Literature review of business cases in 3Os

Project ZOOOM - 3Os and IP Awareness raising for collaborative ecosystems

Grant Agreement No. 101007385

Funded by the European Union

Document Control Information

Information	Value
Document Title:	D2.1 – Literature review of business cases in 3Os
Project:	ZOOOM Project GA #101070077
Document Version:	V03
Dissemination Level:	PU - Public
Date:	29/03/2022
DOI	

Document Author(s)

Organisation	Name(s)
VTT	Jutta Suksi, Pasi Pussinen, Juri Mattila
UNITN	Andrea Caputo, Seckin Celik, Loris Gaio, Alessandro Rossi, Erica Santini, Davide Serpico
KUL	Ivo Emanuilov
HIT	Stefano Menegazzi

ZOOOM Project (GA 101007385) was Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them.

Table of Content

List of Abbreviations

Acronym	Explanation
3Os	Free and open source software, open hardware and open data
4Es	The following four emerging technologies: (1) artificial intelligence (2) quantum technologies and Internet of Trust (3) blockchain (4) robotics
ACM	Association for Computing Machinery
AI	Artificial intelligence
CLA	Contributor license agreement
DIH	Digital Innovation Hub
EDIH	European Digital Innovation Hub
ICT	Information and communication technology
IP	Intellectual property
IT	Information technology
LDA	Latent Dirichlet Allocation
LSA/LSI	Latent Semantic Analysis
OCD	Open Corporate Data
OD	Open Data
OGD	Open Government Data
OH	Open Hardware
OS	Open Software
OSH	Open Source Hardware
OSOR	Open Source Observatory
OSPD	Open-source product development

OSS	Open Source Software
OSSN	Open Source Service Network
PATLIB	Patent Library
pLSA	Probabilistic Latent Semantic Analysis
SaaS	Software as a Service
SME	Small and medium enterprise
SSPL	Server-Side Public License
TTO	Tech Transfer Office
VN	Value Network
WP	Work package
ZOOOM GA	Grant Agreement of the ZOOOM Project
ZOOOM	'3Os and IP awareness raising for collaborative ecosystems' (ZOOOM) Project, Grant Agreement No. 101007385

List of Figures

List of Tables

List of Annexes

Executive Summary

ZOOOM Project aims to increase awareness of the importance of intellectual property generation and management in collaborative innovation ecosystems from the perspective of free and open source software, open hardware and open data (in short, the 3Os).

The core objective of ZOOOM Project lies in supporting the relevant stakeholders in the ecosystems to match business models with appropriate 3Os licensing frameworks. ZOOOM approaches this goal from a case perspective focusing on four emerging technologies: artificial intelligence, quantum technologies and Internet of Trust, blockchain, and robotics (in short, the 4Es).

This Deliverable D2.1 ‘Literature review of business cases in 3Os’ aims to collect the core existing and available information relating to ZOOOM objectives and give a solid basis for implementing the further steps of the project: case studies, ecosystem study and framework building.

The deliverable studies the business aspects relating to 3Os through a two-fold approach. The academic literature review forms the basis for understanding value creation and business models in 3Os. It is supplemented by a review of professional literature (white papers, empirical papers, project reports and other sources of information) that sheds light to issues that are difficult to address from an academic perspective.

The focus of the analytical phase of the academic literature review is on case studies, best practices, main licensing strategies, and main business models, which are collected into tables (Table 5, pp. 43, and Table 6, pp. 45). They give a quick overview of the key papers and practices in the areas of 3Os and 4Es. The analysis is concluded with a discussion of emerging questions, problems and major insights emerging from the literature, focusing mostly on business aspects (models, value creation, value capture), barriers and opportunities, innovation through open knowledge and ecosystem level considerations.

Professional literature aims to complement the academic literature review especially from the perspective of 4Es. It identifies current start-up business activities around 3Os and 4Es and analyses white papers from the perspective of potential value in 4Es. In addition, it supplements the academic literature discussion on the 3Os business models and value creation from the point of view of other than academic sources, such as professional literature, white papers, reports, online sources and highlights some of the concurrent topics and trends.

1. Academic Literature Review

1.1. Research Protocol, Sources and Methods

This academic literature review is based on various methodologies. First, ZOOOM authors conducted bibliometric and topic modelling analyses based on the Scopus database. Second, the ZOOOM authors manually categorized the dataset into the three types of open knowledge and the four technologies on which the ZOOOM Project focuses: open software, open data, open hardware, AI, blockchain, quantum technology and Internet of Trust, and robotics. Finally, ZOOOM authors selected a number of papers and book chapters from the literature and analysed them thoroughly in order to extract information about case studies, best practices, main licensing strategies, and main business models, as well as major insights about ecosystems and value creation/capture aspects. The following sub-sections outline the bibliometric and topic-modelling methodology in more detail.

1.1.1. Bibliometric Analysis Methodology

Bibliometric analysis is a quantitative method used to evaluate scientific publications based on the analysis of bibliographic data. The purpose of bibliometric analysis is to measure the impact of scientific publications, identify trends and patterns in research, and assess the productivity and collaboration of authors, institutions, and countries. Bibliometric analysis relies on the use of various metrics such as citation counts, h-index, and journal impact factors to evaluate research output.

Bibliometric analysis can be used to identify research clusters in a dataset by analysing the co-citation patterns of publications. Co-citation analysis involves identifying publications that are frequently cited together, indicating a common research theme or topic. Identifying research clusters involves several steps.

First, a bibliographic dataset is compiled, which may include all publications in a particular field or topic resulting from a query or string. Next, the dataset is analysed by dedicated software to identify the most frequently cited publications. These highly cited publications are often seminal works in the field and provide a starting point for identifying research clusters. Once the highly cited publications are identified, the co-citation patterns of these publications are analysed. This involves identifying other publications that are frequently cited alongside highly cited publications. These publications are likely to be closely related to the research themes or topics covered by the highly cited publications and form the basis of research clusters. Finally, the research clusters are visualized using a variety of techniques, such as cluster maps or dendrograms, which show the relationships between the publications and research clusters. These visualizations can help researchers and policymakers understand the structure and dynamics of research in a particular field, identify areas of collaboration and potential gaps in research, and inform future research directions.

The analysis below was conducted through the R package Bibliometrix (Aria & Cuccurullo, 2017), based on the database of Scopus inspected in January 2023. Scopus was considered instead of other databases, such as Web of Science (WOS), due to the fact that the topic of interest to the ZOOOM Project can be considered a niche at the intersection of several fields. Scopus offers a wider database also including proceedings and has been found to have a 60% larger coverage than WoS (Zhao & Strotmann 2015), and it was thus selected to cover a wider variety of research outputs. Moreover, bibliometric analyses in the social sciences typically deploy only one database to mitigate data homogenization issues faced when working with multiple databases (see Mariani et al. 2022).

T2.1 search can be replicated through the following Boolean multilevel search string aimed at identifying articles that include key terms in their title, abstract or keywords:

Level 1 - theoretical relevance - “business models”	TITLE-ABS-KEY (("business model*" OR "revenue model*" OR "value captur*" OR "value creat*" OR "value deliver*" OR "moneti*") OR ("licensing"))
Level 2 - contextual relevance - “open data, software, hardware”	AND (TITLE-ABS-KEY ((open W/1 sourc*) AND (software* OR data OR hardware*)) OR TITLE-ABS-KEY ((free OR libre) W/1 (software*)) OR (TITLE-ABS-KEY ((open OR free OR libre) W/1 (data OR hardware*))))
Level 3 - exclusion/inclusion criteria	AND (LIMIT-TO (PUBSTAGE , "final")) AND (EXCLUDE (SUBJAREA , "ENER") OR EXCLUDE (SUBJAREA , "MEDI") OR EXCLUDE (SUBJAREA , "EART") OR EXCLUDE (SUBJAREA , "PHYS") OR EXCLUDE (SUBJAREA , "BIOC") OR EXCLUDE (SUBJAREA , "CENG") OR EXCLUDE (SUBJAREA , "CHEM") OR EXCLUDE (SUBJAREA , "HEAL") OR EXCLUDE (SUBJAREA , "PHAR") OR EXCLUDE (SUBJAREA , "IMMU") OR EXCLUDE (SUBJAREA , "DENT") OR EXCLUDE (SUBJAREA , "NURS")) AND (LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "ch")) AND (LIMIT-TO (LANGUAGE , "English")) AND (EXCLUDE (PUBYEAR , 2023))

This final string was obtained after several attempts using other keywords and operators, and was chosen as the most complete search on the subject (a list of older, alternative strings can be found in *Annex 1*).

The first and second levels of the string allowed to identify the publications that regard business aspects and licensing strategies in the domains of open software, data, and hardware. Synonymic terms such as “free” and “libre” were included. The third level of the string excludes results that are not published yet, results that have been published later than 2022, and results from non-relevant research areas, while including articles, book chapters and conference proceedings.

Using these search criteria, the initial dataset included 1005 items (see *Annex 2*). Considering that several publications were multidisciplinary because of the broad scope of the search string, a filtering process was necessary to ensure adherence to the research question. Therefore, ZOOOM authors carried out a refinement process based on abstracts and keywords. In the file, the authors have highlighted in red the items that had been considered out of scope and excluded from the analyses below.

The refined dataset, where the irrelevant items have been removed, includes 767 items (see *Annex 3*). This represents an arguably complete collection of the academic literature on the topic on which the ZOOOM Project focuses.

Some methodological considerations regarding the analysis performed are listed below:

- Only items that were clearly irrelevant have been coloured in red and removed from the dataset to be used for the analysis. This process was done manually by two researchers and it relied on the information provided by the abstracts.
- The choice of being inclusive in the final dataset may mean that some items that have been retained in the dataset might still turn out to be not particularly relevant for the project.
- The dataset might not include some relevant items due to their absence in the Scopus database, publication after the data collection date, or because the article keywords did not include any of the words included in the query above.
- It is unlikely, however, that the bibliometric analysis below would change drastically if a small number of relevant items were excluded from the dataset or a small number of irrelevant items were retained.

1.1.2. Topic Modelling Methodology

The refined dataset was analysed through topic modelling techniques in order to both verify the data from the bibliometric analysis and potentially identify further research clusters or topics. Topic modelling is a computational technique used to discover patterns or themes in a large corpus of text documents. By identifying the main themes present in the data, researchers can better understand the content of the documents and identify patterns and trends that may not be immediately apparent.

Topic modelling algorithms work by analysing the co-occurrence patterns of words in the text data and clustering them into groups based on their semantic similarity. This approach tries to reduce the dimensionality of the original text, representing texts through the words present in a (set of) topic space(s) rather than in the original set of words; this is pursued by an unsupervised learning approach, where a function of entropy of information is optimized. It can be compared to clustering and, as in the case of clustering, the number of topics is a decision parameter. By doing topic modelling, one builds clusters of words rather than clusters of texts; text is thus a mixture of all the topics, each having a specific weight. Sets of words in topics must be interpreted and "tagged" in order to assign meaning to the set of words that sum up each topic.

There are several existing algorithms one can use to perform the topic modelling. The most common are Latent Semantic Analysis (LSA/LSI), Probabilistic Latent Semantic Analysis (pLSA), and Latent Dirichlet Allocation (LDA). In this analysis, ZOOOM authors used LDA, a generative probabilistic model that assumes each topic is a mixture over an underlying set of words, and each document is a mixture of over a set of topic probabilities. More specifically, LDA is a topic modelling algorithm, devised by Blei et al. (2003), that takes a generative probabilistic approach and treats each document as a probability distribution over the various topics and thus groups documents based on the percentage of textual patterns in commonality. From each resulting group one can extract the most frequent word patterns and, based on them, infer the topic that characterizes each single document.

1.2. Results

1.2.1. Bibliometric Analysis Results

Table 1 Main Information about the Refined Dataset

Description	Results
MAIN INFORMATION ABOUT DATA	
Timespan	1999:2022
Sources (Journals, Books, etc)	496
Documents	767
Annual Growth Rate %	18.44
Document Average Age	8.85
Average citations per doc	10.6
References	23216
DOCUMENT CONTENTS	
Keywords Plus (ID)	3475
Author's Keywords (DE)	1804
AUTHORS	
Authors	1791
Authors of single-authored docs	166
AUTHORS COLLABORATION	
Single-authored docs	185
Co-Authors per Doc	2.8
International co-authorships %	21.38
DOCUMENT TYPES	
article	301
book chapter	46
conference paper	420

s

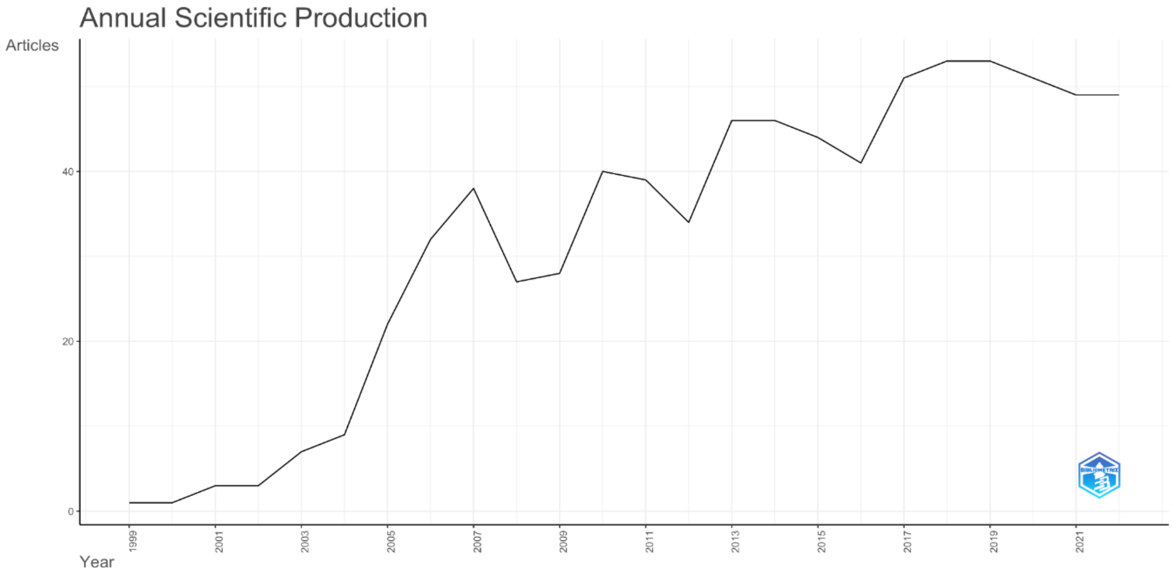

Figure 1 Annual Scientific Production

The analysis of annual scientific production reveals an increasing trend from 1999 to 2022. This indicates that the research community has been actively producing more publications over the years. This growth in scientific output can be attributed to various factors such as increased funding, advances in technology, and the availability of more data sources. The trend of increasing scientific production suggests that the research community is growing and expanding, which could lead to more opportunities for collaboration and the development of new research ideas in the area of open-source projects.

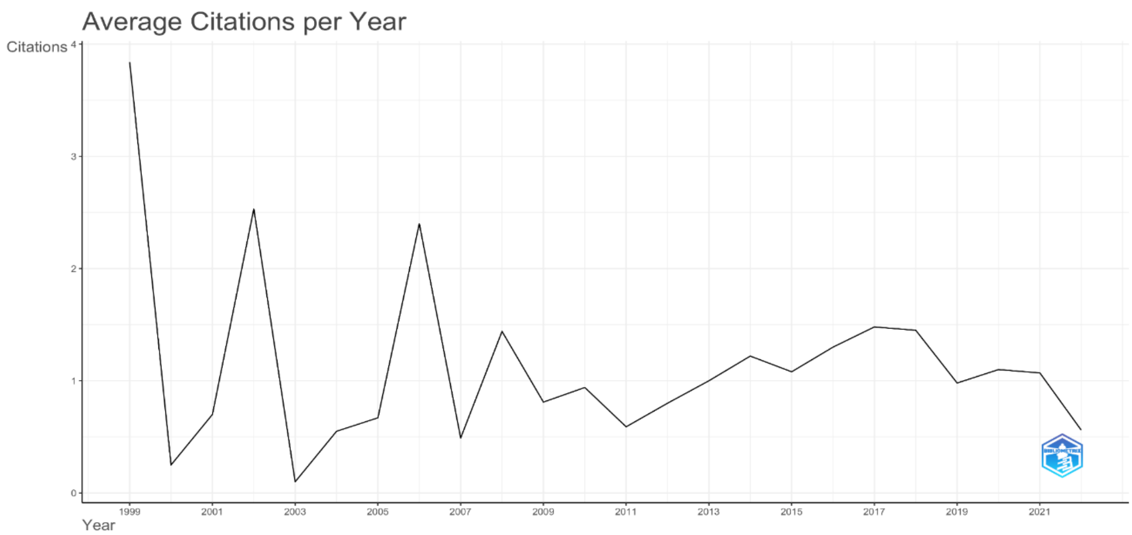

Figure 2 Average Citations per Year

Figure 2 shows that there has been a fluctuating trend in the number of citations received by publications over time. There was a significant decrease in the year 1999, followed by dramatic fluctuations until 2011. After 2011, there was a steadier increase until 2018, after which there was a decrease until 2022. These fluctuations in citations could be attributed to various factors, including the popularity of the research topic, the quality of the research, and changes in citation behaviour. Further investigation into the reasons for these fluctuations could reveal important insights into the research community's dynamics and its impact on the wider scientific landscape.

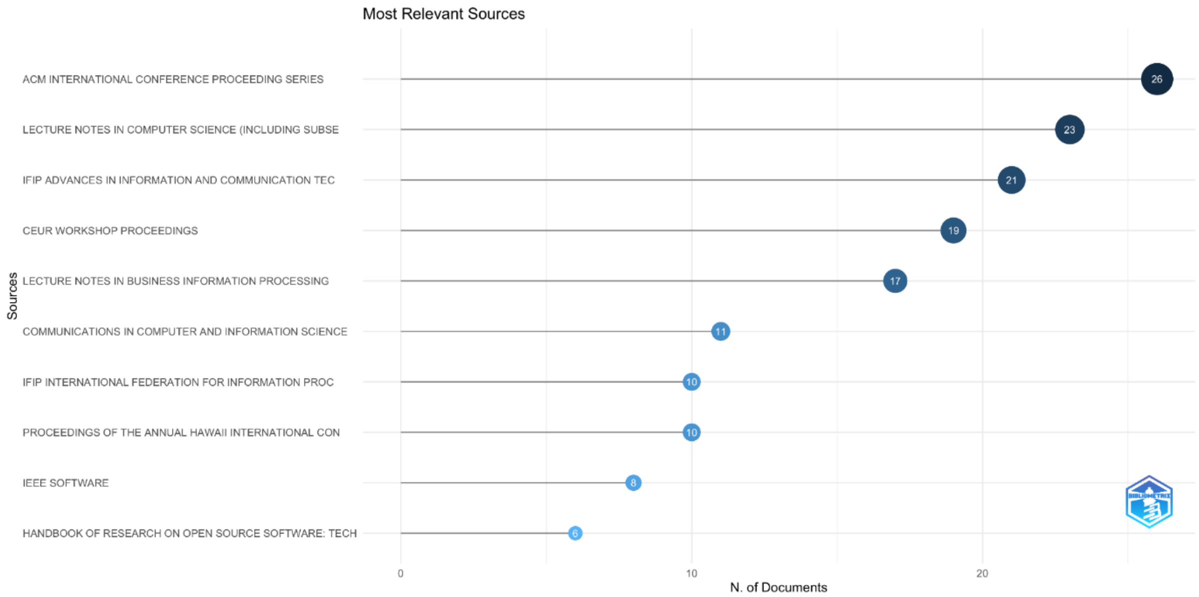

Figure 3 Most Relevant Sources

Figure 3 identifies the sources that have published the most influential and widely cited articles within a particular research community. This metric can help researchers identify important journals that are worth targeting for publication and conference for attendance. The figure for most relevant sources reveals that conference proceedings, particularly those published by the Association for Computing Machinery (ACM), are the most relevant sources in the T2.1 dataset. This suggests that the research community is actively publishing their work in conference proceedings and that these publications are well-regarded by their peers.

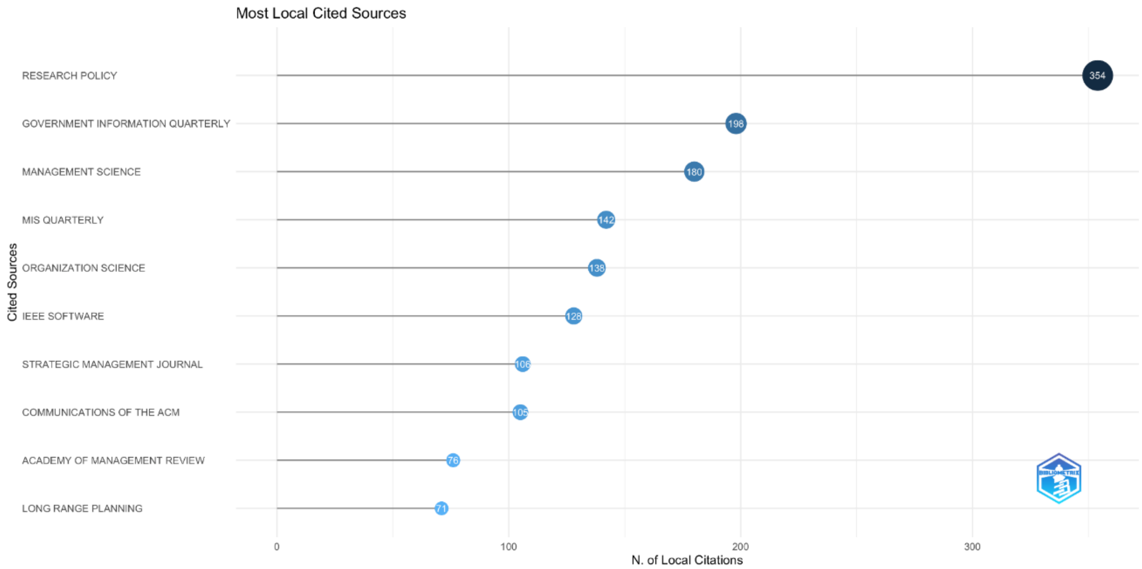

Figure 4 Most Local Cited Sources

Figure 4 identifies the sources that have been cited most frequently by researchers within a particular research community. This metric can help researchers identify the most influential sources within their own community and provide insights into the topics and themes that are of most interest to their peers. The figure for most local cited sources reveals that important journals like Management Science, Organization Science, and Strategic Management Journal are the most cited within the research community. The journal Research Policy has the highest locally cited score of 354, suggesting that it is highly regarded within the community.

Figure 5 Sources' Local Impact by TC Index

Figure 5 measures the impact of individual sources within a particular research community, based on their citation counts and the citation patterns of other sources within the same community. The figure reveals that three journals dominate T2.1 dataset: Management Science, Research Policy, and Information Systems Research. These journals have a high TC index, indicating that they are highly influential within the research community. The dominance of these journals suggests that they are highly respected within the community and that their articles have a significant impact on the field.

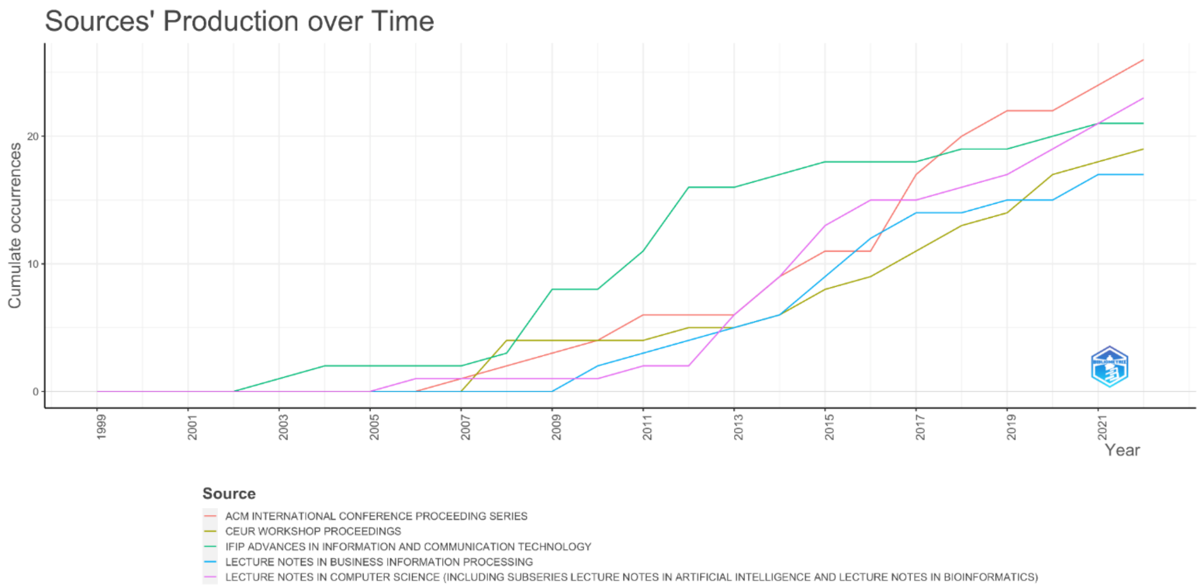

Figure 6 Sources' Production over Time

Figure 6, on sources' production over time, reveals that, according to T2.1 dataset, the production in literature started in 2002 with IFIP Advances in Information and Communication Technology. This suggests that the research community started actively publishing their work in this journal in 2002 and that it has since become an important source of information and ideas within the field.

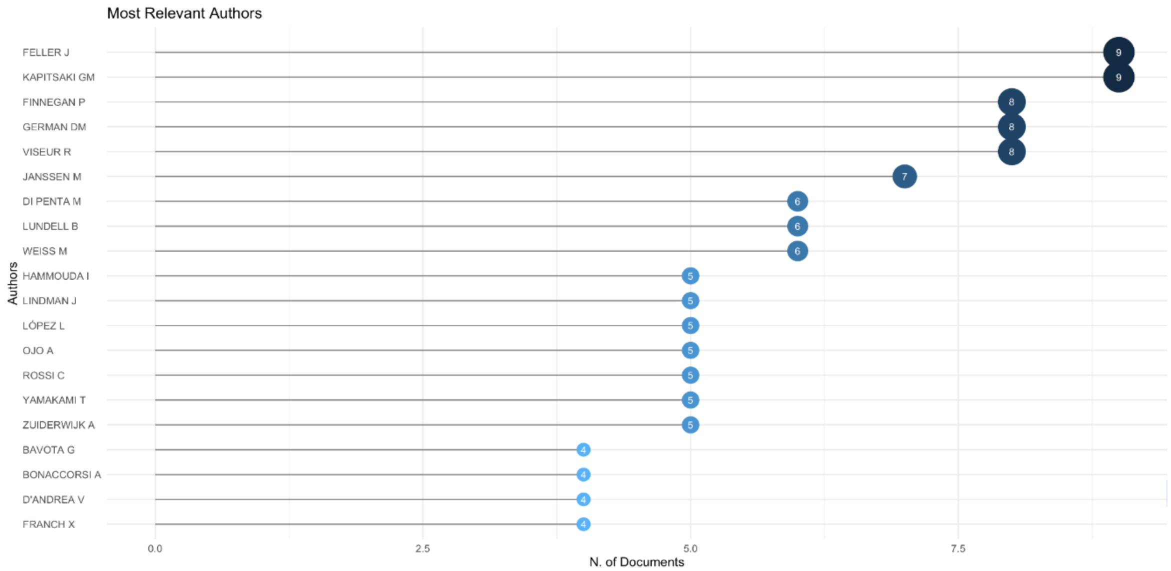

Figure 7 Most Relevant Authors

Figure 7 identifies the authors who have produced the most influential and widely cited articles within a particular research community. The figure reveals that Feller J and Kapitsaki GM are the two most relevant authors in the dataset, followed by Finnegan P, German DM, and Viseur R with the same number of documents. These authors' high relevance suggests that their work is highly regarded by their peers and has made a significant contribution to the field. The insights gained from this figure could be used to identify potential collaborators to develop research collaborations, and to understand the dynamics of the research community.

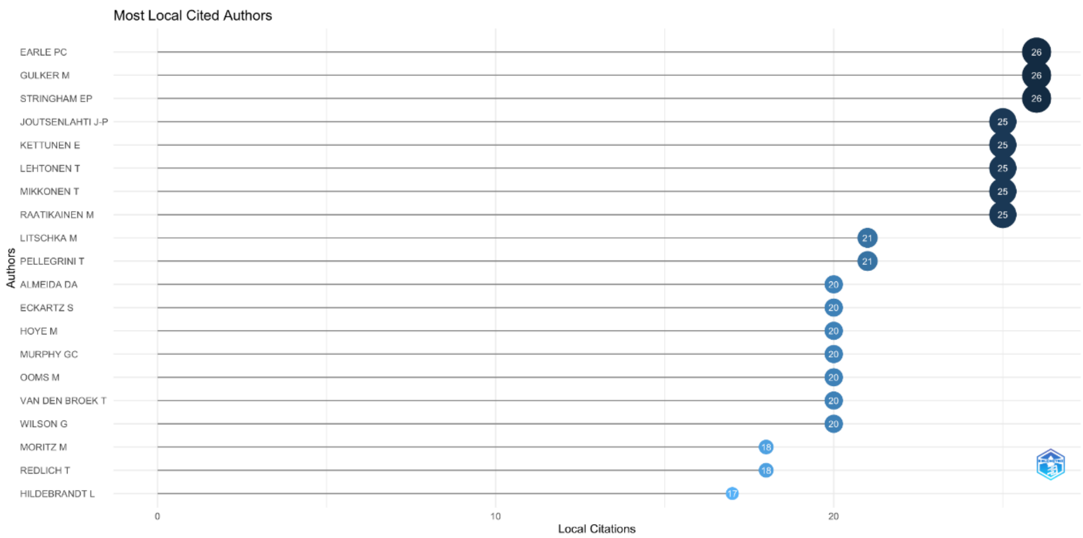

Figure 8 Most Local Cited Authors

Figure 8 identifies the authors who have been cited most frequently by researchers within a particular research community. Most local cited authors are an important group of authors in bibliometric analysis, as they have received significant attention and recognition within a specific geographical area. In the dataset, Earle PC, Gulker M, and Stringham EP are at the top with a local citation score of 26, followed by Joutsenlahti JP, Kettunen E, Lehtonen T, and Mikkonen T with a local citation score of 25. These authors have made important contributions to the literature in their respective fields and have received recognition from the local community. It is worth noting that their impact may not be as significant in other regions, and therefore their influence should be interpreted within the local context.

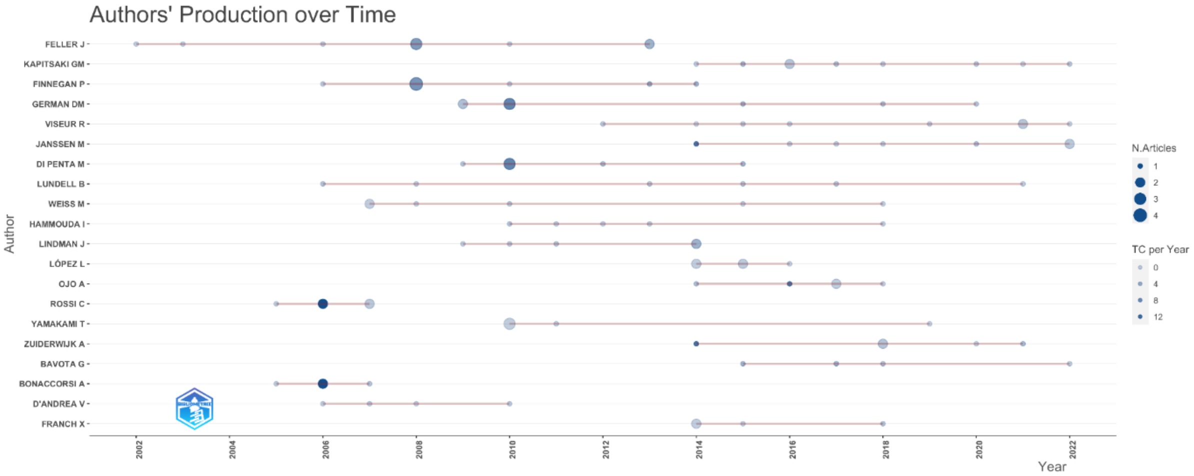

Figure 9 Authors' Production over Time

Figure 9 provides insight into the publishing activity of the most relevant authors in the dataset. According to T2.1 analysis, Kapitsaki GM, Viseur R, and Jannsen M continued to publish until 2022, which is the year of the dataset's end of the scope. On the other hand, Feller J and Finnegan P, who are also among the most relevant authors in the dataset, stopped publishing in 2013 and 2014, respectively. This information is valuable for understanding the trends and dynamics of the research community in a specific field over time.

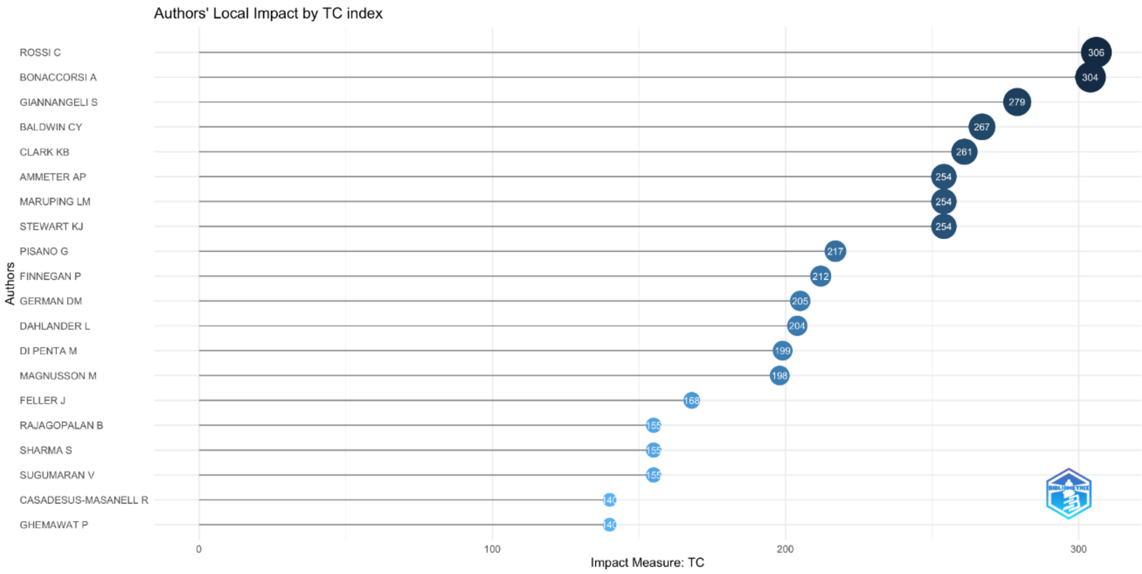

Figure 10 Authors' Local Impact by TC Index

Figure 10 measures the impact of individual authors within a particular research community, based on their citation counts and the citation patterns of other authors within the same community. Rossi C, Bonaccorsi A, Giannangeli S, and Baldwin CY are the top three authors, whose studies have been locally impacted by the TC scores of 306, 304, and 279, respectively. These authors have made significant contributions to their field, and their work has been highly influential within the local context. Their impact on the literature is likely to be significant, and their research is likely to shape the direction of future research in the field.

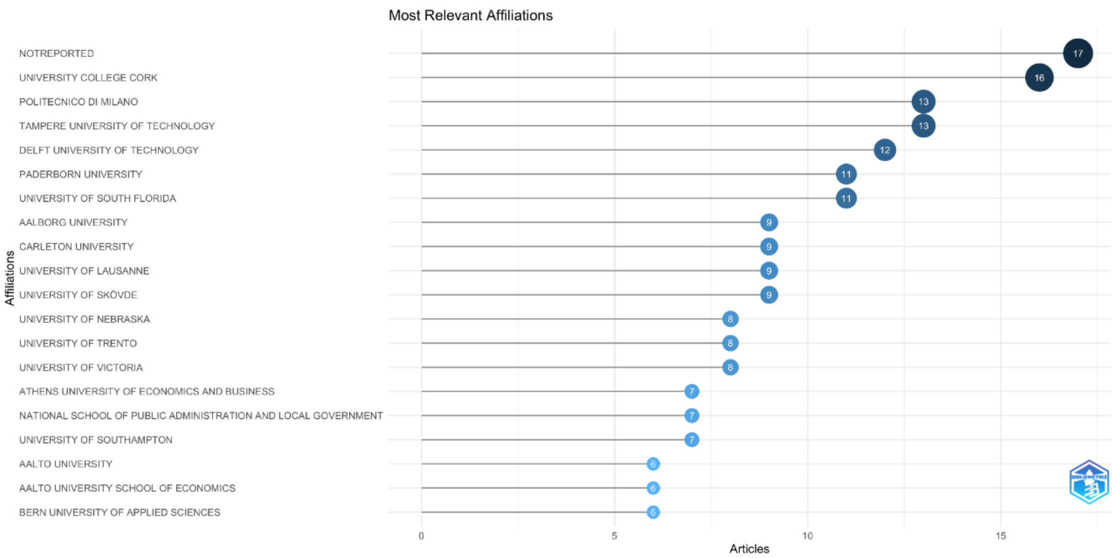

Figure 11 Most Relevant Affiliations

Figure 11 identifies the affiliations (e.g., universities, research institutions, companies) that have produced the most influential and widely cited articles within a particular research community. University College Cork, and Politecnico di Milano are the top three organizations, with the number of articles 17, 16, and 13, respectively. It is worth noting that the number of articles published by an organization does not necessarily reflect the quality or impact of their research.

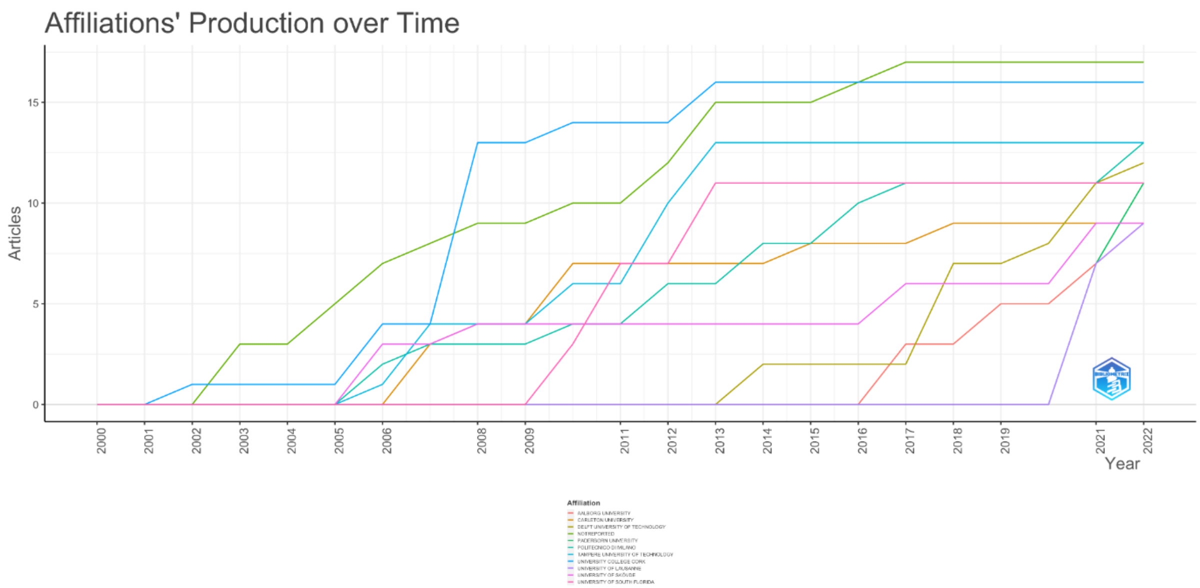

Figure 12 Affiliations' Production over Time

Figure 12, on affiliations' production over time, provides insight into the publishing activity of the most productive affiliations in T2.1 dataset. University College Cork, and Politecnico di Milano started publishing earlier than others and continue to publish until 2022. Not reported started publishing in 2002 with an increasing output until 2017 and kept publishing until 2022. University College Cork started publishing in 2001 and followed an increasing output in published papers until 2013 and kept it steady until 2022. Finally, Politecnico di Milano started publishing in 2005 and increased its production output until 2013, and kept the production output until 2022. This information is valuable for understanding the trends and dynamics of the research community in a specific field over time and can provide insights into the research strategies of these organizations.

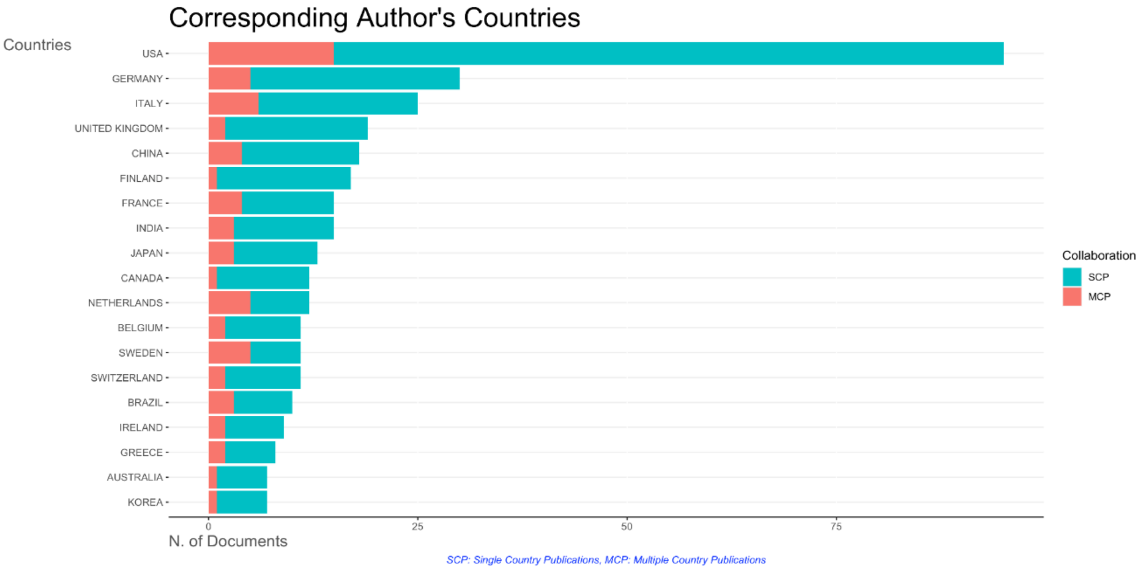

Figure 13 Corresponding Author's Countries

The corresponding author's countries figure (*Figure 13*) highlights the countries with the highest output of published documents in T2.1 dataset. The USA has the highest output, with over 75 documents, followed by Germany and Italy in the 25-50 document interval. The result may indicate that these countries have strong research infrastructures, funding opportunities, and research networks in the area of open innovation. However, it is important to note that this may also reflect the authors' affiliations, funding sources, and collaborations rather than the countries' research capacities.

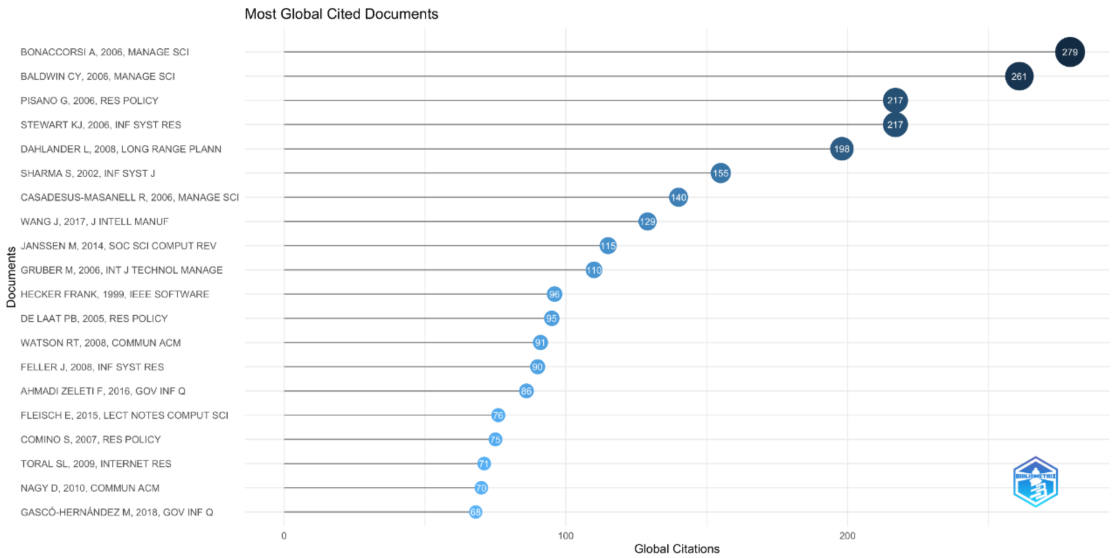

Figure 14 Most Global Cited Documents

The metric on most global cited documents can be used to identify articles that have made significant contributions to the broader scientific literature, identify important research questions, and establish collaborations with researchers from other disciplines. *Figure 14* shows the top four documents with the highest global citation scores in T2.1 dataset. Bonaccorsi, 2006, is the most globally cited document, followed by Baldwin, 2006, Pisano, 2006, and Stewart, 2006 with a global citation score of 261, 217, and 217, respectively. Interestingly, these documents were all published in the 2005-2009 interval, indicating that this period had a significant impact on literature in the field of open innovation. It is important to note that these documents may have influenced subsequent research and shaped the direction of the research in this field.

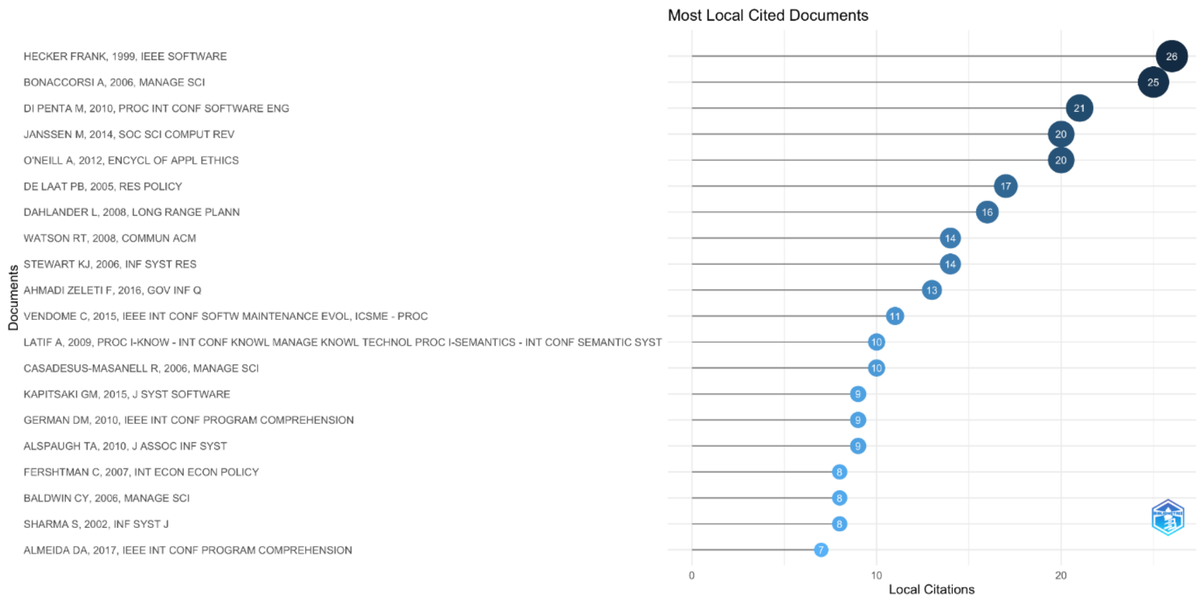

Figure 15 Most Local Cited Documents

Figure 15, on the most locally cited documents, highlights the documents with the highest local citation scores in T2.1 dataset. Hecker, 1999, is at the top with 26 local citations, followed by Bonaccorsi, 2006, and Di Penta, 2010, with local citation scores of 25 and 2, respectively. These documents may have had a significant impact on the local research community, reflecting their relevance to the local context or the quality of their contributions. However, it is important to note that the high local citation scores may also reflect the authors' affiliations, funding sources, and collaborations.

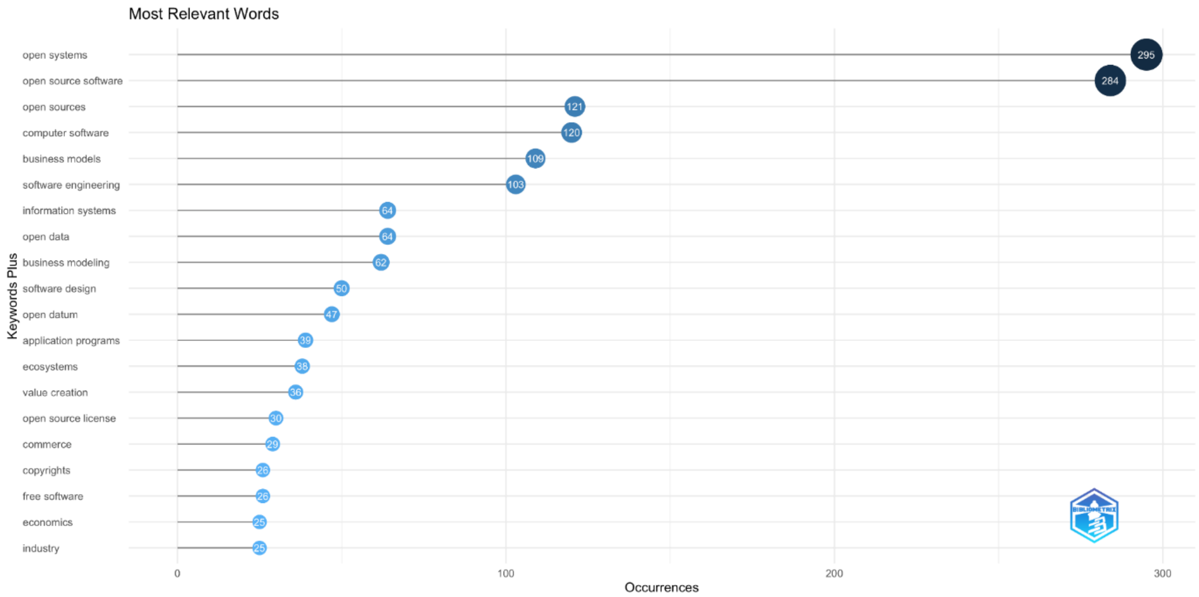

Figure 16 Most Relevant Words

Figure 16 shows the most frequently occurring words in the titles and abstracts of documents in the dataset. “Open systems” and “open-source software” are the two most relevant words that dominate the dataset, indicating their central role in the field. The keyword “open data” also has significant occurrences, reflecting the growing importance of open data in innovation processes. However, “open hardware” is not shown in the most relevant words, indicating that studies concerning open hardware are scarce in the dataset. This may suggest a gap in the literature on the role of open hardware in innovation processes, which could be an interesting area for future research.

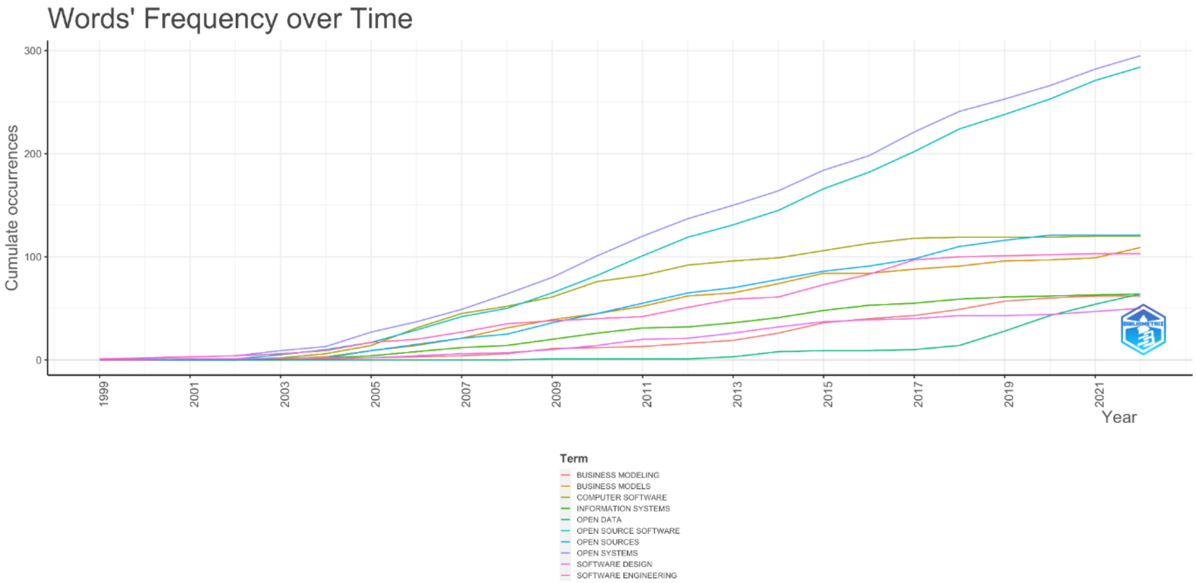

Figure 17 Words' Frequency over Time (Cumulative)

Figure 17 tracks the frequency of specific words or phrases used within publications from a particular research community over time. This metric can be used to identify emerging trends, changing priorities, and shifts in focus within a field. The analysis reveals that the frequency of usage of "open systems" and "open-source software" has dramatically increased over time. This finding indicates that these two terms have been highly popularized and commonly used in the literature in recent years. The other keywords show a steadier increase in their usage, such as "open data", "business models", "information systems", and "software engineering". This finding implies that the topics related to open systems and open-source software have gained more attention in the literature compared to other related topics.

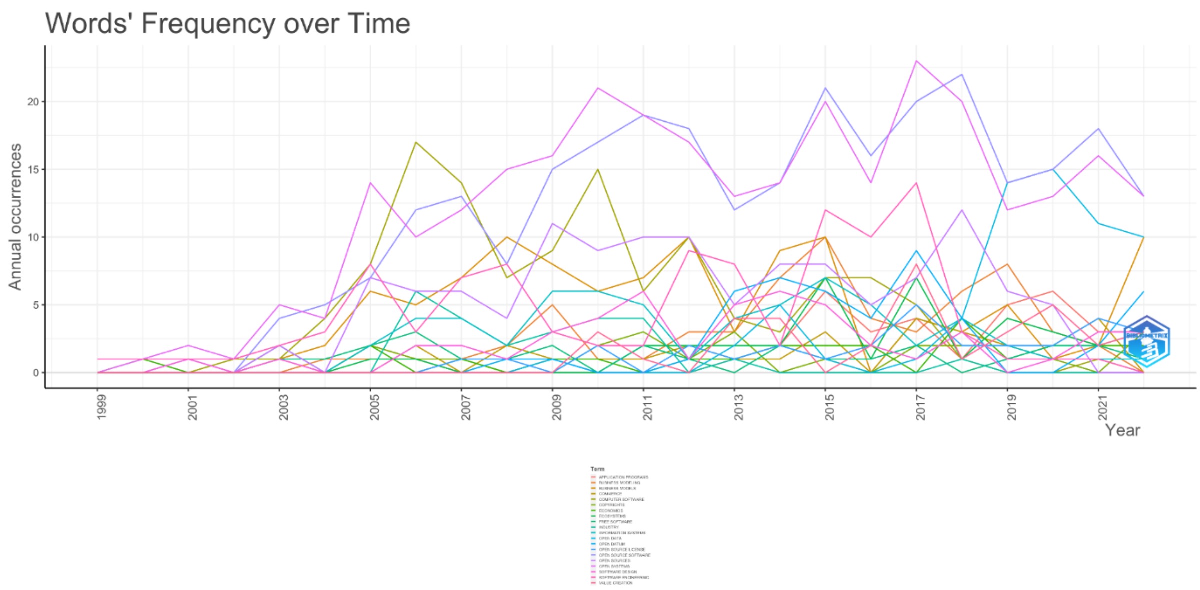

Figure 18 Words' Frequency over Time (Per Year)

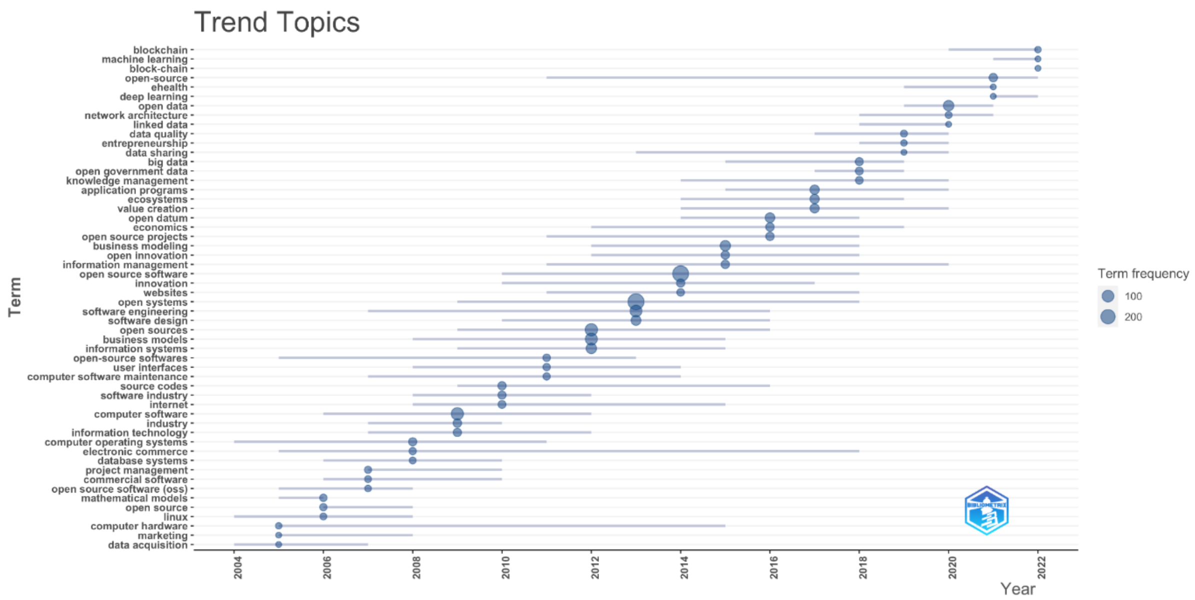

Figure 19 Trend Topics

Figure 19 identifies the topics and themes that are currently trending within a particular research community and field. The results of T2.1 analysis show that the usage of the term “open source software (OSS)” started to increase in 2005 and peaked in 2007. On the other hand, the terms related to open data such as “open data”, “linked data”, “open datum”, and “open government data” emerged later in the literature, starting from 2014 and peaking in 2021. This finding suggests that open data has gained popularity in the literature in recent years. Finally, the trend topics figure shows that terms related to technologies such as “blockchain” and “machine-learning” started to gain attention in the literature more recently.

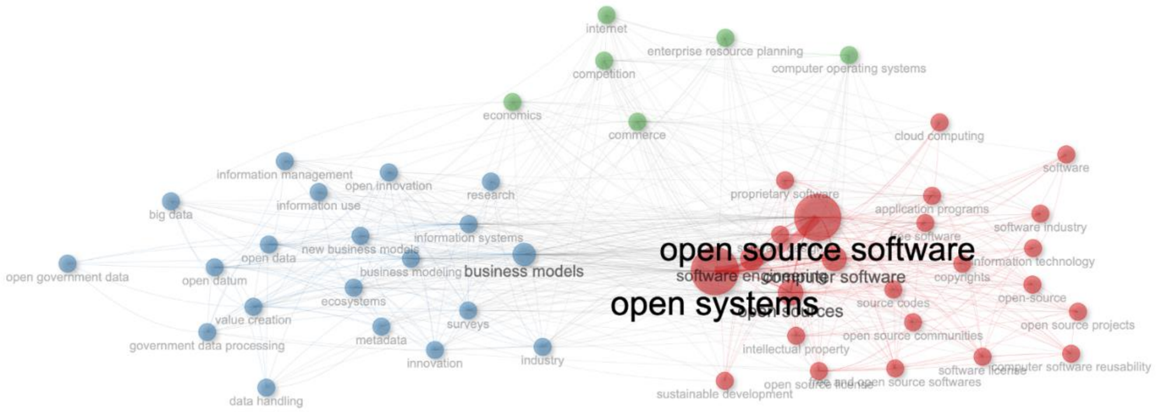

Figure 20 Co-occurrence Network

The ‘co-occurrence network’ metric can be used to identify clusters of related concepts and to evaluate the strength and nature of the relationships between different ideas within a field. The bibliometric analysis reveals a solidly structured network with meaningful links between nodes without any isolated nodes. The well-connected network exhibits three centralization occurrences that help to identify the key themes and research areas within the literature. One of the centralized networks is mostly about open-source software, which suggests that there is a strong focus on this topic within T2.1 dataset. The second network is related to open data, which also has a high level of activity in the literature. Notably, both open-data and open-source software studies intersect at the common ground of “business models” indicating that there is a significant emphasis on exploring the business models that support open-source and open-data initiatives. The third centralized network is more generic and includes terms like “internet”, “competition”, and “commerce”. This network implies that open-source and open-data projects are closely tied to broader concepts related to the digital economy. However, it is important to note that the “open hardware” and its derivatives are underrepresented in the co-occurrence network analysis, highlighting the need for future research in this area. Therefore, scholars and practitioners who are interested in the intersection of open-source projects and intellectual property (IP) management should consider investigating open hardware as an important area for future research.

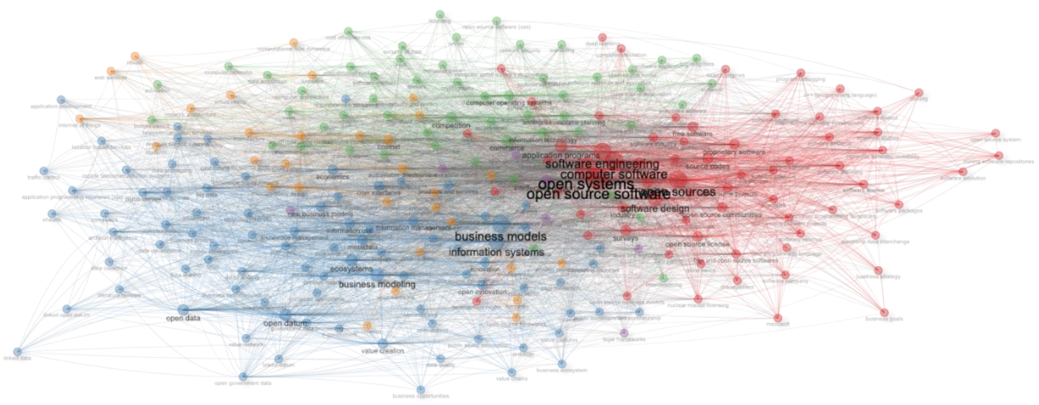

Figure 21 Thematic Map (Network)

Figure 21 represents a more detailed and complex network of relationships between different concepts, topics, or keywords within a particular research community. This picture clearly shows the complexity and interdisciplinarity of the field.

1.2.2. Topic Modelling Results

As introduced above, topic modelling is a natural language analysis model that attempts to bring out the latent topics present in a linguistic corpus by analysing the similarity of the term distribution in the document set with that of a topic or entity identified by the corpus structure. The technique is capable of scanning a document set, identifying patterns in words and phrases by grouping sets of words and expressions that best characterize a set of documents. The goal is to identify, from the groups of documents that have emerged, the topics (not known a priori) that characterize the set of documents and the related patterns of recurring words.

The analysis performed on the dataset of 767 articles can be considered as a kind of pilot test to preliminarily assess the structure of the latent topics present in the dataset. Therefore, it was decided to use the abstracts of the papers in the dataset, rather than full texts, as a starting point for data analysis. In order to make the LDA algorithm more efficient, texts must be prepared in advance. The abstracts were then cleaned to remove stopwords, i.e., lemmas with poor information value, in particular, by removing domain elements (publisher identifiers, words in common usage, etc.). At this stage of the work, it was also decided not to proceed with more sophisticated cleaning techniques, such as linguistic stemming, lemmatization, or the identification of n-grams (bigrams, trigrams, and the like) in order to preserve some of the semantic value of the original lemmas.

The following analysis aims at identifying the most appropriate model to fit the corpus. The authors started with a scan of the space of parameters and number of topics, using performance indicators (for details, see Arun et al. 2010) in order to guess the most meaningful number of topics stemming from the linguistic corpus. *Figure 22* allows to appreciate the results of the analysis performed at this stage; the plot shows the progress of a typical performance indicator (coherence) for a particular model fit to the data, characterized by a set of fixed initial parameters and the permutation in a given interval of the number of topics in this case, the optimal number of topics is associated with the maximum value of the coherence score.

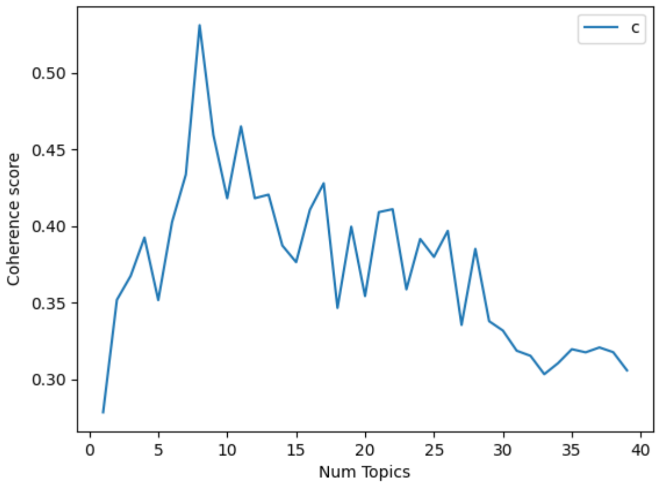

Figure 22 Coherence Score to the Number of Topics in the Corpus (Fixed Parameters Space)

The optimal number of topics stemming from the complete scan of the parameter space and associated with the highest coherence score is 8. The result of fitting the model to the data is shown in *Table 2* below, along with most representative words for each topic.

Table 2 Optimal Topic Model along with Relative Weight and 10 most Representative Words

Topic	weight	word1	word2	word3	word4	word5	word6	word7	word8	word9	word10
1	0.816	business	model	models	research	study	innovation	based	companies	process	knowledge
2	0.526	software	open	source	development	licensing	free	systems	products	industry	product
3	0.259	services	platform	cloud	applications	service	application	systems	system	platforms	users
4	0.184	licenses	license	licensing	code	foss	projects	developers	terms	information	users
5	0.093	system	tools	library	libraries	design	open-source	commercial	gis	computer	licensing
6	0.279	oss	firms	software	projects	communities	market	community	open-source	project	strategy
7	0.229	data	open	public	government	ogd	information	access	quality	sector	datasets
8	0.184	intellectual	resources	property	commons	standards	digital	public	creative	rights	collaborative

The weights column shows the Dirichlet's parameter (alpha) for each topic and allow us to measure the relative importance of the topic given the distribution of linguistic terms in the whole corpus. The weights are not standardized with each other (they do not sum to 1). The composition of topics with respect to linguistic terms, starting with the most relevant words, allows us to provide clues with respect to the semantics of the topic and thus to label them appropriately. For each group in the results, ZOOOM authors decided to assign labels as per *Table 3*. Moreover, the table shows the number of documents in the corpus and the relative share to the whole corpus by the most relevant topic.

Table 3 Labelled Topics and Distribution among Documents in the Corpus

Topic	Inferred Label	# Documents in the Corpus	Relative Frequency
1	business model	60	7.82%
2	open software	82	10.69%
3	cloud platforms	184	23.99%
4	license schema	97	12.65%
5	SW engineering	77	10.04%
6	communities	68	8.87%
7	open data	162	21.12%
8	property rights	37	4.82%

Finally, *Figure 23* below reports the evolution of topics in time, by number of articles referenced each year.

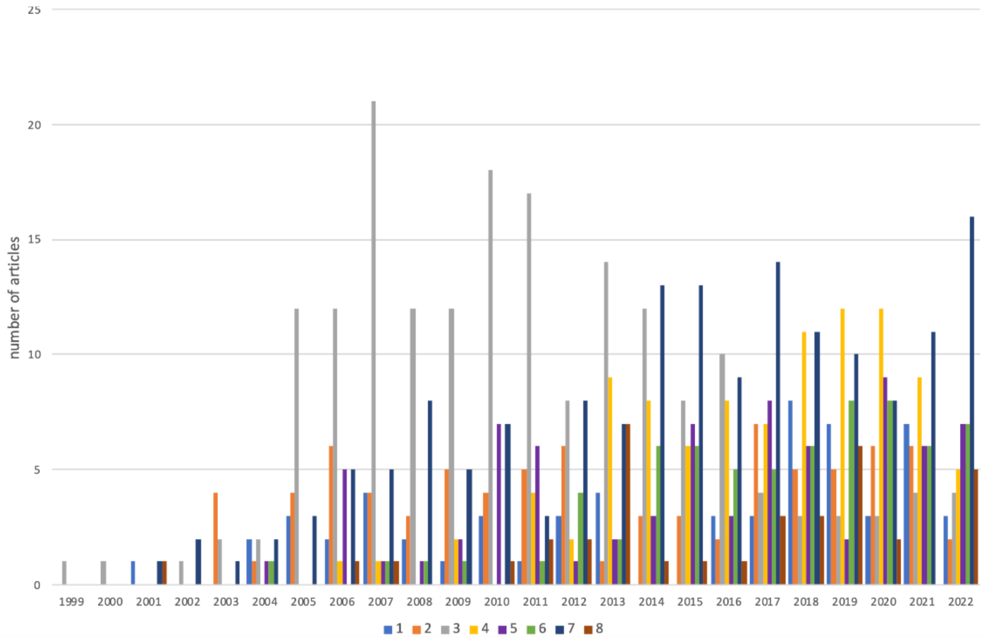

Figure 23 Distribution of Documents by Topic per Year

Topic 3 (cloud platforms) peaks around 2007 and gets less relevant later. Topic 4 (license schema) emerges clearly from 2011. Topic 7 (open data) peaks in recent years.

1.3. Synthesis

1.3.1. Literature Structure

The Bibliometrix Thematic Map function allowed the authors to categorize items into five research clusters:

- C1 Open systems [secondary keywords: "open source software"; "open sources"]
- C2 Open data [secondary keywords: "information systems"; "business modeling"]
- C3 Intellectual property [secondary keywords: "laws and legislation"; "economic and social effects"]

- C4 Competition [secondary keywords: “information technology”; “computer operating systems”]
- C5 Sustainable development [secondary keywords: “distributed computer systems”; “product design”].

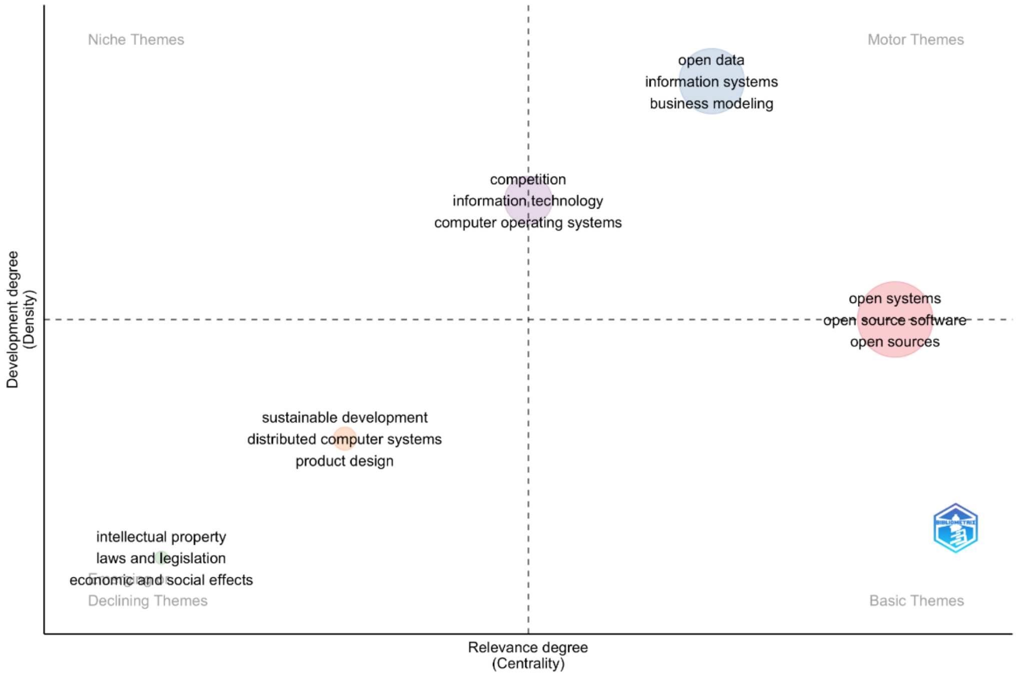

Figure 24 Thematic Map

As shown in *Figure 24*, five clusters are identified by Bibliometric through the Thematic Map function. To identify clusters that are relevant to ZOOOM, ZOOOM authors found it profitable to use the Louvain clustering algorithm.

While the meaning of C1, C2, C3, and C4 is pretty straightforward, it is worth clarifying that C5 (Sustainable Development) concerns good and sustainable practices in Open Hardware production, a field where assets sharing is key (indeed, C5 does include most items relating to Open Hardware, see *Table 6* below).

Another thing to notice is that some items from Open Software (OS) and Open Hardware (OH) appear to be mistakenly assigned by Bibliometrix to C2 (Open Data, in short OD), probably due to the software’s inclusion of two secondary keywords (“information systems” and “business modeling”) in such a cluster (see *Figure 24*). ZOOOM authors highlighted this by colouring in red such items in *Table 6* below.

1.3.2. Project Categorization

In order to inform the literature review and identify relevant case studies, ZOOOM authors categorized manually the items in the clean dataset according to the categories of the ZOOOM Project:

- Types of open elements: Open Software (0), Open Data (1), and Open Hardware (2). Some items have a double characterization (e.g., both OS and OH) if relevant to both, although the main focus of such items may be one of the two.
- Emerging technologies: AI, Blockchain, Cybersecurity, Robotics. In order to identify items that are specifically connected to these areas, ZOOOM authors investigated the dataset through the following keywords: “Additive Manufacturing”; “Artificial Intelligence”; “Blockchain”; “Crypto”; “Embedded”; “Hardware”; “Internet of Trust”; “Quantum”; “Robot”; “Robotics”.

Table 4 Manual Categorization of the Refined Database

Category	Results
Open Software (0)	482
Open Data (1)	262
Open Hardware (2)	38
AI	10
Blockchain	12
Cybersecurity	8
Robotics	6

Two main considerations need to be commented. First, OH appears to be underrepresented in the academic literature. This is somewhat surprising considering that the Open Source Hardware Association has certified 1663 Open Source Hardware (OSH) projects (https://certification.oshwa.org/list.html), the Open Know-How search engine lists 486 OSH projects (https://search.openknowhow.org), and Bonvoisin et al. (2018) analysed over 200 OSH projects.

Second, the categorization revealed that only a small number of items involve the four technological areas. Notably, there is no item in the dataset that is clearly connected with quantum technology and the Internet of Trust, but some items are related to cybersecurity, and this is why ZOOOM authors decided to include this category (however, ZOOOM authors understand that this category may not represent what the ZOOOM Project is looking for when it comes to open knowledge in quantum technology and Internet of Trust).

The finding above might be due to a delay in the academic community or the remote connection with business and licensing aspects. ZOOOM authors expects that more information on these areas will be found in the Practical Literature Review.

1.4. Analysis of the Literature

The clean dataset was investigated in order to identify the most relevant papers, best practices, case studies, main licensing strategies, and main business models adopted across the categories of OS, OD, and OH, as well as the four technological areas. Due to the extensive number of items in the dataset, ZOOOM authors applied a strict selection process based on which they inspected closely several papers for each category.

First, ZOOOM authors investigated the dataset through the following heuristics:

- The authors focused on the 50 top-cited items across all categories (both normalized and non-normalized top-cited);
- The authors focused on the most recent items across all categories (published between 2020-2022);
- The authors investigated every item in the OH category and in the four technological areas, due to the small number of items in these categories.

The authors thus identified 134 items that were worth analysing in detail. Although most items discuss various topics, the The authors initially categorized them as follows: Open Software [44]; Open Hardware [28]; Open Data [28]; AI [6]; Blockchain [6]; Cybersecurity [2]; Robotics [6]; General [14].[1]

After closer examination, the number of relevant items decreased to 48, distributed as follow: Open Software [18]; Open Hardware [12]; Open Data [11]; AI [1]; Blockchain [2]; Robotics [1]; General [3].[2]

[1] Items in the category “General” discuss the following topics more generally: business models and value creation/capture; licensing strategies and IP management; ecosystems; open and digital innovation.

[2] No relevant item was confirmed for the Cybersecurity category.

In this second selection, items have been excluded mostly due to their remote connection with business aspects (value creation/capture) or with the technological areas of interest.

Key information from the papers was collected in a separate file (see *Annex 4*) and categorized into two tables (*Table 5* and *Table 6*), the structure of which is based on the literature structure (*Section 2.3.1*, see above) and Project Categorization (*Section 2.3.2*, see above). The authors decided to collect information into two alternative tables to analyse such information from different angles (*Table 6* is more relevant for bibliometric purposes).

The tables below outline key papers, case studies, best practices, main licensing strategies, and main business models discussed in the selected papers. Such tables make it easy to check state of the art for each domain of the 3Os and technological areas (based on the literature selected, see above) and have a quick overview of what are the key papers and practices in a given area.

Table 5 Intersection between Main Information and ZOOOM Categories, Reflecting on the Manual Categorization of the Dataset Based on the Categories that are Relevant to the ZOOOM Project

	Key Papers	Case Studies / Best Practices	Licensing Strategies	Business Models
OS	Balka et al 2010 Bonaccorsi et al 2006 Casadesus-Masanell & Ghemawat 2006 Dahlander & Magnusson 2008 De Laat 2005 Dell'Era et al 2020 Feller et al 2008 Gruber & Henkel 2006 Höst & Orucevic-Alagic 2011 Morgan et al 2013 Pisano 2006 Rajala et al 2012 Sharma et al 2002 Singh & Phelps 2013 Stewart et al 2006 Teixeira et al 2015 Välimäki & Oksanen 2005 Watson et al. 2008	Apache; Blink; Cendio; CIRS; Cloud Foundry; Free Software Foundation; Freshmeat; GENIVI Alliance; IBM; JBoss; Linux (Embedded); Linux; MySQL; Netscape-Mozilla; ObjectWeb; Open Handset Alliance; OpenStack; OS Consortium; NRC; Osmosoft; Open OSX; OSX; Philips Healthcare; RedHat; Roxen; Sleepycat; SOT; Sourceforge; SpikeSource; Tizen; Trolltech; Webkit; Windows; Xen; ZEA Partners	GPL; LGPL; MPL; NPL	Complementary hardware; Complementary services; Complementary software; Corporate Distribution; Distribution & support; Flexibility in licenses; Leveraging communities; Mixed open/proprietary code; Open Source Service Network (OSSN); Open-source; Proprietary; Public cloud hosting; Second Gen Open; Software platforms; Sponsored Open Source; Traditional value-capture framework

	West 2007			
OD	Cho & Lee 2022 Corrales-Gray et al 2022 Enders et al 2021 Gao & Janssen 2022 Janssen & Zuiderwijk 2014 Kamariotou & Kitsios 2022 Krasikov et al 2020 Monino 2021 Temiz et al 2022 Van Loenen et al 2021 Zeleti et al 2016	COVID-19 Dashboard (Johns Hopkins University); Facebook; Google; Open Geographic data (Netherlands); Open Education data (Denmark); Open Agricultural data (Italy); Open Financial data (Greece); Open Legal data (France); Open Structural Genomics Consortium; Car Spotter; Solar Atlas; OD Municipality Nijmegen; CBS statistics Netherlands; Criminal Chart Netherlands	Dual licensing	Canvas; Demand-oriented Platform; Dual Licensing; Free as branded Advertising; Freemium; Infrastructural Razor and Blades; Open Data Canvas; Open Source; Premium; Sponsorship; Supply-oriented Platform; White-label Development
OH	Antoniou et al 2022 Balka et al 2010 Bonvoisin et al 2021 Chou 2021 Fjeldsted et al 2012 Hildebrandt et al 2022 Li & Seering 2019 Li et al 2021 Moritz et al 2018	3D printers (Makerbot, Prusa Research, Ultimaker, Dagoma); 3DRobotics; Adafruit; Android vs iOS; Architecture for Humanity; Arduino; Betamax vs VHS; Dobot; Face shields (COVID-19); Farmbot; IDEO; Local Motors; Lulzbot; M5stack; OpenMV; OpenROV; Parkfun; RepRap; Sparkfun; Ufactory; Ultimaker	BSD; CERN; CERN OHL; Dual Strategy (Open + Proprietary Platform); GPL; LGPL; TAPR	Canvas; Crowdfunding; Design Mode vs Production Mode; Distribution of designs; Distribution of manufactured third-party products; Dual licensing edition; Freemium; Leveraging communities; Manufacture and sell of products; Online Services; Open-source Development (OSD); Platformization (Closed Supply-chain Platform vs. Open Industry Platform); Proprietary; Retailer

	Viseur & Jullien 2022 Viseur 2012			
AI	Gao & Janssen 2020	Bliq (GE); Governmental Tech Agency of Singapore; Kent Police (UK)	OD licenses; OS licenses	Open-source
Blockchain	Kolade et al 2022 Kumar et al 2022	Coronet Blockchain (IBM BC); Lightency	AGPL; GPL	OS blockchain platform; Smart contracts
Robotics	Bonarini et al 2013	AIRLab (Triskar2 with R2P)	OD licenses; OS licenses	Open-source hardware; Open-source software
General	Chuang et al 2022 Duparc et al 2022 Haim Faridian et al 2021	Apache; IBM; JunoEMR; Linux; Meteor; Microsoft; Moodle; MySQL; Netscape; PrestaShop; Red Hat; Rufus; Sencha; Tendenci; TweetDeck	Apache; Creative Commons; F/LOSS; GNU; GPL2; GPLV3; Open Database; Proprietary	Funding-based; Infrastructure; Open Innovation Network; Open Innovation; Open-Core; Open-source Platform; Proprietary

Table 6 Intersection between Main Information and Research Clusters

	Key Papers	Case Studies / Best Practices	Licensing Strategies	Business Models
C1 Open systems	Fjeldsted et al 2012 Höst & Orucevic-Alagic 2011 Singh & Phelps 2013 Stewart et al 2006	Adafruit; Architecture for Humanity; Arduino; Linux (Embedded); Freshmeat; IDEO; Sourceforge	Open Source	Leveraging communities; Mixing open/proprietary code; Open-source Development Business Model Canvas
C2 Open data	Dell'Era et al 2020	Adafruit; Arduino; Blink; CIRS; Cloud Foundry; COVID-19	Dual Licensing	Canvas; Complementary hardware; Complementary

	Feller et al 2008 Gao & Janssen 2022 Janssen & Zuiderwijk Kamariotou & Kitsios 2022 Krasikov et al 2020 Li et al 2021 Morgan et al 2013 Teixeira et al 2015 Temiz et al 2022 Zeleti et al 2016	dashboard (Johns Hopkins University); Dobot; Facebook; Farmbot; Free Software Foundation; GENIVI Alliance; Google; Netscape-Mozilla; NRC; ObjectWeb; Open Handset Alliance; OpenROV; OpenStack; OS Consortium; Osmosoft; Philips Healthcare; Red Hat; Sparkfun; Structural Genomics Consortium; Tizen; Webkit; Xen; ZEA Partners; OD Municipality Nijmegen; CBS statistics Netherlands; Criminal Chart Netherlands		services; Complementary software; Demand-oriented Platform; Distribution & support; Dual Licensing; Free as branded Advertising; Freemium; Infrastructural Razor and Blades; Open Data Canvas; Open-source Service Network (OSSN); Premium; Public clouds hosting; Sponsorship; Supply-oriented Platform; White-label Development
C3 Intellectual property	Haim Faridian et al 2021	IBM; Meteor; Microsoft; MySQL; Netscape; Red Hat; TweetDeck		Open Innovation Network
C4 Competition	Bonaccorsi et al. 2006 Casadesus-Masanell & Ghemawat 2006 Dahlander & Magnusson 2008 De Laat 2005 Gruber & Henkel 2006 Pisano 2006 Sharma et al 2002 Välimäki & Oksanen 2005	Apache; Cendio; IBM; JBoss; Linux (Embedded); Linux; MySQL; Netscape; Open OSX; OSX; RedHat; Roxen; Sleepycat; SOT; SpikeSource; Trolltech; Windows	GPL; LGPL; MPL; NPL	Corporate Distribution; Open-source; Proprietary; Second Gen Open; Sponsored Open Source; Traditional value-capture framework

	Watson et al. 2008 West 2007			
C5 Sustainable development	Antoniou et al 2022 Bonvoisin et al 2021 Hildebrandt et al 2022 Moritz et al 2018 Viseur & Jullien 2022 Viseur 2012	3D printers (RepRap, Makerbot, Prusa Research, Ultimaker, Dagoma); Arduino; Face shields (COVID-19); Local Motors; Sparkfun	BSD; CERN OHL; CERN; GPL; LGPL; TAPR	Canvas; Crowdfunding; Design Mode vs Production Mode; Distribution of designs; Distribution of manufactured third-party products; Dual licensing edition; Freemium; Manufacture and sell of products; Online Services; Open-source; Platformization (Closed Supply-chain Platform vs. Open Industry Platform); Retailer

Table 6 above reflects the categorization of the dataset into five research clusters identified by Bibliometrix. In red are highlighted items that seem to have been assigned to incorrect clusters: as it was mentioned above (see *Section 2.3*), some items from OS and OH appear to be mistakenly assigned to C2 (Open Data), probably due to the Bibliometrix's inclusion of two secondary keywords ("information systems" and "business modeling"). Some papers in the dataset have not been assigned to any cluster, so that this table is less comprehensive than *Table 5*. This is probably due to the little connection between such papers and the whole corpus.

1.5. Discussion

In the following subsections, The authors discuss emerging questions, open problems, and major insights emerging from the literature. The authors focused mostly on business aspects (models, value creation, value capture), barriers and opportunities, and innovation through open knowledge. Since they are covered in WP1, The authors did not focus much on legal aspects, such as license incompatibility, unless licensing aspects were relevant to business aspects specifically. The information below will help refine the interviews to be performed further in the ZOOOM Project and inform the Ecosystem Report and Legal-Business Framework.

1.5.1. Open Software

Emerging Questions

What approaches and processes are applied by commercial organizations to introduce open-source products in their proprietary products and to provide their software products to the open-source community? What are the main motivations and business incentives for the procedures and processes above?

Major Issues

Emerging from a community-based software development context, Open Source Software (OSS) has become important for firms in a variety of industries, particularly in the context of open innovation. However, the engagement of firms with OSS communities raises concerns about the quality and utility of OSS code, organizational control, and governance. In particular, the shift from proprietary code to OSS requires a reconsideration of the processes that would facilitate value creation and value capture (Morgan et al. 2013).

Business Models

For OSS business models in IT vendor firms, West (2007) found that business buyers (customer segment) enjoyed lower costs and avoided vendor lock-in (value proposition). Because customers expected a richer ‘whole product’ solution than that provided by the OSS project community (integration, customization, support, etc.), such vendors combined priced and unpriced complementary assets to create value.

Watson et al. (2008) distinguish five models of software production or distribution: proprietary, open community, corporate distribution, sponsored OSS, and second-generation OSS. Whereas the first two constitute the extremes of the closed-open continuum, the other three are hybrids of closed and open models:

- *Proprietary Model*: It has dominated the marketplace for decades. Firms employ programmers to develop software, and customers purchase it. The code is considered a major intellectual resource, and traditional software firms protect their code by erecting physical and legal firewalls between their code and the outside world.
- *Open Community Model*: It involves the development and support of software by volunteers with limited or no commercial interest. This model dominates the OSS movement in terms of number of projects.
- *Corporate Distribution*: Quality products are being produced through the open community model, but some entrepreneurs recognize that identifying appropriate products, interacting with open communities for support, and developing the required support skills can be challenging for many potential OSS customers. As a result, firms, such as RedHat, SpikeSource, and OpenOSX, have emerged to create value (and generate revenue) by identifying best-of-breed OSS projects, improving distribution methods for these products, and providing complementary services in order to make these OSS products more accessible to a broader market.
- *Sponsored Open Source*: Corporations and foundations sponsor some OSS projects. For example, the Apache Software Foundation fosters the development of the Apache server and over 50 other OSS projects. Some corporate sponsors directly contribute development resources to OSS projects. IBM is a high-profile example of a corporation contributing developers to Apache’s Web server. In some cases, sponsored OSS projects

have been initiated by corporations releasing previously closed codes and encouraging their employees to continue to work on the now open project. Eclipse, an integrated software development environment, was released as OSS by IBM, whose developers are still primary contributors to the project.

- *Second-Generation Open Source (OSSg2)*: Firms adopting this model are essentially a hybrid between a corporate distribution and sponsored OSS. OSSg2 companies typically generate the bulk of their revenues by providing complementary services around their products but do not sell licenses for their products and own – or tightly control – the software code. OSSg2 models have three main features: accountability, talent based, and ecosystem.

Ecosystems

Feller et al. (2008) investigate the social – rather than legal – mechanisms at the basis of a successful Open Source Service Network (OSSN), an archetype business model that aims to overcome exchange problems in the coordination between firms. Their analysis reveals that, to effectively deliver a "whole product" to customers, OSSNs must enable member firms to access and transfer key strategic resources (e.g., skills/competencies, experiences/knowledge, and customer contacts).

In traditional business networks, legal arrangements are used for coordinating and safeguarding exchanges in order to enable such access to strategic resources. However, the establishment of contracts, and the legal recourse available to participants (litigation) when agreements are violated, incurs substantial overheads in their implementation and lack flexibility.

In contrast, in OSSNs, macroculture (shared assumptions and values that guide actions and create typical behaviour patterns) is central to effectiveness. Similarly, when member firms violate shared norms, values, and goals, the ways in which they are punished (collective sanctions) are a very powerful tool for coordinating exchanges and safeguarding exchanges. At the network design level, restricted access and reputation are also utilized to ensure that only those that are least likely to disrupt the operations of the network (either through incompetence or misbehaviour) become involved; thus facilitating coordinating and safeguarding exchanges.

The authors also point at the importance of restricting network membership for strategic purposes for coordinating exchanges. Coordinating exchanges is facilitated not just by limiting the size of the network (which makes interactions more visible), but by making it easier for members to work together by reducing the variances in the types of participants involved.

Based on case studies such as Mysql, Cendio, Roxen, and SOT, Dahlander & Magnusson (2008) identified three means by which firms exploit communities: (1) accessing communities to extend the resource base; (2) aligning firm strategies with the community; and (3) assimilating communities in order to integrate and share results.

Some firms simply used an adaptive approach, and did not try to change the direction of the development in the communities in any substantial way, but instead focused on using what was developed in the communities and put effort into integrating this work with internally developed components. Others tried to influence community developments by offering incentives to individuals to work with the community, to enhance both its and the firm's reputation, and give the firm more scope for controlling the direction of community developments. By virtue of their positions in the communities, individuals sponsored by firms can help align the firm's strategies through their knowledge of new developments in the community and their ability to influence them. For the firm, this type of effort requires allocation of resources to management of both the firm and the community.

The cases illustrate how firms found it necessary to change their business models to align with the communities. Some firms needed to build a sufficiently large community to create a virtuous development cycle. However, the building and development of communities is not straightforward and just establishing a community does not mean that individuals will necessarily be attracted to becoming members, or that their interest will be sustained over time. Two of the case firms found it difficult to motivate community participants over time, mainly in terms of finding interesting tasks for community participants, and activity in their communities faded away. The nature of the competition also differed among firms, from competition at the product level to fierce competition between communities vying with each other to enrol bright-minded individuals. If a community cannot attract new talent, it is difficult to keep it vital: the development of new functionalities and ideas slows down and the community begins to fade away, which would again necessitate changes in the firms' business models.

When there is a high reliance on external actors, firms need to have the competencies and selection mechanisms to decide which developments are critical – otherwise, there is a risk that the firm will be submerged under a plethora of ideas and suggestions, without the capability to prioritize among them and turn this knowledge into successful products. The reverse can also occur: if too little attention is paid to external actors, firms run the risk of missing out on competitors on critical developments.

Morgan et al. (2013) provide several propositions and constructs for modelling the value creation/capture processes in OSS networks. OSS provides businesses with value through access to knowledge and innovation capacity resident in Value Networks (VNs).[3] VNs enable rich, open, transparent interactions between all members – something that contrasts with the transaction-oriented focus of other traditional inter-organizational relations. In OSS, many firms choose a traditional approach, focusing on bilateral agreements with a limited, familiar number of partners in high-density networks. Nonetheless, partners in OSS high-density VNs are closely acquainted, have formed personal ties, and choose this conventional approach to networking by interacting and collaborating with a limited number of partners. Firms also choose to operate in low-density VNs, always on the lookout for new linkages and relationships to enhance their innovation potential. Knowledge is an extremely important resource that organizations want to take advantage of in these types of networks.

1.5.2. Open Hardware

The concepts of Open Hardware and Open Source Hardware denote tangible artefacts (machines, devices) whose design has been released to the public in such a way that anyone can make, modify, and distribute them. OH projects are about building *platforms*, meaning organizations or meta-organizations that: (1) federate and coordinate agents capable of innovating and competing around the creation of a technology; (2) create value by generating and exploiting economies of scale on the supply or/and demand side; and (3) involve a modular technology architecture composed of a core and a periphery (Viseur & Jullien 2022).

Li et al. (2021) only consider as relevant for OH those start-ups of which the product portfolio includes at least one OH product – the product blueprint, CAD files, software code, and assembly instructions must be available online and licensed under an open-source license. Bonvoisin et al. (2021) focus on tangible and discretely manufactured products of sufficient complexity that are labelled “open source” by their community, thus excluding food, process industry, software products, and components made of single, 3D-printed parts.

Open-source product development (OSPD) projects are described as a combination of five factors:

[3] VNs are entities of several connected individuals/organizational actors that transform and transfer various complementary resources and capabilities.

- a platform (a meeting place for contributors)
- a drive (to motivate contribution)
- a community (the group of contributors)
- a development process
- a business model.

OSPD projects are often also defined by their *degree of openness*, based on three factors: transparency, accessibility, and replicability (see Antoniou et al. 2022; Balka et al. 2010; Bonvoisin et al. 2021; Moritz et al. 2018):

- *Transparency*: The possibility for any interested person to have unrestricted access to information sufficient for understanding the product in detail.
- *Accessibility*: Any person interested is able to actively participate in developing the product by editing design information.
- *Replicability*: the product can be physically reproduced.

Emerging Questions

Why do entrepreneurs choose to open their product design to start their businesses? Is it a strategic decision? Does any “idealistic” or non-commercial goal drive the makers’ decision? How is success defined and measured in OH projects? What is the role of communities and who are the key partners? What about intellectual property and licensing?

Major Issues

A first issue concerns licensing aspects and IP protection in the context of OH. A basic assumption of innovation theory is that the protection of IP grants the inventor market power to exclude other firms from using/imitating the technology without permission, thus capturing the technology’s commercial value. It is especially true for new firms which typically do not have enough resources to contend with imitators. Many studies also have confirmed the effectiveness of IP protection in economic growth and the formation of a new firm (Li et al. 2021).

Licensing is thus a critical issue in OH. It is crucial for all creators and users to know the terms and conditions. The aim of OH licenses is to balance the initial rights of creators of designs derived from copyright law with the principles of the open-source movement. Licenses make sure that ideas may freely circulate and attribution to contributors is given. No one shall be allowed to restrict access to OH.

Prevailing OH licenses (CERN OHL, TAPR OHL) were derived from OSS licenses, but there is a major difference between software and hardware.[4] Copyright covers only the expression of an idea and not the idea itself (or its implementation in a physical product). In OH projects, this is a problem as the aim of an OH project is the implementation of an idea into a useful physical product. Still, the schematics and designs are subject to copyright, the tangible product (and its commercialization), on the other hand, is not. Even though aspects of the distribution of products are considered in some licenses, it is questionable if they would be enforceable in case of infringement (Moritz et al. 2018).

A second issue regards the economic value of OH. Hardware development is considered more complex than software development as the former involves more considerations such as manufacturing, tooling, and supply chain management (Antoniou et al. 2022), which raises the question of how companies can create and capture value from OH.

Moreover, evaluating the economic impact of OH is not easy: contributors in communities are often not getting paid, and people who build products and use them are not buying them from vendors. It is not possible yet to register how many products were built from downloaded designs. So, there is no official registration of value that is being generated/captured besides the revenues of OH businesses. It is thus easier to look at modes of value creation instead of value capture strategies and not to ignore OH projects that create, but do not capture value (Moritz et al. 2018).

Major Insights

Li et al. (2021) interviewed many OH firms and showed that the reasons for going open result from: a) intrinsic factors, such as entrepreneurs' sense of moral obligation, altruism and b) extrinsic motivations, such as market obligations, reduced time-to-market, lowered R&D costs and lowered customer support cost (see *Figure 25*).

[4] See more details in ZOOOM Deliverable D1.1 'Literature review of legal cases in free and open source software, open hardware and open data', p 72-81.

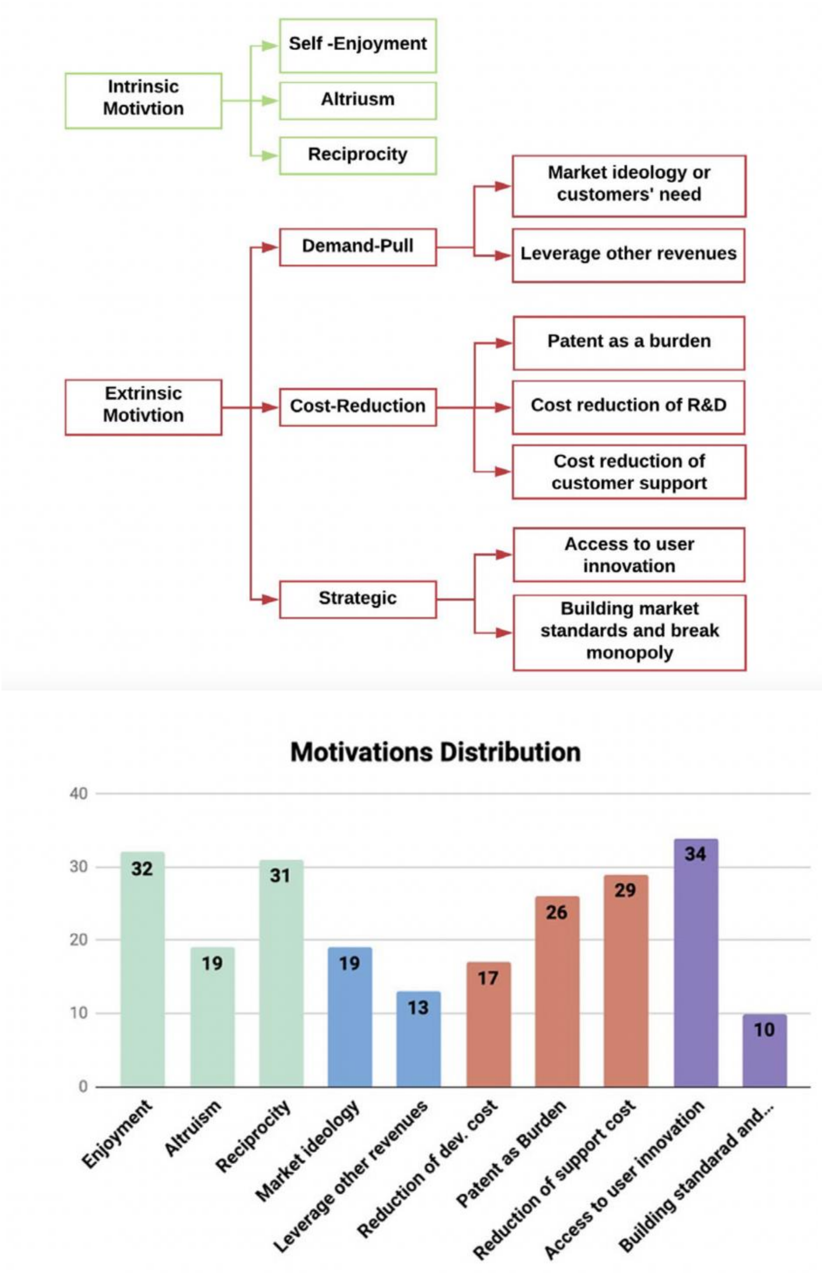

Figure 25 Motivations for openness in OH (from Li et al. 2021).

Antoniou et al. (2022) identify three general characteristics of successful OSH projects: value creation, quality of output, and effective processes. More specifically, successful projects:

- create value for contributors, users, other projects and society
- generate business activity and are sustainable over time

- have a good reputation, which can be demonstrated by the ranking of projects on search engines, the number of projects, documentation, scientific paper citations, the number of views and downloads of project documentation, the number of followers/interested people, and the presence of project communities with a high level of activity (e.g., frequent participation in community forums)
- develop hardware that is highly accessible, reproducible, modifiable, performant, with transparent design, solves a problem/fulfils a need, and offers advantages over alternative products
- create high-quality documentation
- have high process openness
- follow good practices in product development and in project, community, and business management
- are transparent and committed to openness.

Hildebrandt et al. (2022) summarize the main advantages of OH for commercial and private users (see *Figure 26* below).

OSH benefits for commercial use	OSH benefits for private use
• Low purchase prices and operating costs;	• Low purchase prices and operating costs;
• Less training costs because of documentation;	• Low entry barriers because of documentation;
• Fast and cheap support by communities;	• New social community structures;
• Lower R&D-costs and lower or no legal fees;	• Applied education, gain technological literacy;
• Knowledge transfer into own enterprise;	• Resource efficient - option to repair and modify;
• Suitable in minor developed regions.	• Suitable in minor developed regions.

Figure 26 Main Benefits Relating to OH Users (Hildebrandt et al. 2022)

Finally, an interesting legal issue is identified by Moritz et al. (2018) who show that 18% of projects and companies do not use any license at all for their products. This is critical for both makers and users. Entities that use licenses (70%) rely on one of the Creative Commons licenses which provide no effective protection whenever a physical artefact based on copyright-protected documentation of an idea is implemented. Only 17% use proper OH licenses (CERN OHL, TAPR). These licenses cover not only the digital data and documentation, but also deal with aspects regarding the physical implementation, commercial use, and distribution of the products (like a patent). In addition, 5 companies applied for patents (for defensive purposes), which does not necessarily contradict the idea of open source when anyone is granted permission to use the patented idea. Overall, 26 projects and 23 companies (63%) comply with the strict notion of open source by using licenses that explicitly permit commercial use (e.g., CC-BY-SA).

Business Models

Bonvoisin et al. (2021) found that revenue streams of OH projects often come from personal means, external foundation grants, or crowdfunding. However, some projects purposefully adopt a non-commercial strategy aiming merely to sustain their activity. Other projects consider a commercial strategy (e.g., based on product selling) as a way to strengthen their activity as well as the open source movement as a whole. More specifically, Moritz et al. (2018) identified two basic modes of value creation:

- *Design Mode*: One develops a product, shares the design and all relevant information so that other people may make use of it; this mode is mostly used for early-stage projects as a financial investment are low: all you need is a computer, Internet access, and some design skills.
- *Production Mode*: One develops a product, shares the design, manufactures/assembles the product based on that design and sells it as a kit or ready-to-use. Here, capturing value matters: even if the digital representation of the physical product is freely shared, actually making it requires resources (time, money, materials, skills) and access to means of production.

Recently, Viseur & Jullien (2022) identified two types of strategies:

- *Closed Supply-chain Platform* (case: Makerbot): One seeks to control all the components in order to master its architectures, as part of a quality strategy, and therefore favours a closed approach to innovation in order to guarantee the technological consistency of its platform. This involves a razor-blade type business model that relies heavily on the sale of raw materials, but requires strong control over the overall quality of the proposed solutions (the assembly of components), and also over the compatibility between the different elements of the hardware. As a result, the community is excluded from hardware developments. This strategy broadly follows the classic pattern identified in the open-source sector, or the more general framework of open innovation, where the community is exploited temporarily at the launch of the activity, to compensate for the lack of resources of the company, with a gradual closure of the development process, because the company is no longer able to capture the value created by the community. The company then develops (contractual) collaborations with suppliers or its own unitary technologies that guarantee it total control over the innovative solutions it proposes.

- *Open Industry platform* (cases: Prusa Research and Ultimaker): A particular form of innovation ambidexterity, as the companies have developed their capacities to make both incremental innovation and to explore new offers. Then, Ultimaker has even abandoned the open-hardware strategy, while Prusa's open-hardware strategy remains important for it to maintain exploratory capacities. The publication of machine specifications under a free license is necessary to generate contributions, but also to allow users to test configurations and thus report bugs. Finally, and contrary to Makerbot, this strategy does not harm value capture, because the main source of revenue does not come from selling machines to its user-developers, but from printing solutions, for which the company possesses other specific human assets (business experts or experts in 3D technology), which are also expensive to replicate.

Ecosystems

In Bonvoisin et al.'s (2021) study, OH projects range between two archetypes:

- *Isolated Innovators*: They are characterized by a low willingness to co-design. They tend to publish their design only after reaching a first stable state, to strive for transparency and eventually replicability, but not accessibility. Consequently, their surrounding community leans towards a community of followers, replicators or users rather than developers.
- *Development Communities*: They are characterized by a high willingness to co-design with their surrounding community. They appear to adopt an early release policy, both for working documents regarding the product and for the product development process. As a result, the development community usually consists of a core group plus the dynamic, rather unstable participation of community members.

1.5.3. Open Data

The concept of Open Data involves free access to public sector information. Hence, it is often used as a synonym for Open Government Data (OGD). The W3C, the World Bank, and the EU agree that OD must be freely accessible, reusable, and based on open licenses (Chuang et al. 2022). The philosophy behind OD takes inspiration from the OSS movement that has long pushed for computer code to be libre or free (Temiz et al. 2022).

OD are widely considered to have both social and economic value, potentially. Value creation through OGD involves public interest, growth, and securing transparency through data (Cho & Lee, 2020). On the economic side, OD could allow multiple industry sectors to benefit and prosper, including transportation, consumer products, electricity, oil and gas, healthcare, consumer finance, and agriculture. Indeed, entrepreneurs and startups can take advantage of OD to develop new business models, enhance profitability, and improve their overall competitive advantage. OD can also help in the development of new services and products, and boost innovation and the profitability of businesses. According to Temiz et al. (2022), data get immense value as soon as they are combined – for example, by linking customer characteristics to buying behavior – the value can quickly multiply. For example, Facebook and Google are able to target campaigns with unprecedented precision.

Emerging Questions

How to make OGD and Open Corporate Data (OCD) available and reusable for business purposes? How to collect, store, analyse, disseminate, and protect data?

Why and how do open data providers choose and implement open data standards?

Major Issues

Enterprises face remarkable challenges in order to find, access, and select OD. These can make them reluctant to even try. In fact, multiple studies have shown that users find open data's lack of transparency, unknown quality, and unclear licensing unsettling challenges. Krasikov et al (2020) show that less than 50% of analyzed registers provide companies' full legal addresses, while only 10% note their contact information, and conclude that open corporate datasets have only limited use for typical use cases due to their lack of relevant business concepts. Kamariotou and Kitsiosis (2022) summarize the areas where there is a lack of information about the benefits of opening up data and viable business: accessibility and findability of data;

understandability of data; quality; compatibility; linking and merging data; metadata; contact with data providers; openness and the uploading of data; legislation obstacles, technology limitations, information quality. These obstacles describe why firms refuse to share data with developers.

Moreover, organizations refrain from investing in these initiatives due to uncertainty in OD benefits: while data consumers refrain from building services or business models based on open data given its uncertain perpetuity, data providers hesitate to make investments given a lack of evidence of innovation and added value (Enders et al 2021).

Temiz et al. (2022) identified five sets of motives and beliefs about open data: 1) ease and need of use, 2) business potential, 3) capabilities and openness for open data, 4) legitimacy, and 5) business and legal risks. On their view, if open data initiatives are primarily related to factors such as legitimacy-seeking, there is limited incentive to really work towards a productive use of the data. Indeed, while gaining legitimacy may be a necessary condition, it is far from sufficient: if data is primarily opened for the sake of being open, and not for any productive use, the authors are unlikely to see the expected potential come to life. Organizations tend to be overly naïve in believing that OD will automatically lead to value creation. Several complementary investments and capabilities are rather needed. For example, the data must not only be open, but also useful, useable, cleaned, and technically and legally accessible, and it must be matched by investments in information, metadata, software, quality management, and social tools that can cultivate the ecosystem around the open data, in addition to data analytics capabilities.

In short, organizations need to adopt new visions allowing them to have strategic information from this data likely to generate more value. In other words, as Monino (2021) notices, data in itself is not power; it is using them that gives power.

Business Models

Zeleti et al. (2016) provide a taxonomy of business models that are relevant to OD (see Table 4.7).

Table 7 Taxonomy of Open Data Business Models (from Zeleti et al. 2016)

Models	Description
Premium	In the premium business model, the offering is high end open data products and services which are paid for (Huber, 2011).
Freemium	In the freemium model, quality products are provided free of charge for a short period of time after which customers are requested to pay for the products (Teece, 2010).
Open Source	Products in this model are provided in open format that allows free usage and redistribution without any technical barrier (Ferro & Osella, 2013).
Infrastructural Razor and Blades	A razor-blade business model entails selling a product for a low price in order to generate revenues from the complementary products (Pietersz, 2013).
Demand-oriented Platform	This model involves charging consumers of the data products (e.g., developers) for the added value (Howard, 2014).
Supply-oriented Platform	This business model entails the presence of an intermediary business actor providing infrastructural services for data consumers (Ferro & Osella, 2013).
Free as Branded Advertising	This model encourages the audience towards a brand or a company by delivering commercial messages through visualized data which is also called “display advertising” (Ferro & Osella, 2013).
White-label Development	A white-label product is a new product or service developed by one company but acquired and rebranded by another as theirs (Howard, 2014).
Cost Avoidance	This model reduces the cost of data publishing by having a sustainable publishing solution (Epimorphics Ltd, 2012).
Sponsorship	This model entails giving the product for free to customers and obtaining revenue from some sponsors (Casadesus-masanell & Zhu, 2011).
Dual Licensing	Dual licensing is based on the idea of simultaneous use of both open source and proprietary licenses (Välimäki, 2003). Products are given away in an open license for certain purposes and under a closed license for others (Musings, 2012).
Support and Services	This model ensures that the paid packages are given away with guarantees for paying customers (Musings, 2012).
Charging for Changes	In this model, fee is applied for changes made to the product (Musings, 2012).
Increasing Quality Through Participation	This model involves increasing participation to co-creation value with the goal of generating higher margins (Angot, 2010).
Supporting Primary Business	This model entails releasing data towards supporting the primary goals and processes of a business or Organization (Musings, 2012)

According to Enders et al. (2021), firms experience benefits from OD across multiple domains: they contribute to internal improvements, act as an innovation driver, and improve the firm’s visibility towards external stakeholders. As regards BMs, open data can enable new business models for the data provider as well: while open data is revealed free of charge, organizations start implementing freemium payment models. Enders and colleagues found that firms allow access to limited or aggregated open data as a trigger for interested parties to pay for more granular data. Thus, companies establish new business models and thereby monetize their data assets.

Ecosystems

Van Loenen et al. (2021) argued that traditional "one-way street" OD systems should be replaced by OD ecosystems focused not only on data accessibility but also on the larger environment for OD use. The added value of the ecosystem perspective is its focus on the relationships and interdependencies between the social and technological factors that affect the performance of OD activities. In such ecosystems, available data is cycled and recycled among data providers and users. Intermediaries (i.e., actors in-between data providers and users) are essential to OD ecosystems because they form a bridge between the actors and create additional value.

OD ecosystems can be realized at diverging levels, such as the data providers and data user level, or between them at the intermediary's level. Furthermore, the scale of OD ecosystems may vary (within institutions, countries, and regions, as well as within different disciplines and domains). Sustainable OD ecosystem needs to be (1) user-driven, (2) circular, (3) inclusive, and (4) skill-based. Case study analysis by Van Loenen et al (2021) highlights that it is difficult to identify real open data ecosystems: often, existing ecosystems do not balance open data supply and demand, exclude specific user groups and domains, are linear, and lack skill training.

According to Kamarioutou & Kitsios (2022), the lack of a value ecosystem and business models is a crucial barrier to the widespread adoption of available data in services and applications. It is hard for developers to use data to create applications, and they do not have the technical skills to utilize data sources that use complicated data formats or interfaces. Kamarioutou & Kitsios (*Figure 27*) provide an open data ecosystem business model canvas.

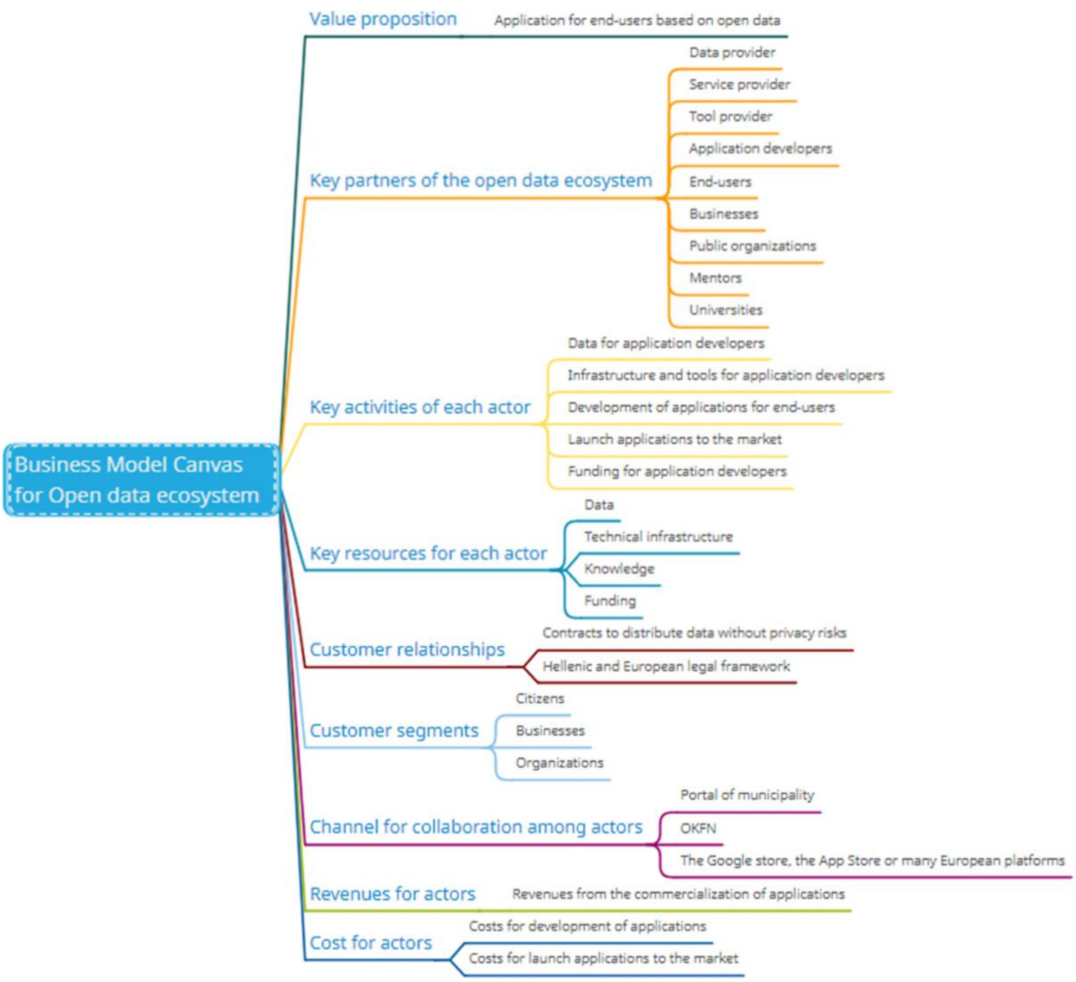

Figure 27 Business model canvas for OD (from Kamarioutou & Kitsios 2022).

1.5.4. Additional References

Some items revealed to be particularly interesting but cannot be easily categorized into one specific domain:

- *Business Models*: Dupart et al. (2022) provide a taxonomy of seven business models that are relevant to the open-source context and discuss how major firms' business models fit such taxonomy.
- *Open Innovation*: Dell'Era et al. (2020) describe how important firms have leveraged openness to achieve strategic and commercial aims. It is relevant to Open Software, but also other fields.

- *Ecosystems*: Haim Faridian & Neubaum (2021) provide a number of propositions for the improvement of exploration- and exploitation-oriented network activities.

2. Professional Literature

2.1. Research Protocol, Sources and Methods

The search for business cases in professional literature was completed by using publicly available sources such as search engines like Google for whitepapers and industry case studies. In addition, company listings by different public agencies (for example, https://www.eu-startups.com/, https://www.seedtable.com/startups-europe, https://startupheatmap.eu/list/, https://app.europeanstartups.co/) were used to identify potential companies (= business cases) according to the three levels of open (3Os) and the emerging technologies (the 4Es). In the search for whitepapers and industry case studies, the following search terms were used:

- The search was limited to companies acting in Europe
  - Very few cases were European, so the cases do involve cases from outside Europe
- Search terms:
  - The three O's of the ZOOOM Project: "open source", "open hardware", "open data"
  - Emerging technologies: "Quantum technologies", "Blockchain technologies", "Artificial Intelligence (AI)" and "Robotics"
- Terms "case", "whitepaper" and "industrial case study" were used as well.

Whitepapers and industrial case studies selected for the analysis were selected based on their relevance to the emerging technology. However, there were not many studies available to incorporate both search terms, so the emerging technologies (the 4Es) were used as a guideline rather than the 3Os.

Table 8 *Number of Selected Whitepapers per Emerging Technology*

Emerging technology	Number of Selected Whitepapers for Analysis
Quantum technologies	17
Blockchain technologies	15
Artificial Intelligence (includes machine learning)	15
Robotics	17

In addition, a search on companies active in the fields of 3Os and 4Es was performed. A list of cases and potential companies in Europe was established and a manual sorting of the case companies was conducted to see if they fit to the original search scope and if the company is still existing. As the emerging technologies used in the search are very new to the marketplace, the list of companies is dominated by start-up companies which evolve quickly and some of the companies listed may not exist anymore. List of suitable companies to be further explored in the project can be found in *Annex 5*. The purpose of the company listing was to provide candidates for the interviews in the next tasks, as well as provide a viewpoint on the companies involved with the technologies.

2.2. 3Os Business Models and Value Creation in Professional Literature

2.2.1. Free and Open Source Software

2.2.1.1. Open Source Based Business Models

Business related to open-source software is perhaps the most well-developed of the 3Os. There are multiple practical case studies (Blind, Knut et al., 2021; Open Source Observatory (OSOR), v1.90.4; Vaughan-Nichols, 2023) related to the business of open source software. The business opportunities can be grouped into these established business models (Blind, Knut et al., 2021; Okoli & Nguyen, 2015):

1. *Auxiliary services*: Selling consultancy services and like training and technical services related to OSS
2. *Corporate development and distribution*: Using corporate resources to develop OSS to accommodate corporate needs, for example hardware compatibility
3. *Software as a Service (SaaS) with distribution of server software*: OSS software components as a part of the SaaS delivery
4. *Dual-licensing/selling exceptions*: Software is released as open source with strong copyleft license, but customers can buy a proprietary license with no copyleft.
5. *Membership and donations*: Donating money for the development effort.
6. *Crowdfunding*: Crowdfunding the development effort.
7. *Advertising*: Advertising with the OSS, either directly in the software or related mediums
8. *Update subscriptions*: Paying for updates.

There are also two, new business models considered as emerging business models:

1. *Selling user data*: OSS components may be used to gather and process user data, or any other valuable data, to be sold later.
2. *Software certification*: Software may be released under an open source license, but some (industrial) use cases require separate certifications in order to be used.

2.2.1.2. The Economics of Open Source and the Role of Copyright Licenses

There is a widespread belief that open source is a business model as such. However, as Amanda Brock has argued, this is not the case and open source was never meant to be a business model (Brock, 2022, p. 329). In the words of Matthew Aslett, open source should be seen as a development and distribution model that is enabled by a licensing tactic, as opposed to a business model (Brock, 2022, p. 342).

As highlighted in ZOOOM Deliverable D1.1, the idea of openness emerged as a reaction to restrictions imposed on sharing and collaboration [5]. This is understandable given that open source is defined in legal terms by reference to the concepts of ownership and licensing (Brock, 2022, p. 330). Indeed, copyright brings software to commercial exchange and licenses are the legal vehicle that creates a relationship between authors and users. They are "the mechanism by which authors give permission to third parties to use their work" (Böhm, 2022, p. 300).

It is the source code, as a subject of intellectual property rights, that links open source as a development model to open source licensing[6]. The protection granted by copyright law to human-readable source code delineates mere information (unprotected) from executable knowledge which is expressed in a human-readable form (protected) (Böhm, 2022, pp. 301–302).

Open source software is integrated as an intermediate or final product into consumer applications and services. As Böhm points out, it is three things simultaneously: state-of-the-art technology, commodity and public good (Böhm, 2022, p. 298). Pragmatically speaking, copyright and licencing are used to restrict the number of available copies of a work (Böhm, 2022, p. 302). This is so because a "good that is available with unlimited supply will converge on a price of zero" (Böhm, 2022, p. 303). Open source uses copyright licensing to enable sharing and provide users with the four essential freedoms. This is the fundamental "hack" on copyright to reverse its impact on the sharing of knowledge (Böhm, 2022, p. 303).

It is a fact that most open source software today is developed in an open development model (Böhm, 2022, p. 307). Open source functionality therefore becomes available to everyone. This is one of the reasons why businesses building open source technologies have struggled to find suitable business models which "succeed in creating additional value at the edge of the commons" (Böhm, 2022, p. 318, citing; Weber, 2005). To generate revenue, open source is typically combined with other approaches, such as dual licensing, trademark licensing programs, patent acquisition, or enforcement of standards covered by standard-essential patents (Böhm, 2022, p. 316).

Business models describe the functioning of a company, how it generates profit and the factors for its success or failure (Schriek et al., 2022, p. 10). Starting companies should assess the viability of different business models, considering particularly their

[5] ZOOOM Deliverable D1.1 'Literature review of legal cases in free and open source software, open hardware and open data', p 8.

[6] ZOOOM Deliverable D1.1 'Literature review of legal cases in free and open source software, open hardware and open data', p 9-10.

competitors in the same market. Scalability is a key factor of success. It tells investors how stable a business is and how likely it is to generate revenue and expand (Schriek et al., 2022, p. 11). Territoriality is another key factor of success. As IP rights are territorial, the business model may not be homogenous across the different countries where a company intends to operate. It may have to adjust to regional or local conditions (Schriek et al., 2022, pp. 11–12).

It has been suggested that ZOOOM authors should distinguish between business models based on goods and business models based on services as the source of revenue generation (Böhm, 2022, p. 315). Indeed, todays it is typical for open source businesses to offer a combination of products and services.

Companies that pursue a product-oriented business strategy may use software as a foundational technology or as a consumer technology (Böhm, 2022, p. 315). In a way, foundational technologies are like means of production. They are the building blocks that provide certain functionality which is then integrated into consumer devices. They do not necessarily satisfy a concrete consumer need and consumers may be and often are completely unaware of them (Böhm, 2022, p. 315). Consumer technologies, in contrast, serve specific needs.

If a company pursues a product-oriented business strategy, it must necessarily exercise control over the source code, e.g., through the use of contributor license agreements (CLAs) (Böhm, 2022, p. 314). These single-vendor business models "require a strong market position usually based on thought leadership and innovativeness that convinces external contributors to participate" (Böhm, 2022, p. 314). Examples of this approach include MySQL, Qt, Asterisk. Alternatively, if a company pursues a service-oriented strategy where the service is related to an open source product, it does not necessarily need exclusive rights over the software but expertise and skills to support code distributed by others.

The service-oriented model can take two forms: vertical integration or complementary services. Vertical integration concerns the integration of open source software into the main product. This is the example of a cloud provider that offers an open source database vertically integrated into the main product, which is, in most cases, running data centres. Complementary services usually concern custom feature development or support and maintenance (Böhm, 2022, p. 315). While vertical integration requires good domain knowledge, complementary services rely on deep technical expertise about the implementation of the software (Böhm, 2022, p. 316).

2.2.1.3. Open Source and Revenue Generation

Revenue generation is distinct from the way open source software is distributed or licensed. Importantly, if no license has been chosen, then the default copyright rules apply, so choosing to be oblivious could have counterproductive outcomes (Brock, 2022, p. 337).

Businesses, especially starting companies, may not be aware of whether going open source is a good choice. They are often tempted by the idea that open source is free of charge (gratis), so using open technologies is seen as a way of reducing upfront costs and generation of value. While this is certainly true, there are key questions that every business should be able to answer comfortably if they are to adopt an open source strategy. Furthermore, businesses should be aware that open source does generate value, but this value is often non-economic (Brock, 2022, p. 337). Amanda Brock has summarized these questions as follows (Brock, 2022, p. 337):

- What if you build your business on software and share it under an Open Source license and somebody else uses it?
- What if you build your business on software and share it under an Open Source license and somebody else uses it to make money?
- What if you build your business on software and share it under an Open Source license and somebody else uses it to make more money than you make or even a lot of money?

In her view, if the business model cannot override any discomfort with the answers to these questions, then the software should not be released under an open source license. Businesses should understand that there is a "risk that a third party may commercialise your code and that 'one day you have a company the next day, that's a feature of a cloud platform" (Brock, 2022, p. 341).

Böhm offers a similar three-step test to help businesses describe the value proposition of any product derived from open source and intended to be place on the market (Böhm, 2022, pp. 326–327):

- What is the revenue model of the product?
- What type of open source-related good is it, e.g., a service (complementary of vertically integrated) or a product (foundational or consumer-oriented)?
- What is different about the offer?

Many open source projects start as community projects before they evolve into a sustainable business. This evolution is about "recognising that there's a clear market gap" (Brock, 2022, p. 340, citing Armon Dadgar, co-founder of Hashicorp). This is why the business or revenue model should be chosen prior to releasing the software under an open source license (Brock, 2022, p. 341).

Studies have recognized business models with vendors selling services around open source, not the code or licenses to the code (Brock, 2022, p. 342). According to Matthew Aslett, these include:

- Commercial licenses, notably dual licensing
- Subscriptions - annual, repeatable support and service agreements
- Service/Support - ad hoc support
- Embedded hardware - software distributed embedded by hardware vendor
- Embedded software - Open Source is embedded within commercial software
- Software as a Service (SaaS)
- Advertising - funded by associated advertising
- Custom development - pay for the software to be customized
- Other products and services.

Similarly, John Koenig refers to seven optimization strategies that include similar business models (Brock, 2022, p. 342):

- Dual licensing
- Support
- Consultancy
- Patronage
- Hosted
- Optimization
- Embedded.

Pureplay open source businesses usually mix open software with services and consultancy. This type of business is usually met with the claim that "support doesn't scale", but as Brock reports, Rancher Labs's founders are proof that a pure support model is possible. A good support model needs two ingredients: useful technology and a quality product (Brock, 2022, pp. 344–345).

A pure copyleft model is another possibility. A good example of a successful story is Nextcloud, whose founder Karlitschek advocates copyleft-based business models because they (as reported by Brock, 2022, p. 345):

- Create the best communities where all contributors are equal
- Imply forks are always possible, which encourages contributors to deliver the best results
- Enable a global upstream/downstream network with no ‘code islands’
- Work with the best developers
- Make ecosystems key because of the level playing field created by the GPL
- Create good community governance
- Remove vendor lock-in
- Imply code ownership is not important because the value is in the people.

Open core models are perhaps the most prevalent business model where open source is foundational. These models are usually referred to as ‘commercial open source’. Most companies whose business model is open core mix software under an open source license and a proprietary license.

Open core models are further classified as tight open core and loose open core. Tight open core models typically offer many of the direct and critical features via proprietary licensing. Maintaining centralized control over source code is therefore essential in a tight core model. In contrast, loose open core concerns features built around the core software. It is these features that are monetized and the software is more useful as a standalone product (Brock, 2022, p. 348).

Enterprise open source and the subscription model are another type of model. In this model, the enterprise “product [that] requires testing, performance tuning, and [to] be proactively examined for security flaws” (Brock, 2022, p. 349, citing Joe Brockmeir). The community may be the main competitor of the business in this model. The subscription model provides add-on services in a similar duality. Red Hat and Tidelift are the best examples of companies that use this model. Importantly, the process of sanitization of the software by a company that assumes responsibility for it and a higher level of liability for some risks can be seen as a revenue source (Brock, 2022, p. 368).

Certification and trademark licensing are another type of model. Moodle is perhaps the best examples of a successful business operating this type of model. Their business model supports the project but achieves the mission to provide the software free of charge by licencing the trademark in a commercial context (Brock, 2022, pp. 349–350).

There are also revenue generation models around foundation and financial fiduciaries. In the first model, foundations may use fiscal sponsors or collaborative or crowdsources financing. One recent example is the GNOME Foundation when it faced patent litigation by Rothschild Patent Imaging.

Finally, there are SaaS business models where users do not get access to code but benefit from access to a service.

Recently, there have been discussions about companies (notably, cloud service providers) using open source technology that have been accused of SaaS squatting. This practice refers to a situation “when a large SaaS or cloud company uses Open Source technology, which is freely available to it but doesn't contribute financially in return despite the Open Source usage generating large revenues, even if this failure to share effectively undermines the company that's developing the Open Source software being used” (Brock, 2022, p. 354).

Examples include MongoDB and their switch from AGPL to the non-open source Server-Side Public License (SSPL). As Brock has argued, the main issue for MongoDB was seemingly their failure to understand that the choice of going open source is irreversible and that open source is not just a good marketing strategy (Brock, 2022, p. 356). Other examples of companies reverting to non-open source licences include Redis Labs and their choice to add the Common Clause to the Apache licence for distribution of the Redis plugins, Cloudera, Confluent, Chef, Cockroach Labs and Elastic (for a discussion of these cases, see Brock, 2022, pp. 354–363).

The conclusion drawn on the basis of these examples is just one: open source cannot be expected to provide the business model of a company. Accusations that certain entities use open source to make (a lot of) money and that this clashes with the principles of open source simply do not hold true. Non-discrimination is one of the basic principles of open source. Indeed, as explained in ZOOOM Deliverable D1.1, the Open Source Definition provides that an open source license must not discriminate against any person or group of persons (Principle 5) and that it must not restrict anyone from making use of the program in a specific field of endeavour (Principle 6). The principle even provides an example that a license may not restrict the program from being used in a business, or from being used for genetic research. In other words, either rights under an open source license exist for all, or they don’t exist for anyone (Brock, 2022, p. 363).

2.2.1.4. Benefits of Open Source for Business

Böhm summarizes the benefits of open source for businesses in three main points (Böhm, 2022, pp. 317–318):

- Open source can be useful without pursuing direct financial benefit, for example, by standardizing the ICT infrastructure and software engineering toolchains
- Open source can be used to generate revenue by being sold, for example, in single-vendor models where one company controls the distribution of source code and uses dual- or multi-licensing approaches
- Open source can be used to reduce the cost of a product, for example, by substituting a proprietary model for an open source implementation (e.g., in embedded systems).

2.2.2. Open Hardware

The rise of Free and Open Source models for software development has catalysed the emergence of “open hardware” or “open source hardware”. The major disadvantage in comparison to open source software is that there will be manufacturing costs involved when using open source hardware. Therefore, it is not easy to compare open source software and open hardware to one another. The differences between them are reflected in the potential business models as well, although there are commonalities for example related to service provisioning. It is also stated that industry virtually ignores OSH for use in commercial products, and contributes little to front-end, back-end, or EDA tools, due to lack of perceived value (Gupta et al., 2016). Some sources also state that the open hardware communities are mostly originating from the US. The separation and differences between open source software and hardware are illustrated below in *Figure 28*.

	Meaningful	Practical	Critical Mass	Deployment
HW vs SW Differences	• Fewer developers and design complexity	• Lack of cheap development infrastructure	• Lack of IP and platforms	• Non-zero deployment cost
Community Role	• Develop more components and increase usability	• Develop effective design tools • Provide free FPGA farms	• Build resusable h/w platforms • Build tools to organize emerging critical mass	• Develop tools for post-manufacture testing
Industry Role	• Contribute more • Develop alternate business models	• Provide freemium tools • Reduce cost and simplify path to chip prototypes	• Take a leap of faith & work with OSH community	• Build customizable appliances

Figure 28 Differences between Hardware and Software, Showing the Role of Industry and Community in Enabling Open Source (Adapted from Gupta et al. 2016).

Business models focusing on open hardware include (Arancio & Molloy, 2021):

- designing the open hardware kits: combining kits of open hardware blueprints and suitable hardware components from selected sources, exampled include Arduino;
- manufacturing the open hardware items: manufacturing of open hardware components, examples include Sparkfun and Adafruit Industries;
- providing services around the items: this is close to equivalent OSS business model where the business is centred around the open core product, services can include for example training and events.

The open hardware movement has been able to progress specifically in small-scale 3D printing / additive manufacturing, including printer manufacturers Lulzbot, Creality, Prusa and Voron Design. In drones, the open hardware movement is also visible in the products of companies or communities like ArduPilo and Flone. The combination of software and hardware to make a drones provides opportunities to use open approach in either of them.

One specific business models which mixes software and hardware concerns the embedded systems domains. In embedded software business models, the source of revenue is savings in building the device and maintenance fees (Brock, 2022, p. 350). Business models for open source software and open hardware can, of course, be very different. There is therefore a “need for EU policy makers to understand better the various differences of OSS and OSH based business models from established business models or licensing frameworks” (Blind, Knut et al., 2021, p. 335).

2.2.3. Open Data

Professional literature relating to open data business models cover mostly public sector information or government data. This is probably due to the history of opening data, which started from the sphere of governmental data. The trend of opening public sector data continues, e.g. through increasing requirements from the European regulations (Open Data Directive) and stakeholder initiatives. An additional concurrent trend is sharing of data, without making it publicly open for everyone, through similar initiatives (European regulation e.g., Data Governance Act, Data Act and Data Spaces

Initiatives) and other market driven initiatives. However, this section leaves out professional literature relating to data sharing and focuses on publicly open data.

2.2.3.1. Open Data Business Models and Value Disciplines

Share-PSI 2.0, the network for innovation in European public sector information, collected best practices needed to build open data business. In their best practices, they have adopted the concept of open data value disciplines introduced by Ahmadi Zeleti et al. (2016) and Zeleti & Ojo (2017), referred to also in the Academic Literature Review above.

The four value disciplines referred to in Zeleti & Ojo (2017) above are:

1. usefulness,
2. process improvement,
3. performance, and
4. customer loyalty.

They are meant to give strategic focus and should be used before defining the business model.

The identification of the value discipline(s) is useful for the organisation to build a best practice for open data business. In addition to the identification of the data discipline(s), an organisation building open data business needs e.g. internal or external expertise; understanding of the existing market, open data products and services; screening of collaborators and competitors; and finding the market niche. (*Share-PSI Best Practice: Open Data Business Models & Value Disciplines*, 2016)

The value disciplines referred to above in Ahmadi Zeleti et al. (2016), are based on the 6-V business model framework consisting of:

1. value proposition
2. value adding process
3. value in return
4. value capture
5. value network
6. value management and related underlying components.

In addition, the following five business model categories were identified based on the framework and analysis of existing open data business models as per Ahmadi Zeleti et al. (2016):

1. freemium
2. premium
3. cost saving
4. indirect benefit
5. parts of tools.

2.2.3.2. Value Proposition Enablers and Roadblockers for Linked Open Government Data

In the Study on business models for linked open government data (inked Government Data (LGD) | Joinup, v1.90.4), value proposition enablers and roadblocks for linked open government data have been identified. These are based on the analysis of 14 case studies.

The study (Linked Government Data (LGD) | Joinup, v1.90.4) identified the following 8 enablers:

1. Efficiency gains in data integration
2. Forward-looking strategies
3. Increased linking and integrated services
4. Ease of model updates
5. Ease of navigation
6. Open licensing and free access
7. Enthusiasm from 'champions'
8. Emerging best practice guidance.

The 7 road blockers identified in the study (Linked Government Data (LGD) | Joinup, v1.90.4) were:

1. Necessary investments
2. Lack of necessary competencies
3. Perceived lack of tools

4. Lack of service level guarantees
5. Missing, restrictive, or incompatible licences
6. Surfeit of standard vocabularies
7. The inertia of the status quo.

2.2.3.3. Sectoral Value and Benefits of Using Open Data

A useful source of information for innovative business models for open data is the project "Open Data 500 U.S." (The Governance Lab, 2023) referred to in the Proceedings of IDIMT 2015 conference paper on Open innovation, open data and new business models. "Open Data 500 U.S." covered 500 US companies using open government data as a key business resource. The project provided analysis of how re-use of open data creates value in different sectors, covering education, transportation, consumer products, electricity value chain, oil and gas value chain, healthcare, and consumer finance (Zimmermann & Pucihar, 2015).

According to Saxena (2014), using open data entails several benefits. One of the most powerful ones is the possibility for disruption of entire industry sectors. In a smaller scale, it can form the basis for new product or service offerings or expand the current ones. The data can be further analysed to provide insight and solve problems. It can even form the basis for data products or data services. Open data can be used internally to make the operations of an organization more efficient or increase the customer understanding.

2.2.3.4. Role of Data in AI Business Models

One interesting area in which open data can be observed to mix with open algorithms is the emerging technology of AI or machine learning. Data and algorithms processing such data are the most essential components generating competitive advantage for AI based business. In a broad sense AI business model archetypes can be identified from a matrix built on the level of openness of data and algorithms, and from such matrix, the following general archetypes can be found (Keller et al., 2018):

1. Proprietary data model (open algorithm, closed data)

   A business restricting access to data that is used in developing the AI system, but using open source AI toolkits and/or giving access to its own algorithms

2. Closed model (closed algorithm, closed data)

   A business residing on closed IP with regard to both algorithms and data

3. Proprietary algorithm model (closed algorithm, open data)

   A business that want to protect their expertise on algorithms at the same time they allow third parties to use their AI model

4. Open model (open algorithm, open data)

   A business finding its competitive advantage from other sources than closed algorithms and closed data, e.g., societal purposes or purposes relating to the development of the AI community

5. Shared model (shared algorithm, shared data)

   A business sharing algorithms and data to selected parties, and finding the competitive advantage from skills, expertise and services

2.3. Business Cases and Potential Value Embedded in 4Es

Our search indicates that practical business cases using both: one of the 3Os and some of the 4Es emerging technologies are very rare. ZOOOM authors were able to identify a list of cases (= companies) who claim to fulfil both conditions in their business, but these companies are still very young, mostly start-up companies. Therefore, extensive knowledge about their business was not available, but the initial identification of the companies outlines potential candidates for interviews in the next tasks of the work package.

In this section, a more general outline of the business potential related to cases in the 4Es emerging technologies are presented based on the results of the practical literature review.

2.3.1. Artificial Intelligence

AI companies can be categorised based on the level of their business using artificial intelligence, and three main categories are stated as:

1. *AI application provider*: Companies whose product simply could not function without AI at its core, whether they serve consumers or enterprises

2. *AI infrastructure provider*: Companies providing AI tools and infrastructure (software & hardware to all companies)
3. *AI adopter*: Companies using AI as part of a broader product or technology stack.

Some details on the potential of AI in the selected industries are detailed on the table below, also addressing the source of value creation potential.

Table 9 Artificial Intelligence: Most Potential Application Areas per Industry Section

AI		
INDUSTRY SECTIONS	**MOST POTENTIAL APPLICATION AREAS**	
Health	**Imaging Technologies**	
	Use case	**Potential value**
	Diagnostics	Faster and more accurate diagnostics
	Health Data Analytics / Diagnostics	
	Use case	**Potential value**
	Virtual drug development	Improved drugs, faster time to market
	Identification of Pandemics	
	Use case	**Potential value**
	AI doctors	Smarter scheduling of appointments and operations
Automotive	**Autonomous Fleets**	
	Use case	**Potential value**
	Ride sharing	On-demand flexibility
	Drivers assist	Safer driving
	Predictive Maintenance	
	Use case	**Potential value**
	Car repairs & maintenance	Improved (resale) value of automobiles
Financial	**Smart Analytics/ Processing**	
	Use case	**Potential value**
	Personalized products	Better Customer Experience
	Credit-checks	Democratization of financial services
	Smart analytics/ processing	Efficiency and robustness of banking processes
	Deep learning and reinforcement learning	New business opportunities
Manufacturing	**Machine Learning, Deep Learning**	
	Use case	**Potential value**
	Intelligent Maintenance	Overall Equipment Effectiveness (OEE)

	Product Quality Control	Avoid loss of production
	Demand Planning	Low space part inventory

2.3.2. Quantum

As a synthesis of the white papers identified in Annex 6, the potential added value from the emerging area of quantum computing and technologies arises from the following three application areas:

1. Quantum computing
    a. Simulation
    b. Machine learning
    c. Optimization
2. Quantum communication, e.g., security, cryptography, encrypting
3. Quantum phenomena and complex system detection.

From the same sources, the authors has identified areas for use cases and the potential value that can be generated from these application areas as follows:

Table 10 Quantum Technology: Most Potential Application Areas per Industry Section

QUANTUM TECHNOLOGY		
INDUSTRY SECTIONS	**MOST POTENTIAL APPLICATION AREAS**	
Health / Pharma	**Quantum Computing**	
	Simulation, Machine learning, Optimization	
	Use case	**Potential value**
	Compound optimization	Better and faster in vitro assessments
		Clinical trial optimization
	Medical diagnoses	More accurate diagnoses
	Molecular and drug solvency simulations	Enhanced drug design
	Quantum Communication	
	Security, Cryptography, Encryption	
	Use case	**Potential value**
	Quantum-safe data handling	Protection of sensitive personal data
		Avoiding data breaches
	Quantum phenomena and sensing	
	Detecting quantum phenomena and complex systems	
	Use case	**Potential value**
	Detection of molecular structure	Faster and more efficient drug discovery
		Better biological target identification
		Compound screening
		Entry into new markets
		Faster and more efficient R&D
	Genome sequencing	Personalized medicines
		New and more efficient treatments
		Increase of quality of life
Financial services	**Quantum Computing**	
	Simulation, Machine learning, Optimization	
	Use case	**Potential value**
	Portfolio optimization and management	Better securities and derivatives valuation
		Better asset analysis and allocation
		More sustainable investment portfolios
	Risk assessment and management	Better market simulation
		Better scenario analyses
		Lower risks
	Pattern recognition and irregular behaviour	Quality and speed of fraud detection
	Synthetic data generation and market simulation	Better and faster financial forecasts
		Better trading strategies
		Recognition of market instabilities

	Credit scoring models	Reducing margin of errors
		Better understanding of risks
		More accurate credit scorings
		Profitability of credit offers
	Quantum Communication	
	Security, Cryptography, Encryption	
	Use case	**Potential value**
	Quantum readiness	Secure communication
		Quantum-proof encryption
Mobility / Transport / Logistics	**Quantum Computing**	
	Simulation, Machine learning, Optimization	
	Use case	**Potential value**
	Route planning and traffic optimization	Real-time and multimodal traffic management
		Better weather forecasting
	Management of fleet and autonomous vehicles	Better management of complex systems
	Shipping, freight and cargo optimization	Better network scheduling
		Optimal cargo loading
		Accurate freight forecasting
	Crash simulation	Safety of traffic
Energy	**Quantum Computing**	
	Simulation, Machine learning, Optimization	
	Use case	**Potential value**
	Modelling of energy processes, climate and weather	Better weather predictions
		More efficient solar cells
	Energy network and grid management	More efficient use of energy
		Optimal network design and energy distribution
	Energy production	Optimal combinations of different energy sources
		Optimal locations to produce energy
	Quantum Phenomena and Sensing	
	Detecting quantum phenomena and complex systems	
	Use case	**Potential value**
	Generation and storage of electricity	Novel batteries
		Better energy storage systems
	Prediction of nuclear reactions	Next generation nuclear power
	Exploration of energy sources	Enhanced energy source exploration
Manufacturing / Chemical industry / Materials /	**Quantum Computing**	
	Simulation, Machine learning, Optimization	
	Use case	**Potential value**

Product design	Simulation of quantum interaction (atoms and molecules)	Green hydrogen, hydrogen-powered steel production
		Replacement of petrochemicals
		Room temperature superconductors
		High-efficiency data storage and transfer
		Energy efficiency and safety
	Design of new molecules	New catalysts
		Increase of energy efficiency (batteries, solar panels)
	Production processes	Minimizing by-products
		Optimizing processes, yields, welding, painting
		Reducing resource use
	Supply chain management	Improved supply chains
	Product design	Improved and high-quality design
		Faster design
		Mesh optimization
		High-quality fluid simulation
		Predictive maintenance
	Quantum Phenomena and Sensing	
	Detecting quantum phenomena and complex systems	
	Use case	**Potential value**
	Novel discoveries	Improved R&D
	Material development	Better materials, fabrics, coatings
		Novel materials, fabrics, coatings
	Carbon sequestration	Solving complex global sustainability challenges
		Decreasing energy costs
Agriculture / Food	**Quantum Computing**	
	Simulation, Machine learning, Optimization	
	Use case	**Potential value**
	Agricultural methods	Improved methods
	Optimizing pesticides and chemicals	Environment friendly use of pesticides and chemicals
	Simulating plant genome	Higher yields
		Carbon footprint minimization
	Quantum phenomena and sensing	
	Detecting quantum phenomena and complex systems	
	Use case	**Potential value**
	Previously intractable phenomena	Energy efficiency of fertilizer production
Other	**Quantum Computing**	

	Simulation, Machine learning, Optimization	
	Use case	**Potential value**
	Democratizing AI	Level playing field for SMEs
		Novel business models
	Marketing	Optimization of advertising campaigns
	Quantum Communication	
	Security, Cryptography, Encryption	
	Use case	**Potential value**
	IoT data and data sharing	Quantum-safety
	Quantum phenomena and sensing	
	Detecting quantum phenomena and complex systems	
	Use case	**Potential value**
	Defence	Sensors for better detection of e.g. submarines

2.3.3. Blockchain

According to the European Union Blockchain Observatory, the biggest impact related to blockchain is anticipated in the following fields:

- supply chain
- government services
- finance
- IoT
- healthcare
- media
- smart cities
- energy
- legal aspects
- cryptography.

As a synthesis of the white papers identified in *Annex 6*, the potential use cases and value are:

Table 11 Blockchain: Most Potential Application Areas per Industry Section

BLOCKCHAIN

INDUSTRY SECTIONS	MOST POTENTIAL APPLICATION AREAS	
Health / Pharma	**Blockchain**	
	Distributed append-only databases, Byzantine fault-tolerance, data exchange interfaces, zero-knowledge proofs	
	Use case	**Potential value**
	Identity management	Better control over data privacy in medical records
	Supply chain provenance	Improved drug safety
	Medical records	Data ownership and improved access
	Disease control	Improved reliability of information in pandemics
	Concurrency control and immutability of records	Improved auditability in drug testing
Financial services	**Blockchain**	
	Distributed append-only databases, cryptocurrency tokens, smart contracts, zero-knowledge proofs	
	Use case	**Potential value**
	Payment processing	Improved customer experience
	Remittance	Democratization of financial services
	Investment contracts	New business opportunities
	Cross-border collateral movement	Efficiency and robustness of banking processes
	Post trade processing and settlement of securities	Improved efficiency and transparency
	Central bank digital currency	More robust monetary system, more versatile fiscal policy
Mobility / Transport / Logistics	**Blockchain**	
	Distributed append-only databases, Byzantine fault-tolerance, RFID-tags	
	Use case	**Potential value**
	Quality management	Improved supply chain risk management
	Provenance	Fair trade through improved transparency
	Transaction transparency	Improved consumer awareness
	Prevention of double marginalization	Improved market efficiency
	Bills of lading	Improved freight finance
	Freight bills	Improved freight finance
	Warehouse management	Waste reduction, improved transport efficiency
Energy / Sustainability	**Blockchain**	
	Distributed append-only databases, cryptocurrency tokens, smart contracts, tokenized assets	
	Use case	**Potential value**
	Visibility to alignment with sustainability goals	Efficient monitoring of sustainability actions
	Financing smart infrastructure	Mobilizing new sources of financing

	Emissions certificate trading / carbon off-setting	Reducing greenhouse gas emissions
	Power sharing economy (P2P)	Improved demand response and energy efficiency
	Certificates of origin for green energy	Efficient monitoring of sustainability actions
Internet of Things	**Blockchain**	
	Distributed append-only databases, smart contracts, data exchange interfaces, RFID tags, sensors	
	Use case	**Potential value**
	Product lifecycle management	Overall equipment effectiveness
		Lower maintenance costs
		Reduction of waste, avoid loss of production
	ICT system integration	Improved accessibility of information
		Low space part inventory
Public sector	**Blockchain**	
	Distributed append-only databases, smart contracts, distributed autonomous organizations	
	Use case	**Potential value**
	Identity management	Improved access to public services
	Notarization	More secure public record
	e-Government	More resilient public services
	Real-time e-voting	Improved transparency of public elections
	Land title registry	More secure public record
Other	**Blockchain**	
	Distributed append-only databases, smart contracts, non-fungible tokens	
	Use case	**Potential value**
	Art work licensing (media)	Verifying the authenticity of works of art
	Settling royalty payments (media)	Improved efficiency in payment processing

2.3.4. Robotics

Based on the European Union'n description on robotics (Robotics | Shaping Europe's Digital Future, 2022; The Impact Of Robotics On The European Economy, 2022), the impact of robotics lies within the manufacturing sector where already millions of jobs are established and robotics has potential to enhance productivity of the industry into new levels by turning the traditional industry into cyber-physical organization with the use of 5G, IoT and data. In addition to manufacturing industries, robotics offers potential new solutions to societal challenges from ageing to health, smart transport, security, energy and environment.

In the table below, The authors has gathered data from different whitepapers and industry reports (*Annex 6*) to highlight the value creation opportunities with use cases in different industries.

Table 12 Robotics: Most Potential Application Areas per Industry Section

ROBOTICS		
INDUSTRY SECTIONS	**MOST POTENTIAL APPLICATION AREAS**	
Health	**Robotised Surgery**	
	Use case	**Potential value**
	Surgery by robotic system	Improved accuracy
		Increased stability (tremor reduction)
		Scale motion
		Increased patient recovery
		Sterilization and resistance to radiation and infection
	Intelligent Prosthetics	
	Use case	**Potential value**
	Prosthetics mimic human functionality through artificial muscles, joints or skeleton parts	Functionality of a natural limb
	Robotized Motor Coordination Analysis and Therapy	
	Use case	**Potential value**
	Robot therapy systems	Controlled support of movement helping the human in making the movement or straining the movement
Manufacturing	**Warehouse Management**	
	Use case	**Potential value**
	Robotic operators	Improved productivity
		Enhanced employee satisfaction, attractiveness, and retention
	Manufacturing Automation robots	
	Use case	**Potential value**
	Manufacturing robots	Improved efficiency
Logistics	**Autonomous Guided Vehicles**	
	Use case	**Potential value**
	Assembly	Moving products through production processes
	Kitting	Collecting parts for assembly
	Transportation	Loading pallets and loose parts
	Staging	Delivering pallets for production processes
Maintenance	**Floor-cleaning Robots**	
	Use case	**Potential value**

	Floor-cleaning robots	Cleaning and disinfection of surfaces
Retail / Restaurants	**Automated checkout and self-service**	
	Use case	**Potential value**
	Automated checkout in retail and self-service kiosks	Automation in retail and restauration

References

Academic Literature Review

Antoniou, R., Bonvoisin, J., Hsing, P.-Y., Dekoninck, E., & Defazio, D. (2022). Defining success in open source hardware development projects: A survey of practitioners. Design Science, 8, e8. https://doi.org/10.1017/dsj.2021.30

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of informetrics, 11(4), 959-975.

Arun, R., Suresh, V., Veni Madhavan, C. E., & Narasimha Murthy, M. N. (2010). On finding the natural number of topics with latent dirichlet allocation: Some observations. In Advances in Knowledge Discovery and Data Mining: 14th Pacific-Asia Conference, PAKDD 2010, Hyderabad, India, June 21-24, 2010. Proceedings. Part I 14 (pp. 391-402). Springer Berlin Heidelberg.

Balka, K., Raasch, C., & Herstatt, C. (2010). How Open is Open Source? - Software and Beyond: HOW OPEN IS OPEN SOURCE? Creativity and Innovation Management, 19(3), 248–256. https://doi.org/10.1111/j.1467-8691.2010.00569.x

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. Journal of machine Learning research, 3, 993-1022.

Bonaccorsi, A., Giannangeli, S., & Rossi, C. (2006). Entry Strategies Under Competing Standards: Hybrid Business Models in the Open Source Software Industry. Management Science, 52(7), 1085–1098. https://doi.org/10.1287/mnsc.1060.0547

Bonarini, A., Matteucci, M., Migliavacca, M., & Rizzi, D. (2014). R2P: An open source hardware and software modular approach to robot prototyping. Robotics and Autonomous Systems, 62(7), 1073–1084. https://doi.org/10.1016/j.robot.2013.08.009

Bonvoisin, J., Thomas, L., Mies, R., Stark, R., Samuel, K., Jochem, R., Boujut, J.-F., & Gros, C. (2021). Current state of practices in open source product development.

Casadesus-Masanell, R., & Ghemawat, P. (2006). Dynamic Mixed Duopoly: A Model Motivated by Linux vs. Windows. Management Science, 52(7), 1072–1084. https://doi.org/10.1287/mnsc.1060.0548

Cho, J. Y., & Lee, B. G. (2022). Creating value using public big data: Comparison of driving factors from the provider's perspective. Information Technology & People, 35(2), 467–493. https://doi.org/10.1108/ITP-04-2019-0169

Chou, C. (2021). Could coexistence of open-source and proprietary platforms be an equilibrium outcome? The Manchester School, 89(3), 297–309. https://doi.org/10.1111/manc.12359

Chuang, T.-R., Fan, R. C., Ho, M.-S., & Tyagi, K. (2022). Openness. Internet Policy Review, 11(1). https://doi.org/10.14763/2022.1.1643

Corrales-Garay, D., Ortiz-de-Urbina-Criado, M., & Mora-Valentín, E.-M. (2022). Understanding open data business models from innovation and knowledge management perspectives. Business Process Management Journal, 28(2), 532–554. https://doi.org/10.1108/BPMJ-06-2021-0373

Dahlander, L., & Magnusson, M. (2008). How do Firms Make Use of Open Source Communities? Long Range Planning, 41(6), 629–649. https://doi.org/10.1016/j.lrp.2008.09.003

de Laat, P. B. (2005). Copyright or copyleft? Research Policy, 34(10), 1511–1532. https://doi.org/10.1016/j.respol.2005.07.003

Dell'Era, C., Di Minin, A., Ferrigno, G., Frattini, F., Landoni, P., & Verganti, R. (2020). Value capture in open innovation processes with radical circles: A qualitative analysis of firms' collaborations with Slow Food, Memphis, and Free Software Foundation. Technological Forecasting and Social Change, 158, 120128. https://doi.org/10.1016/j.techfore.2020.120128

Duparc, E., Möller, F., Jussen, I., Stachon, M., Algac, S., & Otto, B. (2022). Archetypes of open-source business models. Electronic Markets, 32(2), 727–745. https://doi.org/10.1007/s12525-022-00557-9

Enders, T., Benz, C., & Satzger, G. (2021). Untangling the Open Data Value Paradox: How Organizations Benefit from Revealing Data. In F. Ahlemann, R. Schütte, & S. Stieglitz (Eds.), Innovation Through Information Systems (Vol. 48, pp. 200–205). Springer International Publishing. https://doi.org/10.1007/978-3-030-86800-0_15

Feller, J., Finnegan, P., Fitzgerald, B., & Hayes, J. (2008). From Peer Production to Productization: A Study of Socially Enabled Business Exchanges in Open Source Service Networks. Information Systems Research, 19(4), 475–493. https://doi.org/10.1287/isre.1080.0207

Fjeldsted, A. S., Adalsteinsdottir, G., Howard, T. J., & McAloone, T. (2012). Open Source Development of Tangible Products.

Gao, Y., & Janssen, M. (2020). Generating Value from Government Data Using AI: An Exploratory Study. In G. Viale Pereira, M. Janssen, H. Lee, I. Lindgren, M. P. Rodríguez Bolívar, H. J. Scholl, & A. Zuiderwijk (Eds.), Electronic Government (Vol. 12219, pp. 319–331). Springer International Publishing. https://doi.org/10.1007/978-3-030-57599-1_24

Gao, Y., & Janssen, M. (2022). The Open Data Canvas–Analyzing Value Creation from Open Data. Digital Government: Research and Practice, 3(1), 1–15. https://doi.org/10.1145/3511102

Gruber, M., & Henkel, J. (2006). New ventures based on open innovation an empirical analysis of start-up firms in embedded Linux. International Journal of Technology Management, 33(4), 356. https://doi.org/10.1504/IJTM.2006.009249

Haim Faridian, P., & Neubaum, D. O. (2021). Ambidexterity in the age of asset sharing: Development of dynamic capabilities in open source ecosystems. Technovation, 99, 102125. https://doi.org/10.1016/j.technovation.2020.102125

Hildebrandt, L., Moritz, M., Redlich, T., & Wulfsberg, J. P. (2022). Open Source Hardware and Decentralized Urban Production for Urgently Needed Products during the COVID-19 Pandemic. 2022 Portland International Conference on Management of Engineering and Technology (PICMET), 1–10. https://doi.org/10.23919/PICMET53225.2022.9882854

Höst, M., & Oručević-Alagić, A. (2011). A systematic review of research on open source software in commercial software product development. Information and Software Technology, 53(6), 616–624. https://doi.org/10.1016/j.infsof.2010.12.009

Janssen, M., & Zuiderwijk, A. (2014). Infomediary Business Models for Connecting Open Data Providers and Users. Social Science Computer Review, 32(5), 694–711. https://doi.org/10.1177/0894439314525902

Kamariotou, M., & Kitsios, F. (2022). Bringing Digital Innovation Strategies and Entrepreneurship: The Business Model Canvas in Open Data Ecosystem and Startups. Future Internet, 14(5), 127. https://doi.org/10.3390/fi14050127

Kolade, O., Adepoju, D., & Adegbile, A. (2022). Blockchains and the disruption of the sharing economy value chains. Strategic Change, 31(1), 137–145. https://doi.org/10.1002/jsc.2483

Krasikov, P., Obrecht, T., Legner, C., & Eurich, M. (2020). Is Open Data Ready for Use by Enterprises? Learnings from Corporate Registers: Proceedings of the 9th International Conference on Data Science, Technology and Applications, 109–120. https://doi.org/10.5220/0009875801090120

Kumar, A., Gupta, A., Sanagavarapu, L. M., & Reddy, Y. R. (2022). An approach to Open-Source Software License Management using Blockchain-based Smart-Contracts. 15th Innovations in Software Engineering Conference, 1–5. https://doi.org/10.1145/3511430.3511448

Li, Z., & Seering, W. (2019a). Does Open Source Hardware Have a Sustainable Business Model? An Analysis of Value Creation and Capture Mechanisms in Open Source Hardware Companies. Proceedings of the Design Society: International Conference on Engineering Design, 1(1), 2239–2248. https://doi.org/10.1017/dsi.2019.230

Li, Z., & Seering, W. (2019b). Build Your Firm With Strangers?: Longitudinal Studies on Open Source Hardware Firm Growth. Volume 7: 31st International Conference on Design Theory and Methodology, V007T06A027. https://doi.org/10.1115/DETC2019-97618

Li, Z., Seering, W., Yang, M., & Eesley, C. (2021). Understanding the motivations for open-source hardware entrepreneurship. Design Science, 7, e19. https://doi.org/10.1017/dsj.2021.15

Mariani, M. M., Perez-Vega, R., & Wirtz, J. (2022). AI in marketing, consumer research and psychology: A systematic literature review and research agenda. Psychology & Marketing, 39(4), 755-776.

Monino, J.-L. (2021). Data Value, Big Data Analytics, and Decision-Making. Journal of the Knowledge Economy, 12(1), 256–267. https://doi.org/10.1007/s13132-016-0396-2

Morgan, L., Feller, J., & Finnegan, P. (2013). Exploring value networks: Theorising the creation and capture of value with open source software. European Journal of Information Systems, 22(5), 569–588. https://doi.org/10.1057/ejis.2012.44

Moritz, M., Redlich, T., & Wulfsberg, J. (2018). Best practices and pitfalls in open source hardware. In Proceedings of the International Conference on Information Technology & Systems (ICITS 2018) (pp. 200-210). Springer International Publishing.

Pisano, G. (2006). Profiting from innovation and the intellectual property revolution. Research Policy, 35(8), 1122–1130. https://doi.org/10.1016/j.respol.2006.09.008

Rajala, R., Westerlund, M., & Möller, K. (2012). Strategic flexibility in open innovation – designing business models for open source software. European Journal of Marketing, 46(10), 1368–1388. https://doi.org/10.1108/03090561211248071

Sharma, S., Sugumaran, V., & Rajagopalan, B. (2002). A framework for creating hybrid-open source software communities. Information Systems Journal, 12(1), 7–25. https://doi.org/10.1046/j.1365-2575.2002.00116.x

Singh, P. V., & Phelps, C. (2013). Networks, Social Influence, and the Choice Among Competing Innovations: Insights from Open Source Software Licenses. Information Systems Research, 24(3), 539–560. https://doi.org/10.1287/isre.1120.0449

Stewart, K. J., Ammeter, A. P., & Maruping, L. M. (2006). Impacts of License Choice and Organizational Sponsorship on User Interest and Development Activity in Open Source Software Projects. Information Systems Research, 17(2), 126–144. https://doi.org/10.1287/isre.1060.0082

Teixeira, J., Robles, G., & González-Barahona, J. M. (2015). Lessons learned from applying social network analysis on an industrial Free/Libre/Open Source Software ecosystem. Journal of Internet Services and Applications, 6(1), 14. https://doi.org/10.1186/s13174-015-0028-2

Temiz, S., Holgersson, M., Björkdahl, J., & Wallin, M. W. (2022). Open data: Lost opportunity or unrealized potential? Technovation, 114, 102535. https://doi.org/10.1016/j.technovation.2022.102535

Välimäki, M., & Oksanen, V. (2005). The impact of free and open source licensing on operating system software markets. Telematics and Informatics, 22(1–2), 97–110. https://doi.org/10.1016/j.tele.2004.06.008

Van Loenen, B., Zuiderwijk, A., Vancauwenberghe, G., Lopez-Pellicer, F. J., Mulder, I., Alexopoulos, C., Magnussen, R., Saddiqa, M., Dulong de Rosnay, M., Crompvoets, J., Polini, A., Re, B., & Casiano Flores, C. (2021). Towards value-creating and sustainable open data ecosystems: A comparative case study and a research agenda. JeDEM - EJournal of EDemocracy and Open Government, 13(2), 1–27. https://doi.org/10.29379/jedem.v13i2.644

Viseur, R. (2012). From Open Source Software to Open Source Hardware. In I. Hammouda, B. Lundell, T. Mikkonen, & W. Scacchi (Eds.), Open Source Systems: Long-Term Sustainability (Vol. 378, pp. 286–291). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-33442-9_23

Viseur, R., & Jullien, N. (2022). 3D printer builders' Open-hardware strategies. The 18th International Symposium on Open Collaboration, 1–8. https://doi.org/10.1145/3555051.3555069

Watson, R. T., Boudreau, M.-C., York, P. T., Greiner, M. E., & Wynn, D. (2008). The business of open source. Communications of the ACM, 51(4), 41–46. https://doi.org/10.1145/1330311.1330321

West, J. (2007). Value Capture and Value Networks in Open Source Vendor Strategies. 2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07), 176–176. https://doi.org/10.1109/HICSS.2007.600

Zeleti, F., Ojo, A., & Curry, E. (2016). Exploring the economic value of open government data. Government Information Quarterly, 33(3), 535–551. https://doi.org/10.1016/j.giq.2016.01.008

Zhao, D., & Strotmann, A. (2015). Analysis and visualization of citation networks. Synthesis Lectures on Information Concepts, Retrieval, and Services, 7(1), 1–207.

Professional Literature

Arancio, J., & Molloy, J. (2021). Open hardware is ready to help technology transfer offices maximise the impact of academic research. Google Docs. https://drive.google.com/file/d/1E9p0_HOj9Q3QHsIr8I-NlD2bYkYPxiKD/view?usp=embed_facebook

Blind, Knut, Böhm, Mirko, Grzegorzewska, Paula, Katz, Andrew, Muto, Sachiko, Pätsch, Sivan, & Schubert, Torben. (2021). The impact of Open Source Software and Hardware on technological independence, competitiveness and innovation in the EU economy [Final Study Report.]. Publications Office of the European Union. https://ec.europa.eu/newsroom/dae/redirection/document/79021

Brock, A. (2022). Business and Revenue Models and Commercial Agreements. In A. Brock (Ed.), Open Source Law, Policy and Practice (New Edition, Second Edition, New Edition, Second Edition). Oxford University Press. https://global.oup.com/academic/product/open-source-law-policy-and-practice-9780198862345?cc=gb&lang=en&

Böhm, M. (2022). Economics of Open Source. In A. Brock (Ed.), Open Source Law, Policy and Practice (New Edition, Second Edition, New Edition, Second Edition). Oxford University Press. https://global.oup.com/academic/product/open-source-law-policy-and-practice-9780198862345?cc=gb&lang=en&

Gupta, G., Nowatzki, T., Gangadhar, V., & Sankaralingam, K. (2016). Open-source Hardware: Opportunities and Challenges. CoRR, abs/1606.01980. http://arxiv.org/abs/1606.01980

Linked Government Data (LGD) | Joinup. (v1.90.4). Linked Government Data (LGD). What Is Linked Government Data (LGD)? Retrieved March 16, 2023, from https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/linked-government-data-lgd

Robotics | Shaping Europe's digital future. (2022, December 6). Robotics. https://digital-strategy.ec.europa.eu/en/policies/robotics

Saxena, A. (2014, December 18). 5 Ways Open Government Data Can Inspire Startup Innovation. Entrepreneur. https://www.entrepreneur.com/science-technology/5-ways-open-government-data-can-inspire-startup-innovation/240187

Share-PSI Best Practice: Open Data Business Models & Value Disciplines. (2016, July 25). Best Practice: Open Data Business Models & Value Disciplines. https://www.w3.org/2013/share-psi/bp/odbm/ Retrieved January 18, 2023, from https://www.w3.org/2013/share-psi/bp/odbm/

Schriek, U., Juette, M. F., & Golkowsky, S. (2022). Basics of intellectual property rights with a focus on start-ups. In S. Golkowsky (Ed.), Intellectual property strategies for start-ups: A practical guide. Edward Elgar Publishing.

The Governance Lab. (2023, March 16). Open Data 500 Global Network. How to Make Open Government Data More Relevant, Accessible, and Actionable? Retrieved March 16, 2023, from https://thegovlab.org/project/project-open-data-500-global-network

The Impact Of Robotics On The European Economy. (2022, February 8). The European Business Review. https://www.europeanbusinessreview.com/the-impact-of-robotics-on-the-european-economy/

Zeleti, F. A., & Ojo, A. (2017). The 6-Values Open Data Business Model Framework. In A. Ojo & J. Millard (Eds.), Government 3.0 – Next Generation Government Technology Infrastructure and Services (Vol. 32, pp. 219–239). Springer International Publishing. https://doi.org/10.1007/978-3-319-63743-3_9

Zimmermann, H.-D., & Pucihar, A. (2015). Open Innovation, Open Data and New Business Models [SSRN Scholarly Paper]. https://doi.org/10.2139/ssrn.2660692

Keller, J. R., Chauvet, L., Fawcett, J., & Thereaux, O. (2018). #OPEN The role of data in AI business models. Google Docs. Retrieved January 17, 2023, from https://docs.google.com/document/d/14g0p6KSyH1r1J_PrykJIXUX-rdeP1B4CLIffAyFPOnk/edit?usp=embed_facebook

Emanuilov, I., Lasota, L., Singhal, N., Suksi, J., Zirnstein, E. (2023). ZOOOM Deliverable 1.1 'Literature review of legal cases in free and open source software, open hardware and open data' (unpublished)

This document is an output of the ZOOOM Project

(GA 101007385)

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them.