results align with the broader demographic trends. Specifically, the number of respondents aged 61 to 65 (55) exceeds those aged 22 to 30 (50), supporting [75] Rodríguez's findings from Universidad Complutense de Madrid that highlight the aging trend among Spanish university faculties.

To summarize the key points of our analysis, overall there are no significant differences in knowledge of Open Government Data by gender. Instead, the main factor influencing the use and understanding of these data is scientific discipline, with Social Sciences showing the highest engagement with these activities. Age appears to have a minimal impact on this knowledge, as does contract type.

## 5. Conclusions

This study demonstrates that the surveyed scientific community reinforces previous general observations. This group not only possesses the expertise needed for searching, processing, and analyzing information but also frequently acts as a group of potential reusers of open data. They often develop services and conduct research based on this data,

creating significant value for society. In addition to open research data, general open data has become an essential tool in both research and university teaching. However, it remains underutilized for various reasons [12], including the limited technical capacity of academic staff [55].

As previously explained, the sample was representative, with 783 responses collected. Consequently, the results and conclusions can be extrapolated to the broader demographic of the Spanish university faculty. The sample includes representation from 34 of Spain's 50 provinces and both autonomous cities, covering 47 universities and 21 research centres.

In terms of research areas, there is a significant response bias from researchers in Area 1 (Social Sciences), which garnered the highest number of responses. This is due to two key reasons: first, researchers in this field are particularly interested in quantitative studies of social impact and have shown a strong engagement with them. Second, the Social Sciences encompass a broad range of disciplines related to social phenomena, which makes them more involved with government and open data initiatives compared to natural sciences and technology fields. Nonetheless, the data gathered from other research areas is also valuable and informative, with nearly all areas being represented in the sample.

There is parity of responses in terms of gender: 41.76% of responses from women and 56.58% of responses from men guaranteeing equal representation.

One of the study's objectives was to assess researchers' awareness of open data. The results show that 79.31% of respondents acknowledge open data as potentially reusable resources in science, indicating a high level of familiarity. However, attention should be given to the 20.69% of researchers who either do not know about open data or chose not to answer. This group highlights a gap in understanding regarding information sources, data usage, and their potential, as well as a possible deficiency in the dissemination and promotion of open data by the authorities. As we explained in the Introduction, the importance of the movement of Open Data lies in reuse [8,76–78]. The percentage of reuse obtained was low (50.57%) if we take into account that university lecturers are qualified personnel and are familiar with the scientific method and sources of information, as well as having the capacity to work with and interpret data for the development of new studies or value-added services.

Several reasons were cited for the lack of data reuse. Primarily, researchers pointed to the scarcity of relevant data in their fields. Notably, Area 19 (Biomedicine) shows high awareness of open data but exhibits low reuse levels. This discrepancy may be attributed to the field's tendency to generate and use its own data. Similarly, disciplines such as Area 10 (Physical Sciences), Area 14 (Chemical Sciences and Technologies), Area 8 (Psychology), and Area 4 (Mind, Language, and Thought) also report minimal reuse of public sector data, indicating a general shortage of applicable information.

Another important outcome of our investigation is the need to prioritize the standardization of certain datasets. Standardization of datasets would facilitate the integration and use of data from a variety of sources, leading to increased reuse with reduced effort and enhancing overall interoperability.

In conclusion, we can assert that open data from government sources are valuable for scientific research, particularly in fields such as Social Sciences, Industrial Production, Civil Engineering, Engineering for Society, Information and Communication Technologies, Economics, and Environmental Sciences and Technologies. These are the areas where researchers most frequently reported reusing such data.

**Author Contributions:** Conceptualization, A.F.-S.; Methodology, C.V.-C., E.A.S.-P. and A.F.-S.; Validation, E.A.S.-P.; Formal analysis, A.F.-S.; Investigation, C.V.-C., E.A.S.-P. and A.F.-S.; Resources, C.V.-C. and A.F.-S.; Data curation, C.V.-C.; Writing—original draft, C.V.-C., E.A.S.-P. and A.F.-S.; Writing—review & editing, C.V.-C., E.A.S.-P. and A.F.-S.; Visualization, C.V.-C.; Supervision, E.A.S.-P.; Project administration, A.F.-S.; Funding acquisition, A.F.-S. All authors have read and agreed to the published version of the manuscript.

**Funding:** The third author has been supported by Gran PID-2019-105708RB/Stable methodologies to evaluate and measure quality, interoperability, blockchain and reuse of open data in the agricultural field. Funded by: MCIN/AEI/10.13039/501100011033.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The original data presented in the study are openly available in FigShare at https://figshare.com/articles/dataset/Encuesta_de_datos_abiertos_para_investigadores_Espa_a_/26798023?file=48697102 accessed on 7 October 2020.

**Acknowledgments:** Funding from the Universidad Católica de Valencia San Vicente Mártir is gratefully acknowledged.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Soeharjono, S.; Roche, D.G. Reported individual costs and benefits of sharing open data among Canadian academic faculty in ecology and evolution. *BioScience* **2021**, *71*, 750–756. [CrossRef]
2. Stieglitz, S.; Wilms, K.; Mirbabaie, M.; Hofeditz, L.; Brenger, B.; López, A.; Rehwald, S. When are researchers willing to share their data? Impacts of values and uncertainty on open data in academia. *PLoS ONE* **2020**, *15*, e0234172. [CrossRef]
3. Beverungen, D.; Hess, T.; Köster, A.; Lehrer, C. From private digital platforms to public data spaces: Implications for the digital transformation. *Electron. Mark.* **2022**, *32*, 493–501. [CrossRef]
4. Jetzek, T.; Avital, M.; Bjorn-Andersen, N. The Sustainable Value of Open Government Data. *J. Assoc. Inf. Syst.* **2019**, *20*, 702–734. [CrossRef]
5. Pereira, G.V.; Macadar, M.A.; Luciano, E.; Testa, M.G. Delivering public value through open government data initiatives in a Smart City context. *Inf. Syst. Front.* **2017**, *19*, 213–229. [CrossRef]
6. Bonina, C.; Eaton, B. Cultivating Open Government Data Platform Ecosystems Through Governance: Lessons From Buenos Aires, Mexico City and Montevideo. *Gov. Inf. Q.* **2020**, *37*, 101479. [CrossRef]
7. Hein, A.; Engert, M.; Ryu, S.; Schaffer, N.; Hermes, S.; Krcmar, H. Building open government data platform ecosystems: A dynamic development approach that engages users from the start. *Gov. Inf. Q.* **2023**, *40*, 101878. [CrossRef]
8. Scupola, A.; Mergel, I. Co-production in digital transformation of public administration and public value creation: The case of Denmark. *Gov. Inf. Q.* **2022**, *39*, 101650. [CrossRef]
9. Lnenicka, M.; Nikiforova, A.; Luterek, M.; Milic, P.; Rudmark, D.; Neumaier, S.; Bolívar, M.P.R. Identifying patterns and recommendations of and for sustainable open data initiatives: A benchmarking-driven analysis of open government data initiatives among European countries. *Gov. Inf. Q.* **2024**, *41*, 101898. [CrossRef]
10. Okunleye, O.J. The Role of Open Data in Driving Sectoral Innovation and Global Economic Development. *J. Eng. Res. Rep.* **2024**, *26*, 222–243. [CrossRef]
11. Gascó-Hernández, M.; Martin, E.; Reggi, L.; Pyo, S.; Luna-Reyes, L. Promoting the use of open government data: Cases of training and engagement. *Gov. Inf. Q.* **2018**, *35*, 233–242. [CrossRef]
12. Hernández-Leo, D.; Amarasinghe, I.; Beardsley, M.; Hakami, E.; García, A.R.; Santos, P. Responsible Educational Technology Research: From Open Science and Open Data to Ethics and Trustworthy Learning Analytics. In *Data Cultures in Higher Education: Emergent Practices and the Challenge Ahead*; Springer International Publishing: Cham, Switzerland, 2023; pp. 179–200.
13. Raffaghelli, J.E.; Sangrà, A. Data Cultures in Higher Education: Acknowledging Complexity. In *Data Cultures in Higher Education: Emergent Practices and the Challenge Ahead*; Springer International Publishing: Cham, Switzerland, 2023; pp. 1–39.
14. Meijer, A.J.; Curtin, D.; Hillebrandt, M. Open government: Connecting vision and voice. *Int. Rev. Adm. Sci.* **2012**, *78*, 10–29. [CrossRef]
15. Wirtz, B.W.; Birkemeyer, S. Open Government: Origin, Development, and Conceptual Perspectives. *Int. J. Public Adm.* **2015**, *38*, 381–396. [CrossRef]
16. Nikiforova, A.; McBride, K. Open government data portal usability: A user-centred usability analysis of 41 open government data portals. *Telemat. Inform.* **2021**, *58*, 101539. [CrossRef]
17. Calderón, C.; Lorenzo, S. *Open Government: Gobierno Abierto*; Algón Editores MMX: Jaén, Spain, 2010; Available online: https://libros.metabiblioteca.org/handle/001/163 (accessed on 13 March 2022).
18. McDermott, P. Building open government. *Gov. Inf. Q.* **2010**, *27*, 401–413. [CrossRef]
19. Corojan, A.; Campos Domínguez, E.M. *Gobierno Abierto: Alcance e Implicaciones*; Fundación Ideas: Montequinto, Spain, 2011; Available online: http://uvadoc.uva.es/handle/10324/42441 (accessed on 15 March 2022).
20. Ferrer-Sapena, A.; Peset, F.; Aleixandre-Benavent, R. Acceso a los datos públicos y su reutilización: Open data y open government. *Prof. Inf.* **2011**, *20*, 260–269. [CrossRef]

21. Obama, B. Transparency and Open Government, Memorandum for the Heads of Executive Departments and Agencies. 2009. Available online: https://web.archive.org/web/20161215181923/https://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment (accessed on 5 May 2021).
22. Garriga-Portolà, M. ¿Datos abiertos? Sí, pero de forma sostenible. *Prof. Inf.* **2011**, *20*, 298–303. [CrossRef]
23. Obama, B. *Memorandum for the Heads of Executive Departments and Agencies: Open Government Directive*; Executive Office of the President, Office of Management and Budget: Washington, DC, USA, 2009. Available online: https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/memoranda_2010/m10-06.pdf (accessed on 10 June 2021).
24. Acosta y Plata. Participación, Colaboración, Transparencia y Datos Abiertos. *Teuken Bidikay* **2016**, *7*, 209–222.
25. Attard, J.; Orlandi, F.; Scerri, S.; Auer, S. A systematic review of open government data initiatives. *Gov. Inf. Q.* **2015**, *32*, 399–418. [CrossRef]
26. Okamoto, K. Introducing open government data. *Ref. Libr.* **2017**, *58*, 111–123. [CrossRef]
27. Çaldag, M.T.; Gokalp, M.O.; Gokalp, E. Open government data: Analysing benefits and challenges. In Proceedings of the 2019 1st International Informatics and Software Engineering Conference (UBMYK), Ankara, Turkey, 6–7 November 2019; pp. 1–6. [CrossRef]
28. Çaldag, M.T.; Gökalp, E. The maturity of open government data maturity: A multivocal literature review. *Aslib J. Inf. Manag.* **2022**, *29*, 1007–1030. [CrossRef]
29. Cingolani, L. The survival of open government platforms: Empirical insights from a global sample. *Gov. Inf. Q.* **2021**, *38*, 101522. [CrossRef]
30. Alexopoulos, C.; Zuiderwijk, A.; Charalabidis, Y.; Loukis, E.; Janssen, M. Designing a second generation of open data platforms: Integrating open data and social media. In *Electronic Government: 13th IFIP WG 8.5 International Conference, EGOV 2014, Dublin, Ireland, 1–3 September 2014, Proceedings*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 230–241. [CrossRef]
31. Open Government Partnership. Open Government Partnership Global Report—Executive Summary. 2019. Available online: https://www.opengovpartnership.org/wp-content/uploads/2019/05/Global-Report_Executive-Summary_EN.pdf (accessed on 1 June 2021).
32. Hansson, K.; Belkacem, K.; Ekenberg, L. Open government and democracy: A research review. *Soc. Sci. Comput. Rev.* **2015**, *33*, 540–555. [CrossRef]
33. Altayar, M.S. Motivations for open data adoption: An institutional theory perspective. *Gov. Inf. Q.* **2018**, *35*, 633–643. [CrossRef]
34. Çaldag, M.T.; Gökalp, E.; Alkis, N. Analyzing determinants of open government based technologies and applications adoption in the context of organizations. In Proceedings of the International Conference on E-Learning, e-Business, Enterprise Information Systems, and e-Government (EEE'19), Las Vegas, NE, USA, 29 July–1 August 2019; pp. 50–56. Available online: https://www.researchgate.net/publication/335014598_Analyzing_Determinants_of_Open_Government_Based_Technologies_and_Applications_Adoption_in_the_Context_of_Organizations (accessed on 10 October 2021).
35. Schmidthuber, L.; Stütz, S.; Hilgers, D. Outcomes of open government: Does an online platform improve citizens' perception of local government? *Int. J. Public Sect. Manag.* **2019**, *32*, 489–507. [CrossRef]
36. Kassen, M. Open data and its peers: Understanding promising harbingers from Nordic Europe. *Aslib J. Inf. Manag.* **2020**, *72*, 765–785. [CrossRef]
37. Matheus, R.; Janssen, M. A systematic literature study to unravel transparency enabled by open government data: The window theory. *Public Perform. Manag. Rev.* **2020**, *43*, 503–534. [CrossRef]
38. Janssen, M.; Charalabidis, Y.; Zuiderwijk, A. Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Inf. Syst. Manag.* **2012**, *29*, 258–268. [CrossRef]
39. Borgman, C.L.; Brand, A. The Future of Data in Research Publishing: From Nice to Have to Need to Have? Harvard Data Science Review, (Special Issue 4). 2024. Available online: https://hdsr.mitpress.mit.edu/pub/imfn7i7f/release/2 (accessed on 18 March 2022). [CrossRef]
40. Borgman, C.L.; Pasquetto, I.V. Why Data Sharing and Reuse Are Hard to Do Irene Pasquetto the Big Data to Knowledge (BD2K) Guide to the Fundamentals of Data Science. Available online: https://escholarship.org/uc/item/0jj17309#author (accessed on 21 April 2017).
41. Ramachandran, R.; Bugbee, K.; Murphy, K. From open data to open science. *Earth Space Sci.* **2021**, *8*, e2020EA001562. [CrossRef]
42. Zuiderwijk, A.; Rhythima, S.; Wei, J. What drives and inhibits researchers to share and use open research data? A systematic literature review to analyze factors influencing open research data adoption. *PLoS ONE* **2020**, *15*, e0239283. [CrossRef]
43. Abella, A.; Ortiz de Urbina Criado, M.; De Pablos Heredero, C.; Vidal-Cabo, C.; Ferrer-Sapena, A. *La reutilización de datos abiertos en España II*; Desidedatum, Universidad Rey Juan Carlos y Cátedra de Transparencia y Gestión de Datos UPV: Madrid, Spain, 2019; Available online: https://ctranspaupv.com/informes/la-reutilizacion-de-datos-abiertos-en-espana-ii/ (accessed on 15 January 2020).
44. Van Hesteren, D.; van Knippenberg, L. Open Data Maturity Report 2020. 2021. Available online: https://op.europa.eu/en/publication-detail/-/publication/66c72f75-a40e-11ec-83e1-01aa75ed71a1/language-en (accessed on 11 January 2022). [CrossRef]
45. European Data Portal. Available online: https://data.europa.eu/catalogue-statistics/currentState/countries?locale=en (accessed on 11 January 2022).
46. DataEurope.eu. Open Data Maturity Report. 2021. Available online: https://www.capgemini.com/es-es/wp-content/uploads/sites/16/2022/01/Open-Data-Maturity-2021.pdf (accessed on 16 March 2022).

47. Huyer, E. Capgemini I. *The Economic Impact of Open Data Opportunities for Value Creation in Europe*. 2020. Available online: https://op.europa.eu/en/publication-detail/-/publication/1021d8a7-5782-11ea-8b81-01aa75ed71a1/language-en (accessed on 11 January 2022). [CrossRef]
48. ASEDIE. Informe del Sector Infomediario. 2021. Available online: https://static1.squarespace.com/static/600a99c4d2a8133c3599fc67/t/607ff089596f9b27bfc11636/1618997397041/ISI21+VF++web.pdf (accessed on 18 March 2022).
49. Estrada Zamora, C.; Vargas-Hernández, J.G. Percepción ciudadana del open data y la innovación en México. *Rev. Investig. Tecnol. Inf.* **2018**, *6*, 86–95. Available online: https://dialnet.unirioja.es/servlet/articulo?codigo=7107376 (accessed on 10 October 2021). [CrossRef]
50. Davies, T. Open Data, Democracy and Public Sector Reform. A Look at Open Government Data Use from data.gov.uk. Master's Thesis, University of Oxford, Oxford, UK, 2010; pp. 1–47.
51. Schmidt, B.; Gemeinholzer, B.; Treloar, A. Open Data in Global Environmental Research: The Belmont Forum's Open Data Survey. *PLoS ONE* **2016**, *11*, e0146695. [CrossRef] [PubMed]
52. Coughlan, T. The use of open data as a material for learning. *Educ. Technol. Res. Dev.* **2020**, *68*, 383–411. [CrossRef]
53. Raffaghelli, J.E.; Manca, S. Exploring the social activity of open research data on ResearchGate: Implications for the data literacy of researchers. *Online Inf. Rev.* **2023**, *47*, 197–217. [CrossRef]
54. Quarati, A.; Raffaghelli, J.E. Do researchers use open research data? Exploring the relationships between usage trends and metadata quality across scientific disciplines from the Figshare case. *J. Inf. Sci.* **2022**, *48*, 423–448. [CrossRef]
55. Santos-Hermosa, G.; Quarati, A.; Loría-Soriano, E.; Raffaghelli, J.E. Why Does Open Data Get Underused? A Focus on the Role of (Open) Data Literacy. In *Data Cultures in Higher Education: Emergent Practices and the Challenge Ahead*; Springer International Publishing: Cham, Switzerland, 2023; pp. 145–177.
56. Santos-Hermosa, G. Impact and Implementation of UNESCO's Recommendation on Open Educational Resources in Academic Libraries: SPARC Europe Case Study. *Res. Learn. Technol.* **2024**, *32*, 3183. [CrossRef]
57. Xu, Q. Modeling college faculty users' potential acceptance of library data services for research and teaching. *Inf. Discov. Deliv. ahead-of-print*. **2024**. [CrossRef]
58. Guenduez, A.A.; Frischknecht, R.; Frowein, S.C.; Schedler, K. Government-university collaboration on smart city and smart government projects: What are the success factors? *Cities* **2024**, *144*, 104648. [CrossRef]
59. Guenduez, A.A.; Mergel, I.; Schedler, K.; Fuchs, S.; Douillet, C. Institutional work in smart cities: Interviews with smart city managers. *Urban Gov.* **2024**, *4*, 80–90. [CrossRef]
60. Fiebig, T.; Gürses, S.; Gañán, C.H.; Kotkamp, E.; Kuipers, F.; Lindorfer, M.; Sari, T. Heads in the clouds: Measuring the implications of universities migrating to public clouds. *arXiv* **2021**, arXiv:2104.09462.
61. Williamson, B.; Hogan, A. Pandemic Privatisation in Higher Education: Edtech and University Reform. QUT ePrints. 2021. Available online: https://eprints.qut.edu.au/216578/ (accessed on 16 March 2021).
62. Lo, C.K. What is the impact of ChatGPT on education? A rapid review of the literature. *Educ. Sci.* **2023**, *13*, 410. [CrossRef]
63. Ministerio de Universidades. Datos y Cifras del Sistema Universitario Español. Publicación 2020–2021. 2021. Available online: https://www.universidades.gob.es/stfls/universidades/Estadisticas/ficheros/Datos_y_Cifras_2020-21.pdf (accessed on 12 November 2021).
64. Ahmed, A.; Al-Khatib, A.; Boum, Y.; Debat, H.; Gurmendi Dunkelberg, A.; Hinchliffe, L.J.; Jarrad, F.; Mastroianni, A.; Mineault, P.; Pennington, C.R.; et al. The future of academic publishing. *Nature human behaviour* **2023**, *7*, 1021–1026. [CrossRef] [PubMed]
65. Lichtenauer, N.; Schmidbauer, L.; Wilhelm, S.; Wahl, F. A Scoping Review on Analysis of the Barriers and Support Factors of Open Data. *Information* **2024**, *15*, 5. [CrossRef]
66. Piwowar, H.; Priem, J.; Larivière, V.; Alperin, J.P.; Matthias, L.; Norlander, B.; Farley, A.; West, J.; Haustein, S. The state of OA: A large-scale analysis of the prevalence and impact of Open Access articles. *Peer J.* **2018**, *6*, e4375. [CrossRef] [PubMed]
67. Borgnman, C.L. Open data, Grey Data, and Stewardship: Universities at the Privacy Frontier. *Berkeley Technol. Law J.* **2018**, *33*, 365–412. [CrossRef]
68. Cole, N.L.; Kormann, E.; Klebel, T.; Apartis, S.; Ross-Hellauer, T. The societal impact of Open Science: A scoping review. *R. Soc. Open Sci.* **2024**, *11*, 240286. [CrossRef]
69. Fox, J.; Pearce, K.E.; Massanari, A.L.; Riles, J.M.; Szulc, Ł.; Ranjit, Y.S.; Trevisan, F.; Soriano, C.R.; Vitak, J.; Arora, P.; et al. Open science, closed doors? Countering marginalization through an agenda for ethical, inclusive research in communication. *J. Commun.* **2021**, *71*, 764–784. [CrossRef]
70. Umbach, G. Open Science and the impact of Open Access, Open Data, and FAIR publishing principles on data-driven academic research: Towards ever more transparent, accessible, and reproducible academic output? *Stat. J. IAOS* **2024**, *40*, 59–70. [CrossRef]
71. Aleixandre-Benavent, R.; Vidal-Infer, A.; Alonso-Arroyo, A.; Peset, F.; Ferrer-Sapena, A. Research Data Sharing in Spain: Exploring determinants, practices and perceptions. *Data* **2020**, *5*, 29. [CrossRef]
72. Ley Orgánica 3/2007, del 22 de Marzo, para la Igualdad efectiva de Mujeres y Hombres. BOE nº 71, de 23 de Marzo de 2007. Referencia: BOE-A-2007-6115. Available online: https://www.boe.es/buscar/pdf/2007/BOE-A-2007-6115-consolidado.pdf (accessed on 7 October 2020).
73. Mestre, T. Cuotas voluntarias y legales en España: La paridad a examen. *Rev. Esp. Investig. Sociol. (REIS)* **2008**, *123*, 123–150. Available online: http://www.reis.cis.es/REIS/PDF/REIS_123_051215167211333.pdf (accessed on 10 October 2021).

74. EP Data (S.F.). Profesores, Investigadores y Resto de Personal Docente en la Universidad de España, Datos y Estadísticas. Portal de Visualización de Datos Actualizado en Noviembre de 2021. Available online: https://www.epdata.es/datos/profesores-investigadores-resto-personal-docente-universidad-espana-datos-estadisticas/412 (accessed on 15 November 2021).
75. Hernández Armenteros, J.; Pérez García, J.A. Demografía Universitaria Española: Aproximación a Su Dimensión, Estructura y Evolución. 9. 2018. Available online: https://issuu.com/efse/docs/ct9-studia-xxi (accessed on 11 January 2022).
76. Kassen, M. Understanding motivations of citizens to reuse open data: Open government data as a philanthropic movement. *Innovation* **2021**, *23*, 44–70. [CrossRef]
77. Weerakkody, V.; Irani, Z.; Kapoor, K.; Sivarajah, U.; Dwivedi, Y. Open data and its usability: An empirical view from the Citizen's perspective. *Inf. Syst. Front.* **2017**, *19*, 285–300. [CrossRef]
78. Ansari, B.; Barati, M.; Martin, E.G. Enhancing the usability and usefulness of open government data: A comprehensive review of the state of open government data visualization research. *Gov. Inf. Q.* **2022**, *39*, 101657. [CrossRef]