D4.2 Drivers and barriers to uptake of Open Science resources in industry

Observing and Negating Matthew Effects in Responsible Research and Innovation Transition

Version 1.0
Public

This deliverable presents the insights gained from interviews and a questionnaire about drivers and barriers of the uptake of Open Science resources in SMEs and industry, including a summary of recommendations for action to inform policy advice.

ON-MERRIT - Grant Agreement 824612

The research leading to these results has received funding from the European Union's Horizon 2020 Research and Innovation Programme, under Grant Agreement no 824612.

Document Description

D4.2 - Drivers and barriers to uptake of Open Science resources in industry

D4.2 Drivers and barriers to uptake of Open Science resources in industry			
WP4 - Innovation and Industry			
Due date	30.9.2021	Actual delivery date:	30.9.2021
Nature of document	Report	Version	1.0
Dissemination level	Public		
Lead Partner for deliverable	Know-Center GmbH		
Authors	Angela Fessl (KNOW), Katharina Maitz (KNOW)		
Reviewers	Antónia Correia (UMINHO), Nicki Lisa Cole (TUG), Tony Ross-Hellauer (KNOW)		

Revision History

Issue	Item	Comments	Author/Reviewer
V 0.1	Draft version	Initial Draft	Angela Fessl, Katharina Maitz
V 0.1	Review	Review	Antónia Correia, Nicki Lisa Cole, Tony Ross-Hellauer
V 0.2	Final version	Integration the reviewers' inputs	Angela Fessl, Katharina Maitz

Table of Contents

Tables

Figures

Abbreviations

CEO	Chief Executive Officer
CTO	Chief Technology Officer
R&D	Research and Development
RQ	Research Question
RRI	Responsible Research and Innovation
SME	Small and Medium Enterprises

Executive summary

Open Access to scientific results such as publications or data is a key goal of policy-makers to make scientific results reusable and to facilitate growth and innovation in SMEs and industries. However, it is still unclear if and how open research results are actually being taken up in the private sector (Huber, Wainwright, and Rentocchini 2020; Fell 2019; Europäische Kommission 2016). This deliverable report addresses this issue by presenting the results from an interview study and a questionnaire study.

The deliverable report D4.1 - Information Seeking Behaviour and Open Science Uptake in Industry: A Literature Review (Fessl et al. 2020) served as the starting point for this deliverable report D4.2. In T4.1 of ON-MERRIT, we conducted an extensive literature review on the current state of the art regarding information seeking behaviour in SMEs and industries, as well as the current status of the uptake of Open Science resources in this regard. Our findings with regard to T4.1 were reported in D4.1.

Based on the insights gained from D4.1, we started our investigation of the uptake of responsible research and innovation (RRI) and Open Science in SMEs/industries by conducting an interview study in Austria. We then followed up this work with a questionnaire study across Europe. Our overall findings show that i) individuals with a university education are more familiar with the concept of Open Science than others, and that ii) Open Science resources already play an important role in the companies (of our interview partners), but their uptake depends on the company characteristics, including the company's domain, and the products and services offered.

From both studies, we identified drivers that support the uptake of Open Science resources: i) the employment of people with a university background, ii) offering incentives and support for uptake, iii) offering targeted training to increase uptake, iv) learning from trans- and interdisciplinary cooperations, and v) exploiting the wisdom of the crowd. We also identified the following barriers that hinder the uptake of Open Science resources: i) scarcity of health-related data, ii) licence restrictions for the commercial use of some data sets, iii) the reliability and validation of data, and iv) limited number of Open Access publications and expensive fees for publishing Open Access.

Taking these drivers and barriers into account, we postulate the following two policy recommendations to make scientific results (re)usable in SMEs as well as industries.

- Make Open Science, its opportunities and benefits more visible, especially outside the university context.
- Increase the number of Open Access publications available for all interested stakeholders, especially in domains with no strong tradition in Open Access.

Overall, we found that Open Science resources are already used by SMEs and industries. However, there is still a lot of work to do to raise awareness of the Open Science endeavours within the private sector and to increase benefits from Open Science resources to help spur growth and innovation in SMEs and industries.

1. Introduction

Spurring growth and innovation in SMEs and industries is a key goal of policy-makers. A commonly stated advantage of Open Access to publications and data is greater return on investment for funders, as results are made reusable to a range of societal actors including industry (Huber, Wainwright, and Rentocchini 2020; Fell 2019; European Commission 2016). To what extent are Open Science resources such as Open Access publications, Open Data or Open Source code actually being taken up by SMEs or industries, though? In this deliverable, we will present results which address this question using data from an interview study with 11 interviews conducted in Austria and results from a questionnaire study with 108 respondents located across Europe, focussing on attitudes, best practices as well as drivers and barriers for the uptake of Open Science resources in SMEs and industries.

The deliverable report D4.1 - Information Seeking Behaviour and Open Science Uptake in Industry: A Literature Review (Fessl et al. 2020) served as the starting point for this deliverable report D4.2. In T4.1 we conducted an extensive literature review on the current state of the art regarding information-seeking behaviour in SMEs and industries as well as the current status of the uptake of Open Science resources in this regard. A summary of the resulting deliverable is presented in the section Summary of D4.1.

Based on the insights gained from D4.1, we started to investigate the uptake of responsible research and innovation (RRI) and Open Science in industry by conducting an interview study in Austria. The goal of this study was to capture the current state of play in Austrian SMEs and industries in the domains of health, climate and agriculture on the uptake of Open Data/Science and research in industry. To do so, we invited different types of stakeholders (e.g. CEOs, head of departments, etc.) from SMEs and industries situated in the corresponding domains to participate in our interview study. Based on the interviews and the results gained, we set up a questionnaire that we sent out to SMEs and industries situated in our domains of interest across Europe. With the questionnaire we aimed to obtain broader insights across Europe about the uptake of Open Science resources in SMEs and industries so far.

The goal of both studies was to get deeper insights into the uptake of Open Science resources in SMEs and industry and answer the following research question:

> **RQ: To what extent do actors in the private sector (SMEs and industry) make use of (Open) Science resources in different domains (climate, health, agriculture)?**

Additionally, we defined four research sub-questions:

- RQ1: What is the common knowledge, baseline and understanding of Open Science?
- RQ2: How is the search for scientific information currently conducted within SMEs/industry?
- RQ3: Which levels of uptake already exist in the different domains and what differences can be observed across domains?
- RQ4: What are challenges/barriers that hinder the uptake of Open Science resources?

The evidence presented in this deliverable gives some deeper insights about the uptake of Open Science in today's SMEs and industries in Europe. Key findings are as follows:

- (RQ1) In general the knowledge about Open Science is rather low and to a certain degree related to respondents' educational background. We found that people having a university education are more familiar with the concept of Open Science than others. The majority of those who are familiar with the concept already use Open Science resources and practices.
- (RQ2) With regard to the search for information, we found that literature/publications, social contacts, and company internal information are the most relevant sources for domain related knowledge. Further, we found the following barriers and challenges in relation to the search for information such as i) (domain-specific) knowledge that is not searchable, ii) the large amount of data available (although search engines are very powerful today), iii) health related anonymized data is scarce, iv) lack of time to conduct a search, and v) large differences in the quality of information found and vi) difficulties in distinguishing between low- and high-quality information.
- (RQ3) With regard to the uptake of Open Science resources, in the interview study we found that Open Data, Open Access and also Open Source code already play an important role in the respective companies. However, uptake depends on the characteristics of the companies and - as is shown in the survey results- on the educational levels of the employees involved.
- (RQ4) As barriers to the uptake of Open Data, we found that i) health related data is scarce, ii) there are licence restrictions for the commercial use of some data sets and iii) the reliability and validation of data. With regard to Open Access, we found that i) that the number of Open Access publications is limited and ii) the fees for publishing Open Access are rather high, which poses a significant barrier for increasing the number of available Open Access publications.

Based on our findings from the interview and questionnaire studies, we could derive 5 drivers that facilitate the uptake of Open Science resources: i) the employment of people with a university background, ii) offering incentives or support for uptake, iii) offering trainings to increase uptake, iv) learning from trans- and interdisciplinary collaboration with other companies, academia, and other stakeholders and v) exploiting the wisdom of the crowd in relation to Open Source code. Taking these existing drivers into account, we postulate the following two policy recommendations:

- Make the Open Science endeavours, its opportunities and benefits more visible especially outside the university context.
- Increase the number of Open Access publications available for all interested stakeholders, especially in domains with no strong tradition in Open Access.

Summing up our findings, we found that Open Science resources are already used by SMEs and industries, however, there is still some work to do, so that more actors in the private sector become aware of the Open Science endeavours and can benefit from Open Science resources and activities in the future. This can be done by making Open Science more visible outside the university context and raising the public's awareness of the core advantages of Open Science, like transparency, verifiability, replicability and general openness.

This deliverable report is structured as follows. In Section 2, we will present the state of the art including a short summary of D4.1, literature about academic engagement with industry, and stakeholder maps. Then we present the interview study and its results in Section 3, followed by the questionnaire study and its results in Section 4. In Section 5, we will discuss our findings in relation to our research questions and present our stakeholder map. In Section 6, we summarize the drivers and barriers for the uptake of Open Science resources, and suggest policy recommendations to conclude our work.

2. State of the Art

2.1. Information Seeking Behaviour and Open Science Uptake in Industry

Spurring growth and innovation in SMEs is a key goal of policy-makers. A commonly stated advantage of Open Access to publications and data is greater return on investment for funders, as results are made reusable to a range of societal actors including industry. Is open research data actually being taken up by industry, though? D4.1 - Information Seeking Behaviour and Open Science Uptake in Industry: A Literature Review (Fessl et al. 2020) addressed this broad question by semi-systematically summarizing the existing evidence on how scholarly resources are used in industry, with a special focus on Open Science practices.

Crucial for understanding whether industrial actors are able to benefit from Open Science resources, such as research papers or data, is the concept of absorptive capacity, i.e. "…the ability of a firm to recognize the value of new, external information, assimilate it, and apply it to commercial ends" (Cohen and Levinthal 1990). Recent research (e.g. Huber, Wainwright and Rentocchini 2020) has highlighted that particularly SMEs struggle to benefit from Open Data. However, increasing the absorptive capacity would increase the overall uptake of (Open) scientific resources.

Finding relevant resources is a crucial step in recognizing and assimilating new external information. D4.1 therefore semi-systematically reviewed the literature regarding how companies satisfy their information needs. Common barriers in this regard were found to include difficulties in explicating information needs and finding relevant information, as well as lack of time, accessibility and concerns regarding content quality. Several studies have found that the accessibility of information is the most relevant factor in finding information (Guo 2009; Yitzhaki and Hammershlag 2004; Kwasitsu 2003; Su and Contractor 2011). Although recent studies (e.g. Kraaijenbrink and Groen 2006) have highlighted how searching for information on the internet has become ubiquitous, personal and social contacts still play an important role. Considering demographic factors, there exists a clear research gap regarding gender differences in information seeking behaviours, since only 1 out of 30 publications took this into account (Le et al. 2016) .

The literature on information-seeking behaviours among industrial actors indicates that research outputs currently play a somewhat peripheral role in general information seeking behaviour in many industrial sectors. The evidence collected points to a general lack of information-seeking skills amongst employees. Exploiting scientific resources for commercial ends also requires skills specific to the subject area. Companies commonly acquire these skills by either hiring graduates or directly collaborating with academia (Fell 2019; Starasts 2015). Open Access to research findings is found to provide efficiency gains (i.e. time and cost savings associated with accessing research), as well as enabling the development of new products, services, and companies, by lowering the barriers for companies of all sizes (from large firms to start-ups) to accessing basic research.

The evidence assembled in D4.1 lays the groundwork for the activities conducted in T4.2 and guided our interview study as well as questionnaire study reported here.

2.2. Academic Engagement with Industry

While we summarize the current status of the uptake of Open Science resources in deliverable D4.1, the other side of the coin - namely how academics collaborate with non-academic organisations in the industrial domain - is also of interest for us and is presented in two publications by Perkmann et al. (2013) and Perkmann et al. (2021). Perkmann et al. (2013) conducted a literature review of papers from 1989-2011 to investigate "academic engagement" which they define as "knowledge-related collaboration by academic researchers with non-academic organisations" (Perkmann et al. 2013, 424). Respective activities consist of "... collaborative research, contract research, and consulting, as well as informal activities like providing ad hoc advice and networking with practitioners" (Perkmann et al. 2013, 424) and play an important role in bringing academic research into the industrial world (Cohen, Nelson, and Walsh 2002). Perkmann et al. (2013) suggest that academic engagement is a multi-level phenomenon in that it takes characteristics from the individual, institutional and organisational context into account. On the individual level they could show that academic engagement is strongly linked to researchers who are well-established and -connected, who are more senior, who have more social capital and who have a high number of publications and government grants. Thus, academic engagement goes directly hand in hand with academic success. They also state that academic engagement is closely related to the 'Matthew effect' in academia (Merton 1968), "...according to which individual success is reinforced through a virtuous cycle of achievements and returns on those achievements" (Perkmann et al. 2013). This pattern matches also with the findings that male academics are more likely to engage with industry, take up more prominent positions, and consequently are able to mobilise more resources and establish wider networks. On the institutional or organisational level, the results are not that clear. In some studies, academic engagement is negatively associated or not correlated with the research quality of academic institutions, while in other studies there is a positive relationship between engagement and research quality of the organisation. However, what the results have in common is that academic engagement strongly depends on highly motivated and successful individual researchers mostly independent of their affiliated institution or organisation. In 2021, Perkmann et al. (2021) published an updated literature review that confirms the following insights gained from the first literature review (Perkmann et al. 2013): i) academic engagement is complementary and consistent with the promotion of academic research activities, ii) highly engaged researchers are more likely to be committed to academic engagement, iii) academic engagement is positively correlated with acquiring research funding and iv) academic engagement is driven by the characteristics of the individual. Furthermore, additional new factors that influence academic engagement were derived: First, with regard to gender, they found that women engage less in academic engagement than men. Second, academic engagement is also socially conditioned by peer influence and by characteristics of the discipline. And third, there seems to be evidence from activities relating to academic entrepreneurship into subsequent academic engagement with companies.

2.3. Stakeholder Maps

Stakeholder analysis as a tool or method have become increasingly popular in recent years. Managers, policy-makers and researchers recognize the central role of stakeholders (individuals, groups or organisations) and their interest (stake) in influencing activities or goals in an organisation, project or policy direction (Brugha 2000). A stakeholder is a group or an individual who affects or is affected by the achievement or activities of organisations (Freeman 2010; Slabá 2016), including an organization's functioning, goals, development or even survival (Chinyio and Olomolaiye 2009). Stakeholder analysis aims at generating knowledge about the stakeholders to get a deeper understanding about their behaviour, intentions, interrelations and interest in a specific topic, project, organisation or policy direction (Varvasovszky 2000; Freeman 2010; Slabá 2016). In

this regard, stakeholder maps - a way to visualise the results of stakeholder analysis - have been shown to be useful to detect key relations between stakeholders, however, the construction of these maps strongly depends on the purpose, for example detecting stakeholder positions around an organisational objective orto assist in designing educational programs from different perspectives (Brugha 2000; Varvasovszky 2000). For example, Lelea et al. (2014) use a stakeholder map to present the flow of a product in a food supply change from producers to consumers. Another example presented by Axelsson, Melin, and Lindgren (2013) conducted a stakeholder analysis to identify different stakeholder groups and their perceptions relevant for developing an e-service development.

There exists a lot of work regarding stakeholder analysis in general, and stakeholder analysis or stakeholder maps are also used in the context of Open Science. For example, the National Initiatives for Open Science in Europe – NI4OS Europe, aims to be a core contributor to the European Open Science Cloud (EOSC) service portfolio[1]. They released a stakeholder map, showing the existing Open Science stakeholders in 15 partner countries. Gonzalez-Zapata and Heeks (2015) conducted a stakeholder analysis regarding open government data in Chile, uncovering the identity, power, motivations, and worldviews of key actors, including the different meanings of open government data. Another case study in this regard is presented by (Kassen 2018), who studied the multi-institutional and multi-layer nature of open data-driven communication processes. The results show that the Open Data concept could be a promising collaboration platform for different stakeholders to speed up technology-driven public reforms, if a fair and equal contribution from public and private sectors of the economy can be ensured. Stakeholder analysis and maps seems to be a powerful tool in uncovering relations between stakeholders and a specific topic. Therefore, we will apply a stakeholder analysis using a stakeholder map to bring together stakeholders from our target domains and their characteristics relating to the uptake and drivers or and barriers to the use of Open Science resources within industry.

3. Interview Study

In this chapter we present the interview study, including methodology, sample and results. The results are structured along the following topics: i) Information seeking behaviour, ii) the uptake of Open Science resources including barriers and challenges to it, iii) absorptive capacity and business model archetypes, iv) GDPR, ethics and data ownership, and v) knowledge risks. For the first three topics, we additionally analysed the results according to the three domains of health, climate and agriculture (the three domains of special interest to the ON-MERRIT project).

3.1. Methodology

3.1.1. Procedure

First, we derived and developed the questions for our semi-structured interview guide based on the literature review presented in D4.1. After the interview guide was finished, we started the recruiting process of interview participants in Austria. Starting the interviews in Austria has the following reasons: First, the interviews were planned as contextual inquiries, thus, conducting the interviews at the interview partners' company to get insights about the working environment and as the researchers of this work package were located in Austria this would have been very convenient. Second, A common political and societal context

[1] https://ni4os.eu/2020/03/23/ni4os-europe-open-science-stakeholder-map-now-available-online/

the companies are situated in facilitate the synthesis and aggregated analyses of the interviews. And third, the analysis and interpretation of the interview data would have been more difficult with interview partners from different countries unless we manage to recruit a considerable number of interview partners from each country and therefore could do a comparative analysis. Due to the already cumbersome recruitment process in Austria, we decided to take the least risky approach - focus on just one country and try to recruit as many participants from that country as possible."

To do so, we contacted more than 50 companies (in the domains of climate, health, and agriculture) in Austria via email in two invitation rounds - one in summer 2020 and a second one in autumn 2020. Additionally, Know-Center employees contacted people via the Know-Center network on a personal level who met our domain requirements. Although we made a great effort to recruit interview participants, due largely to the effects of the CoVID-19 pandemic, recruitment proved to be difficult and only 13 people agreed to participate in an interview. Originally, the interviews were planned as contextual inquiries, meaning to conduct the interviews in-place, thus, in the corresponding SME to get some insights about the workplace and the company's environment. However, due to the CoVID-19 pandemic all interviews had to be conducted online.

3.1.2. Interview Guide

The interview guide contained questions about the following major topics: data and information usage at work, information seeking behaviour, usage of Open Science/Access/Data, absorptive capacity, business models and knowledge risks. Additionally, we asked about barriers and challenges where feasible. During the interview, the interviewer asked the questions along the developed guidelines, however, depending on the answers given, additional ad-hoc questions were posed, and less-relevant guide questions omitted. The direction of the interview and the questions posed were strongly influenced by the interview partner and the answers given. All interviews were conducted in German. The analysis of the interviews was conducted on the German transcripts; illustrating interview statements cited in this work were translated into English by the researchers. The interview guide can be found below in the Annex: Interview Guide.

3.1.3. Participants and Companies

Altogether, 13 persons (2 female, 11 male) from 11 different companies participated in an interview. In two interviews, two people were interviewed simultaneously. The duration of the interviews was from 30 minutes to 1 hour. From the interviewed participants, 3 are founders and CEOs (Chief Executive Officers) of the company they represent (P2, P8, P9) , 1 is a founder (P4), 3 more are CEOs (P1a, P5, P11), 2 are CTOs (Chief Technology Officers) (P1b, P6) and the remaining are leaders of a department (P3, P7, P10a, P10b). All of them hold a master's degree and 4 of them hold a PhD degree.

Out of the 11 companies, 7 are SMEs (C1, C2, C4, C5, C6, C8, C9, C11) and 3 are large enterprises (C3, C7, C10). Five of these companies are situated in the domain of health (C1, C5, C6, C7, C11), 4 in the domain of climate (C2, C8, C9 and C10) and 2 in the domain of agriculture (C3 and C4).

The companies' products and services in the domain of health encompass software applications related to dizziness (C1), diabetes (C6), and the support of elderly people (C11); a food supplement product (C5), and a product related to hearing implants (C7). The companies' products and services situated in the domain of agriculture deal with the maintenance, administration and organisation of a big database for livestock

farming (C3) and a sensor for animal health (C4). The companies' products and services in the climate area are related to the development of sensors for measuring PH, oxygen and temperature under different conditions (C2), data analysis of wind turbines (C8), and earth observation (C10).

C9's work was related to climate with the goal to develop small hydro-power plants, however, this company was closed in 2014 which was not known by the researchers prior to starting the interview. Although the interview was nevertheless completed, the results of this interview will not be reported. Table 1 gives an overview of the companies of the interview partners.

Company/ Participant	Domain	Founded	Number of Employees	Product
C1 / P1a, P1b	Health	End of 2019	4	Software applications related to dizziness
C2 / P2	Climate	2017 founded in Austria (German company founded in 2021)	Not available	Sensor for oxygen, PH and temperature measurements
C3 / P3	Agriculture	1960	> 500	Database for livestock farming
C4 / P4	Agriculture	2009	60	Sensor for animal health
C5 / P5	Health	Founded in 2016, since 2019 at market	40	Food supplement product
C6 / P6	Health	2016	4	Software applications related to diabetes
C7 / P7	Heath	1970s	1500	Product related to hearing implants
C8 / P8	Climate	2019	12	Data analysis of wind turbines
C9 / P9	Climate	closed 2014	-	Development of small hydro-power plants
C10 / P10a, P10b	Climate	Unknown	> 500	Earth observation
C11 / P11	Health	2019	1	Software applications related to the support of elderly people

Table 1: Company overview of the interview partners

3.1.4. Data collection and analysis

One interview was conducted via MS Teams[2] and nine interviews were conducted via GoTo Meeting[3]. All Interviews were audio-recorded by using the recording functionality of the corresponding tool. The interview conducted via phone was not recorded due to the wish of the interview partner, thus the researcher conducting this interview took notes. All interviews were conducted in German. The audio recorded interviews were then transcribed by a professional transcription service. Afterwards, the interviews were anonymized, meaning that all person-related and company-related information were replaced with neutral

[2] https://www.microsoft.com/de-at/microsoft-teams/group-chat-software

[3] https://www.gotomeeting.com/

values, for example, participant's name was replaced with P1 (=participant 1) and the corresponding company's name was replaced with for example C1 (=company 1).

For the qualitative content analysis of the interviews, we used MAXQDA[4], a software tool for analysing data from qualitative and mixed method research. We developed a coding schema in a two-step approach: First, one of the researchers developed a coding schema (deductive/top-down) along the interview guide used. Then, two researchers independently applied this coding schema to two interviews with the result that the coding schema was not comprehensive enough to capture all the information from the data that was relevant to our research questions. Therefore, in a second step, we started our coding process again and enhanced our existing coding schema with codes that could be derived inductively during the analysis of the interviews (inductive/bottom-up). These additional codes were then compared and synthesized in a reflective research meeting. The resulting final coding schema then was used by one of the researchers to code all interviews. When summing up our findings from the interviews, relevant paragraphs, sentences, or sentence snippets were translated on the fly. The English translations were inserted into the text, while the German original statement was added as footnote.

All interview documents - the recordings, the transcripts as well as the analysis files created during the coding with MAXQDA - are secured with a password and are securely stored on an internal database at the Know-Center.

3.2. Interview Results

3.2.1. Information Seeking behaviour

Information Seeking Behaviour

In the interview, we asked, "How do you keep yourself and your company up-to-date with research?" that is related to their products and services. With this question we wanted to find out which information channels they use to get up-to-date information. Overall, 12 different information channels were mentioned by our participants. The channels mentioned were sorted according to the frequency of mentions (meaning how often those channels were explicitly named by the participants) in the interviews starting with the most often mentioned channel.

Literature/Publications: The most relevant information channel to stay up to date in their specific domain, and which was mentioned by all interview participants, is literature and publications (27 times mentioned). For all partners, it is of crucial relevance to read or at least skim literature that is related to their research topics or domain knowledge. As mentioned by P1a (developing an application for dizziness), "*[...] if it's about a certain scientific question, then of course you have to read the papers on it [...]*".[5] P2 (developing a sensor for measuring PH, oxygen and temperature) also mentioned that they conduct very specific literature research while P3 (working on a big database for livestock farming) stated that they have subscribed to certain journals and that they are happy if some papers are available via Open Access. Otherwise, they try to get publications from their academic partners, which is also in line with P4 (working on a sensor for animal

[4] https://www.maxqda.de/

[5] P1a: "[...] wenn es um eine gewisse wissenschaftliche Fragestellung geht, dann muss man natürlich die Papers dazu lesen [...] "

health). P5 (working on a food supplement product) highlights that each month, two to three new research studies appear in relation to their specific research topic and which they need for their product. This means "*[...] that more or less every day we look at what other products are coming onto the market that also deal with cell research*".[6] While P5 exactly tailored their search to their product development, P6 (developing an app on diabetes) conducts their literature research on a more general level: "*Literature research simply in general on the topic, clinical guidelines and simply studies that take place in the area, there are very few comparable systems*".[7] P6 highlighted that they have instantiated "*[...] a quasi process within the framework of our medical device development that the literature searches are also regularly updated*".[8] P7 (working on hearing implants), P8 (analysis of wind turbine data), P10b (working on earth observation) and P11 (working on an app to support elderly people) also confirm that they use publications on a regular basis.

Social Contacts: Social contacts were mentioned (18 times mentioned) as the second most important source of information by 8 interviewees (P1a, P3, P4, P5, P6, P7, P8, P10a). Here we can differentiate social contacts with business partners versus social contacts with scientific partners. P1a, P4, P7 and P8 especially mentioned social contacts with business partners such as customers or company partners. P1a stated that not all information "can be googled" and that some information is knowledge of the people and business relationships. For estimations about the market or information in relation to technology one needs to talk to people and therefore strongly depends on good networking. For example, about their last meeting, P1a remarked "*[...] where we got to know a new person, hey, he's dealing with a similar topic, we could make a call and that's how you get into conversation with people*".[9] P4 comments on the value of social contacts from two sides "*[...] in the meantime, we are well known and, I would say, established in our industry, that we are very often contacted when it comes to new projects, that is very much word of mouth*".[10] On the other hand, they actively contact their end customers and farmers so that they " *… find out what they are doing with our sensors and offer them support there as well*".[11] P7 stated that "*A great deal is actually achieved through direct personal communication with the companies we work with, i.e. on the one hand there are suppliers who provide us with services and components, assemblies and the engineers there on site, but also through communication with companies who supply us with production facilities or measuring and testing equipment, which is how a great deal of information comes to us in the company*".[12]. P8 also mentioned the discussions with their company partners and word of mouth as well as their regular contacts with customers.

P3, P5, P6 and P10a strongly emphasized their connections to researchers. P3 highlighted that "*Yes, we work very intensively together with various scientific partners, nationally and internationally, and yes, we have*

[6] P5: "dass wir mehr oder minder jeden Tag schauen, welche anderen Produkte auf den Markt kommen, die sich auch mit Zellforschung beschäftigen"

[7] P6: "Literaturrecherche einfach allgemein zum Thema, klinische Leitlinien und einfach Studien, die sich in dem Bereich abspielen, es gibt ja ganz wenig vergleichbare Systeme"

[8] P6: "[...] einen quasi Prozess im Rahmen von unserer Medizinproduktentwicklung, dass man das auch regelmäßig, die Literatursuchen aktualisiert."

[9] P1: "[...] wo wir eine neue Person kennengelernt haben, hey der beschäftigt sich mit einer ähnlichen Thematik, da könnten wir mal einen Call aufsetzen und so kommt man auch mit den Leuten ins Gespräch"

[10] P4: "mittlerweile sind wir durchaus bekannt und sage ich mal, etabliert in unserer Branche, dass wir sehr oft auch kontaktiert werden, wenn es um neue Projekte geht, das ist sehr viel ist Mundpropaganda"

[11] P4: " dass wir herauskriegen, was die mit unseren Sensoren machen und denen auch da Unterstützung anbieten".

[12] P11: "sehr viel über eigentlich direkte persönliche Kommunikation mit Firmen, mit denen wir zusammenarbeiten, das heißt, einerseits sind es Zulieferfirmen, die für uns eben Dienstleistungen und also Komponenten, Baugruppen liefern und den Ingenieuren dort vor Ort, dann aber auch aus der Kommunikation mit Firmen, die uns Fertigungsanlagen oder Mess- und Prüfanlagen liefern, auf dem Weg kommt sehr viel Information zu uns in die Firma herein."

various joint projects"[13] and "*[...] we are also represented in international boards*"[14] where international knowledge exchange takes place and where you get corresponding information. P5 stressed especially the collaboration with scientific partners as extremely important, as they are distributed all over the world and they really know where which kind of research takes place. In this regard they can serve as a kind of filter "*[...] because then I have a scientist who deals with cell death, a scientist who deals with cell renewal, an external scientist who deals with cell protection and so it is easier for us to filter the information*".[15] P6 also sees the value of researchers in their relevant domain as important and here especially physicians from medical universities since "*[...] these are actually the people who go to conferences, bring back all the news and from whom we get the information, where we mutually simply pass on interesting things to each other*"[16]. P10a also stated that he is closely connected with scientific partners all over the world "*[...] where you exchange ideas again and again and read the work of your colleagues, in order to not only continue to educate yourself, but also to question your own work*".[17]

Conferences: Conferences are the third most important information sources, which were mentioned by 8 participants (13 times mentioned). P1a, P3, P4, P7, P10, P11 mentioned that they regularly attend expert conferences. P6 highlighted that their partners, and here especially physicians, attend conferences and convey the ongoing research to them. P8 also pointed out that they not only attend expert conferences but also serves sometimes as a chairman or speaker.

Webtools: Overall, most of the participants mentioned Google (P1, P5, P6, P7, P11), Google Scholar (P2), Google Alerts (P4, P5) as the most relevant webtools (11 times mentioned) used for their search activities. Additionally, Pubmed (P1a, P5) and SciFinder (P2) were named, and also the company internal library (P7) was mentioned.

Patent research: P4, P5, P7, P8 indicated (6 times mentioned) that they check patent databases regularly. P5 summarized their purposes in one sentence: "*On the one hand, to see if we have possibilities to patent something somewhere, but on the other hand, to see if we would now infringe any patent with a new product or a new idea*".[18] The latter reason for looking at patents was also mentioned by P6, who furthermore stated that he looks especially at patents that were submitted by competitors in order to see what is going on in their sector.

Other: There is also a list of other information sources that were mentioned by one interview partner each: customers (2 times mentioned), library (1), competitors (1), newsletter (1), social media (1), statistical information (1) and webinars (1). P1a mentioned newsletters as an important source that is often

[13] P3: "Ja wir arbeiten sehr intensiv mit den verschiedenen Wissenschafts-Partnern zusammen, national und international und ja, haben da eben verschiedene gemeinsame Projekte"

[14] P3: "wir sind auch in internationalen Gremien vertreten"

[15] P5: "Weil dann habe ich einen Wissenschaftler, der sich mit Zelltod beschäftigt, einen Wissenschaftler, der sich mit Zellerneuerung beschäftigt, einen externen Wissenschaftler, der sich mit Zellschutz beschäftigt und so ist es dann für uns leichter, die Informationen zu filtern."

[16] P6: ".. das sind eigentlich die, die auf Konferenzen fahren, jede Neuigkeit mitbringen und von denen wir die Infos bekommen, wo man sich einfach gegenseitig dann interessante Dinge weiterleitet und berichtet"

[17] P10a "[...] wo man sich immer wieder austauscht und die Arbeiten der Kollegen liest, um auch und vor allem was ganz wichtig ist, nicht nur weiter sich weiterbildet, sondern auch die eigene Arbeit hinterfragt."

[18] P5: " Einerseits um zu sehen, haben wir Möglichkeiten, dass wir irgendwo etwas patentieren, aber andererseits um zu sehen, ob wir jetzt gegen irgendein Patent verstoßen würden mit einem neuen Produkt oder einer neuen Idee."

underestimated: "*You think to yourself, as a student I always thought, leave me alone with the newsletters, today I'm actually happy about this time saving, if you have good newsletters that point you in the right direction a bit*".[19] P4 also mentioned social media as an important information resource. P8 also looked at their competitors, especially "*[...] what kind of information is published by these companies and what do they do, what is their strategic orientation*".[20] He is in close contact with his customers to discuss "*[...] how they can get the most value out of our technology, through this exchange we are always made aware of what are the current priority issues*".[21] Additionally, P10b mentioned webinars and online courses and P11 named statistical information platforms as valuable information sources for his business.

Barriers and Challenges

Although information seeking is a common task and is more or less part of the daily business of all interview participants, finding relevant information is still challenging from time to time. While five of our interviewees stated that they always find what they are looking for, others hesitated and stated that it also depends on what they are looking for.

Overall, the interviewee responses can be categorized into 5 specific types of barriers or challenges related to their information seeking process that were mentioned by more than one interviewee: domain specific knowledge (4 interviewees), amount of information (2 interviewees), patient data (2 interviewees), paywalls (2 interviewees), and keywords (2 interviewees).

Domain specific knowledge: There is a type of knowledge that cannot be found through online search because it consists of compound work of individual people or experts, or domain-specific knowledge that is not published yet or is not intended for publication at all. For example, P1a stated that there is knowledge which cannot be directly found "in Google" like, for example, to prepare an estimation about their health application for the market, where you need input from different people, e.g. technicians, economists, etc., so "*[...] you have to put things together a bit, you have to talk to people, make phone calls, which means that you are very dependent on good networking [...]*".[22] P11, dealing with the support of elderly people, stated that supporting elderly people with an application leads to the saving of time and costs by relatives and also by the affected persons themselves. However, she stated that "*[...] everyone is always talking about it, but it has never been proven. So there's still no study [...] that really confirms it, that's where I think to myself, well, AAL research has been around for 10 years now [...] I don't think there's enough for that either [...]*".[23] Then there is knowledge that is currently under research, as stated by P3, wherein such topics will only be presented at conferences or not at all, if *"[...] of course internal information, where someone then wants to protect something or so, is just not presented [...]*".[24] This is also in line with the statement of P6, who refers

[19] P1a: "Man denkt sich, als Student habe ich mir immer gedacht, lasst mich in Ruhe mit den Newslettern, heute freue ich mich eigentlich über diese Zeitersparnis, wenn man da gute Newsletter hat, die einen da ein bisschen in die richtige Richtung weisen. "

[20] P8: "[...] was für Information wird veröffentlicht von diesen Firmen und was tun sie, wie ist ihre strategische Ausrichtung."

[21] P8: "[...] können am meisten Mehrwert aus unserer Technologie holen, durch diesen Austausch werden wir immer darauf aufmerksam gemacht, was sind die derzeitigen Prioritätsthemen."

[22] P1a: "[...] da muss man sich die Sachen ein bisschen zusammendichten, da muss man mit Leuten sprechen, Telefonate führen, das heißt, hier ist man sehr abhängig von einem guten Networking.."

[23] P11: "[...] alle reden immer davon, aber es ist noch nie bewiesen worden. Also da gibt's noch immer keine Studie, [...] dass man das wirklich bestätigt, das ist das, wo ich mir denke, naja, AAL Forschung gibt's jetzt seit 10 Jahren [...] dazu finde ich auch zu wenig.[...]"

[24] P3: "[...] natürlich interne Informationen, wo dann jemand was schützen will oder so, wird halt nicht so präsentiert [...] "

to patents search stating that "*[...] competing products, for example, we have of course looked at the patents, but what is in the patent does not describe their algorithm or anything like that [...]*".[25]

Amount of information: Although search engines are very powerful today, the large amount of data available is still seen as a challenge or barrier to finding the needed information as mentioned by three of our interviewees. P5, working on a food supplement product, highlights that "*[...] in the US alone, there are over 600 patent applications with and on [our topic]*".[26] Additionally, he mentioned that "*[...] on a scientific basis and on a preclinical and clinical basis, two to three new studies on [our topic] appear per month, that is actually only our core*".[27] As a drawback due to the large amount of data he stated, "*[...] that we overlook things, that we don't see new competitors, that we are late to new scientific discoveries [...]*".[28] For him the major barrier is to be able to filter all relevant data to find the essential information needed. This is also in line with P7, developing a product related to hearing implants, who said that "*The hardest part is really targeting, when you have very specific topics, to get to the titles that are relevant to my or my working group's work*".[29] As a solution to his problem he wishes, "*[...] a dream now would be if you could go in somewhere with the very specific search terms that I have [...] and then simply get a broad overview of what is there and then at the same time get [...] a rating [...] from other users of these sources of information[...]*".[30]

Patient data: Especially with the health domain, another challenge mentioned is to find anonymized statistical information and patient data relating to different diseases. For example, P1a mentioned that "*[...] it is impossible to get data from patients with dizziness, because dizziness is always, always associated with other diseases and not only per se, and so you simply don't have much data to draw from*".[31] He also mentioned that "*[...] there are already publications, but concrete data, statistics, they vary and also differ*".[32]Thus, to get data relevant and useful for their work in order to be able to, for example, calculate some market forecasts, is really difficult and only possible if one knows the correct persons or institutions. P6 also stated that, for example, with (anonymized statistical) data about the starting dose for insulin "*[...] there you would probably get further with a huge amount of data [...]*".[33] However, such information is not available to them.

Paywalls: Two of our interview partners mentioned paywalls, specifically the access to journal or conference papers, as a barrier. P1a stated that paywalls to access journals are a barrier for SMEs. He stated, "*[...] if you*

[25] P6: "[...] Konkurrenzprodukte jetzt zum Beispiel, da haben wir uns die Patente natürlich angeschaut, aber also das was im Patent steht, beschreibt jetzt deren Algorithmus oder so nicht [...] ".

[26] P5: "[...] allein in den USA gibt es über 600 Patentanmeldungen mit und zu [unserem Thema] [...] ".

[27] P5: "[...] auf wissenschaftlicher Basis und auf präklinischer und klinischer Basis erscheint pro Monat ja zwei bis drei neue Studien zu Spermidin, das ist eigentlich nur unser Kern [...]"

[28] P5: "[...] dass wir Sachen übersehen, dass wir neue Konkurrenten nicht sehen, dass wir neue wissenschaftliche Entdeckungen zu spät mitbekommen[...]"

[29] P7: "[...] Das Schwierigste ist wirklich zielgerichtet, wenn man sehr spezielle Themen hat, auf die Titel aufmerksam zu werden, die für meine oder die Arbeit meiner Arbeitsgruppe relevant sind [...] "

[30] P7: "[...] ein Traum wäre jetzt, wenn man da sozusagen mit den ganz spezifischen Suchbegriffen, die ich habe [...], irgendwo hineingeht und dann einfach einen breiten Überblick kriegt über das was da ist und dann gleichzeitig eine [...] eine Bewertung [...] von anderen Nutzern dieser Informationsquellen bekommen würde [...]"

[31] P1: "..und da ist es unmöglich, dass man zu Daten kommt, die jetzt von Schwindelpatienten zu tun haben, weil Schwindel ist immer, kommt immer mit anderen Krankheiten in Verbindung und nicht nur per se und so hat man da einfach keine großen Daten aus denen man abschöpfen kann."

[32] P1: "[...] es gibt schon Veröffentlichungen, aber konkrete Daten, Statistiken, die variieren und unterscheiden sich auch [...]"

[33] P6: "[...] da würde man wahrscheinlich mit einer großen Masse an Daten weiterkommen[...]"

talk to medical doctors, they don't see the challenges because they usually have access through their hospital but just because you have access there, you're not allowed to use it for a company and that kind of thing, so that's information where technically you would have access but legally you wouldn't, that's why you don't do it and you're a bit frustrated".[34] This is also the opinion of P2, who stated that as an employee of a company they cannot access all the journals they would like to access, but as an employee of a university it is easier to access relevant journals and publications.

Query formulation (keywords): For example, P7 stated,"*[...] I try to get publications, then I definitely come up against limits, that when I try to describe these special fields with keywords, that no relevant information comes back, because they are simply niche fields.*"[35] And another interview partner (P11) mentioned a challenge with fitting keywords, such as "*I think there must be something ... that's a point that I'm actually looking for and not finding with the resources that I'm using*".[36]

Additionally, the following two topics, one related to interoperability and one to credibility, were only mentioned by one interviewee each. P4 raised as a barrier to finding relevant information is the problem of different databases existing as stand-alone solutions that are not connected to each other. He stated "*[...] that there are simply different databases and sources, and that some of them are all cooking their own soup and don't want to network, and others are simply not technically designed for this*".[37]
P11 mentioned a barrier related to trust - trustworthiness of resources - which especially in the health domain is of crucial relevance. After finding some relevant information she stated that sometimes she is not sure "*[...] how correct is the information now, who really wrote or authored it?*".[38] From her perspective there is some uncertainty regarding the reliability of documents found: "*[...] is this really a scientific paper or is it now a bit of a half-hearted paper supported by someone [...]*".[39]

Analysis per Domain

Figure 1 provides an overview about the information channels/sources used for staying up to date per domain, based on how often these channels/sources were mentioned in the interviews. Information channels/sources were mentioned by our interviewees from the health domain (*n*=6) 39 times in total, by the interview partners from the agricultural domain (*n*=2) 19 times and 25 times by the interviewees from the climate domain (*n*=4). While for the health domain (38.5% of the 39 mentions) and the climate domain (32% of the 25 mentions) literature/publications are the most used information sources, in agriculture literature/publications (21.1% of the 19 mentions) seem to be a little bit less relevant. For the agricultural domain, social contacts (36.8%) are mentioned as the most relevant information channel followed by

[34] P1: "[...] wenn man mit Medizinern redet, sehen Sie die Herausforderung nicht, weil die haben meistens Zugang über ihr Spital aber nur weil Sie dort Zugang haben, dürfen Sie es nicht für eine Firma verwenden und so was, das heißt, das sind Informationen, wo man technisch Zugang hätte, aber legal nicht, deswegen macht man es nicht und man ist ein bisschen frustriert[...]".

[35] P7: "[...] versuchen an Publikationen zu kommen, dann stoße ich da durchaus auch an Grenzen, dass wenn ich diese Spezialgebiete mit Schlagwörtern versuche zu umreißen, dass dann keine relevanten Informationen mehr zurückkommen, weil es einfach zum Teil Nischengebiete sind"

[36] P11: "Ich denke, irgendwas muss es schon geben [...], das ist jetzt gerade so ein Punkt, den ich eigentlich – dass ich suche und nicht fündig werde mit den Ressourcen, die ich verwende"

[37] P4: "dass es einfach verschiedene Datenbanken und Quellen gibt und die halt zum Teil alle ihr eigenes Süppchen kochen und nicht, zum Teil sich nicht vernetzen wollen, zum Teil einfach technisch nicht dafür ausgelegt sind"

[38] P11: "[...] wie richtig ist die Information jetzt, wer hat das jetzt wirklich verfasst oder geschrieben [...]"

[39] P11: "[...] ist das wirklich eine wissenschaftliche Arbeit oder ist das jetzt ein bisschen so Halbherzigeres, von irgendwem unterstütztes Papier [...]"

conference attendances (26.3%). Webtools (5.3%) and Patent Research (5.3%) play only a minor role. For the health domain especially social contacts (20.5%) are highly relevant whereas conferences (12.8%), various Webtools (10.3%) or Patent Research (10.3%) play only a tangential role. In contrast, for the climate domain Webtools (24%) are highly relevant, Social Contacts (12%) and Conferences (12%) play a tangential role, while Patent Research (4%) is hardly relevant. These results give some first insights about the different search behaviours to keep oneself up to date across the three domains. However, due to the different group sizes of participants of the three domains, these results need to be interpreted with caution.

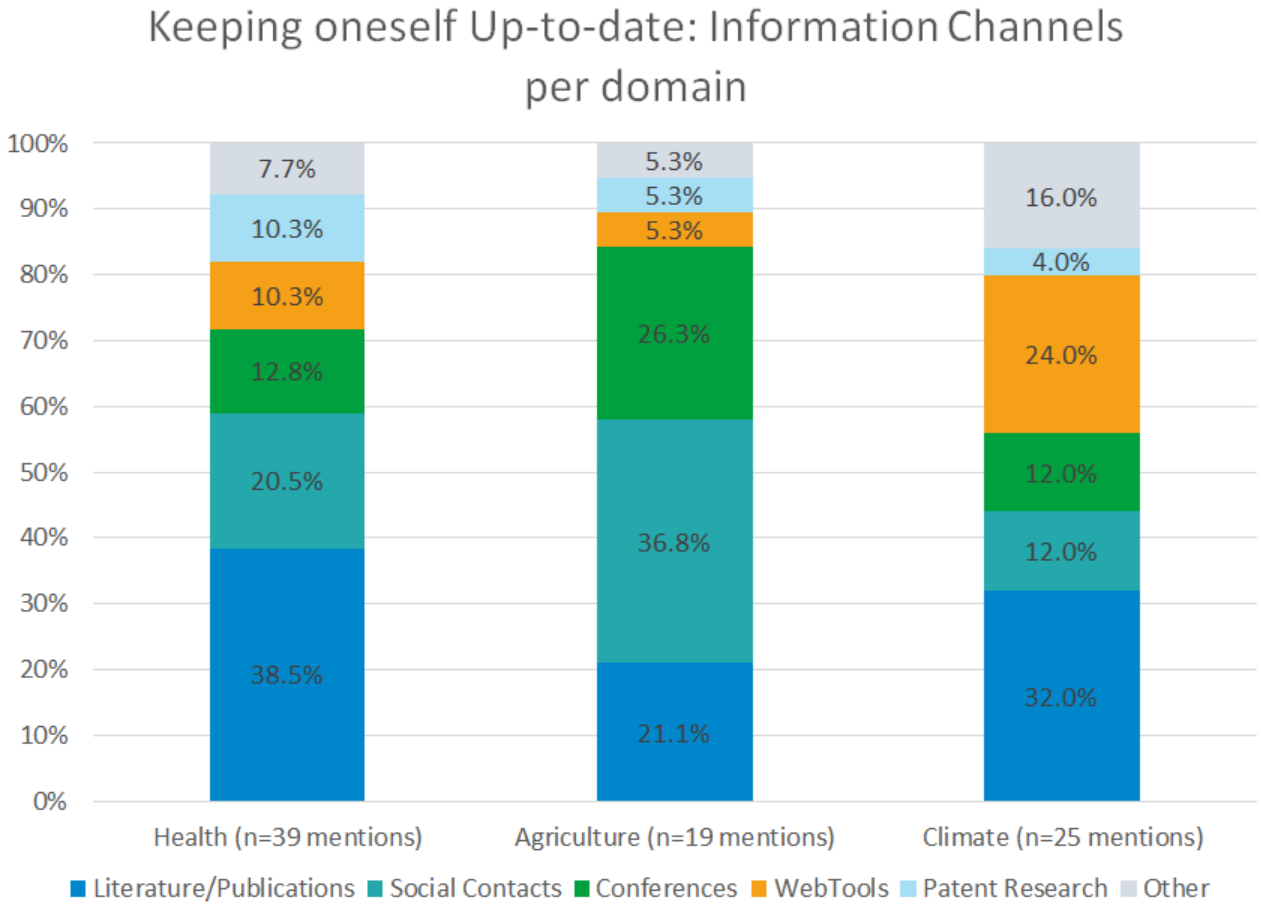


Figure 1: Importance of different information channels used per domain

3.2.2. The Uptake of Open Science Resources

We asked our participants what they know about the concept of Open Science and related terms Open Data, Open Access and Open Source code (altogether referred to as Open Science resources). Additionally, we asked them if they use Open Science resources and if they do, which types they use and if not, why not. Furthermore, we asked them about challenges and barriers which they faced with the topic and the usage of freely available resources. We found that all of our interview partners were already aware of Open Science and most of them have already used some OS resources.

Open Science: Some of our interview partners (P1b, P3, P6, P7 and P11) really highlighted the value of the Open Science endeavours. P1b, developing a software application related to dizziness, values the openness of the machine learning community in general and that everything is more or less accessible: "*[...] the*

machine learning community, one of the nicest things in artificial intelligence is that the whole community is embraced to really share that, a lot of it goes through arXiv [...] almost everything is accessible there and of course you're very grateful for that because it's also one of the most complicated fields".[40] P3, who is working on a big database for livestock farming, highlights that the animal breeding community has been very open and that people want to learn from each other, but it has changed over time, especially because of the investment of money for developing new products or algorithms. P6, who is developing an application related to diabetes, positively highlighted Open Science endeavours as a whole. For software development, he especially highlights the value of the crowd, because many people are using the same libraries and therefore, they are also looking for errors and support the further development of the code. P7, who is working on hearing implants and P11, who is working on supporting elderly people, do not use Open Science for their work but see the value of this endeavours for their companies, as stated by P7: "*[...] I am actually almost certain that this is used in the areas in the company for which it is relevant*".[41]

Open Source code: Five of our interview partners (P1b, P4, P6, P8, P10a, P11) mentioned using Open Source code. P1b highlighted the value of Open Source code, and emphasised here the role of licencing, especially the licenses that allow them to use Open Source code for commercial use: "*[...] we work with Open Source code, licensing is very important here, because Open Source is not just Open Source, of course we only work with things where the licence is also suitable for commercial use*".[42] Additionally he stated, "*Yes, we rely on Open Source and try to return improvements that do not directly interfere with the company's USP*".[43] P4, who is working on a sensor for animal health, P6, who is working on an application dealing with diabetes, and P10a, who is working on earth observation data, use Open Source code for their algorithms and analysis, thus, for their software development. P6 sees Open Source as a "*model of how to develop software, because either you make it proprietary, in which case you have to do all the maintenance and troubleshooting and everything yourself, or you make it more open and do it with libraries that have a general benefit, publish them and then you have the advantage that you have a large mass of users who will also look for errors and perhaps participate in the further development*".[44]

Open Data: Open Data is already used by four of our interviewees (P3, P5, P10a, P10b). P3, who is working on a big database for livestock farming, mentioned the use of climate data from the ZAMG (*Zentralanstalt für Meteorologie und Geodynamik* - the Austrian meteorological research institution), which is not completely Open Data because one has to ask for permission in order to use it and location code data (data about where municipalities are). In both cases they not only use the data for their work but also investigate how they benefit from the usage of the data. P5, who is working on a food supplement product, uses some (closed) data from pharmacies about sales numbers for improving their sales and marketing strategy. They

[40] P1b: "[...] die Machine Learning Community eine der schönsten Eigenschaften in künstlicher Intelligenz ist, dass die ganze Community das sehr embraced ist wirklich zu teilen, viel geht da über arXiv [...], da ist fast alles zugänglich und dafür ist man natürlich sehr dankbar, weil es ist auch eines der kompliziertesten Gebiete."

[41] P7: "[...] aber ich bin mir eigentlich fast sicher, dass das in den Bereichen in der Firma, für die das relevant ist, durchaus auch zum Einsatz kommt."

[42] P1b: "wir arbeiten einerseits mit Open Source Code, hier ist die Lizenzierung sehr wichtig, weil Open Source ist nicht gleich Open Source, wir arbeiten natürlich nur mit Dingen, wo die Lizenz auch passt für Commercial Use"

[43] P1b: "[...] ja man stützt sich auf Open Source und man versucht Verbesserungen zurückzugeben, die jetzt nicht direkt irgendwie in den USP der Firma eingreifen"

[44] P6: "ein Modell, wie man Software entwickeln kann, weil man macht es eben entweder proprietär, dann muss man auch die ganze Wartung und die ganze Fehlersuche und alles selber machen oder man macht es eben offener und macht das jetzt vor allem halt mit Bibliotheken, die allgemein einen Nutzen haben, veröffentlicht sie und hat halt dann die Vorteile, dass man eine große Masse an Usern hat, die dann auch Fehler suchen und sich vielleicht doch an der Weiterentwicklung beteiligen."

do not use Open Data for their product development, marketing or sales, but for internal business organisations. He stated, "*With this public data, when it comes to the topic of COVID, just to take the example, we use it purely to introduce internal organisational control mechanisms. That means that when we see [...] that the number of cases is increasing in Styria, then we know as a company that we have to act in a way that we divide our production shifts differently, so that people don't see each other and there are no problems*".[45] Also P10a and P10b confirmed that they not only use Open Data but also provide Open Data through their organisation. On the one hand, they offer meteorological data from 70 world meteorological organisation stations from Austria's Alpine region. On the other hand, P10b confirmed to use available Open Data for national as well as international projects, like for example in Asia. In this project they do analysis and assessments of "*[...] landslides, of course it is difficult to find information in place, and I go to the area of OpenStreetMap, [...] but it doesn't matter whether it's in Austria or wherever, but these data are now highly frequent and also very dense. So if I only look at this subject area, the Open Street Layers [...], which are spatial data that are available in vector form, can be included in analyses, just as one aspect*".[46]

Open Access: Since literature research is a generic part of all interviewees' everyday work and publications are the most important source of information for most of the interviewees, thus, they are all strongly interested in Open Access publications. However, the interviewees referred to Open Access mainly in relation to barriers, which are described in the following section.

Barriers and Challenges

In the following, we will present barriers and challenges that are referred to the uptake of Open Source code and Open Data.

Open Source Code Barriers/Challenges: Using Open Source code not only has advantages, but also raises some challenges or barriers according to P1b, P4, P8 and P11. P1b mentioned that using Open Source software is a kind of a double-edged sword. He highlighted that, for example, Google or Facebook share their code, however, this code is licenced with non-commercial licences, meaning that he does not even dare to have a look at the code to learn from it, which he perceived as a huge barrier. "*On the one hand, you are very grateful that these big companies like Nvidia and Google, Facebook just publish their results and also their code and stuff like that. On the other hand, it's also frustrating because it's like a pudding that's in front of you but you're not allowed to eat it because it's a code, it's always non-commercially licensed*".[47] P4 also confirmed the use of Open Source, however, he stated that they have to invest some effort to decide what to share with others and what stays with them. P8 emphasized that when he founded his company, he decided explicitly not to use Open Source data to ensure and guarantee the reliability of his software because

[45] P5: "Bei diesen öffentlichen Daten, wenn es jetzt um das Thema Covid nur mal aufzugreifen, um das Beispiel zu bringen, da ist es so, das verwenden wir rein, um interne Organisation Steuerungsmechanismen einzuführen. Das heißt, wenn wir sehen, [...], in der Steiermark steigen die Fallzahlen, dann wissen wir als Unternehmen, okay wir müssen jetzt irgendwie handeln, dass wir unsere Produktionsschichten anders aufteilen, dass sich die Leute nicht sehen und es da zu keinen Problemen kommt."

[46] P10b: "Hangrutschungen, da tun wir uns natürlich schwer vor Ort Informationen zu haben und ich gehe da auf den Bereich OpenstreetMap, [..] ob das aber in Österreich ist oder wo auch immer, ist egal, aber diese Daten sind inzwischen hochfrequent und auch sehr sehr dicht. Also wenn ich nur diesen Themenbereich anschaue, sind ja die Open Street Layer [...], das sind raumbezogene Daten, die in Vektorform vorhanden sind, in Analysen einfließen zu lassen, nur als ein Aspekt."

[47] P1b: "Im Sinne von Open Source Software sind die Hindernisse, ist das ein zweischneidiges Schwert. Einerseits ist man sehr dankbar, dass diese großen Firmen wie Nvidia und Google, Facebook ihre Ergebnisse und auch ihren Code und so was einfach publizieren, andererseits ist das auch frustrierend, weil das ist so wie ein Pudding, der vor einem steht, aber man darf ihn nicht essen, weil es ist ein Code, der ist immer non-commercial lizenziert"

"*[...] we have a certain obligation, sometimes very expensive things are done based on our results, [...] and therefore we have to provide a certain quality guarantee. So that means we can't have code that is changed or adapted by someone else, if then something went wrong, who is responsible for it? That is why I decided against an Open Source model and therefore we charge money for the use of our software*".[48] P11 mentioned that "*[...] if developers don't care about it anymore and don't want to use it, then it falls to the ground, that's the problem [...]*".[49]

Open Data Barriers/Challenges: Several barriers/challenges were mentioned by five of our interviewees (P1b, P6, P7, P8, P10b) in relation to Open Data. P1b mentioned a similar barrier with regard to the licenses of Open Source code protecting data for commercial use. Also, P6 expressed several reservations relating to the use of Open Data. He mentioned that for his health-related application, "*[…] I don't know of any data source that could be useful for us now*"[50], and if there were data from diabetes therapy, he is very sceptical that he could use it, and if then maybe only for very specific questions. He also clearly stated that "*[...] I don't know of any source where you can get free data from the health sector anywhere. If there were, I would take a look at it, but as I said, I don't know anything about* it".[51] This is also in line with P7 stating that it is not usual to use Open Data for the development of hearing implants. P8 also does not use Open Data, because they get data from their customers and "*[…] this is very, very, very sensitive information, actually, which means, precisely, that it remains bilateral between us and our customers*"[52], while P10b stated that they use, for example, OpenStreetMap data and "*[...] of course the question of validation is very big [...]*".[53]

Open Access Barriers/Challenges: Besides the already described barriers, namely the limited number of Open Access publications available and high publication fees for publishing Open Access, no additional barriers and challenges in relation to Open Access were mentioned by our interviewees. P5 and P6 both explicitly refer to the usage of Open Access publications, but also report challenges related to the limited number of Open Access journals in their domain and expensive fees for publishing in such journals. P5 stated that, for example in "Pubmed", "*[...] the abstract or summary is accessible [...]*"[54] but to get the whole paper, he needs to get in touch with colleagues working at the medical university. P6 also highlighted the value of Open Access in general but criticized the high fees to publish their own work Open Access: "*[...] we haven't always been able to afford to openly publish all the latest publications ourselves, but now we have invested in the more*

[48] P8: "[...] wir haben eine gewisse Pflicht, anhand der Ergebnisse werden zum Teil sehr teure Dinge gemacht, [...] und daher müssen wir eine gewisse Qualitätsgarantie liefern. Also das heißt, wir können nicht Code der von jemand anderem geändert oder angepasst wird, wenn dann irgendwas ist, wer ist dann dafür verantwortlich? Das heißt, ich habe mich gegen ein Open Source-Modell entschieden und daher verlangen wir Geld für die Nutzung unserer Software."

[49] P11: "[...] wenn sich keine Entwickler mehr darum kümmern und es nicht mehr verwenden wollen, dann fällt es auf den Boden, das ist ja das Problem[...]"

[50] P6: "[...] bei Daten kenne ich keine Datenquelle, die jetzt für uns nutzbar sein könnte.."

[51] P6: "[...] also mir wäre keine Quelle bekannt, wo man jetzt einen Daten aus dem Gesundheitsbereich irgendwo frei beziehen kann. Wenn es das gäbe, würde ich es mir schon anschauen, aber wie gesagt, ist mir nix bekannt."

[52] P8: "[...] das ist ganz, ganz, sehr sensible Informationen eigentlich, das heißt, genau, es bleibt dann bilateral zwischen uns und unseren Kunden"

[53] P10b: [...] .natürlich die Frage der Validierung steht ja ganz groß im Raum, werden aber definitiv verwendet, um Einschätzungen einmal abgeben zu können."

[54] P5: "[...] .dass das Abstrakt oder ein Summary zugänglich is [...] "

important ones simply because it is public relations work if this publication is simply freely available from the publisher".[55]

Analysis per Domain

In the following, we present an overview of how often different Open Science resources were explicitly mentioned and discussed per domain. In the health domain, Open Source code, Open Data, and Open Access were used and discussed, and different related barriers were mentioned as well. In the health domain, two interviewees (P1b and P6) strongly highlight the value of Open Science endeavours in general, while two other participants (P7 and P11) did not use Open Science so far. Open Source code is used and valued for software development by four interviewees (P1b, P6, P7 and P11). Two of them (P1b and P11) also mention some barriers in this regard (e.g. non-commercial licenses). Two other interviewees (P5 and P6) from the health domain mentioned that they would like to engage in Open Access more but report barriers with regard to lack of availability of relevant publications Open Access and high costs for publishing Open Access themselves.

In agriculture, one of the two interviewees (P4) uses Open Source code for the development of algorithms and analysis, but consciously takes care of what to share. With regard to Open Data, the other interviewee (P3) uses - to a certain degree open - weather data from ZAMG.

In the climate domain, the interview partners were aware of Open Science resources. One of the four interviewees from this domain (P8) explicitly decided against using Open Source code because of compliance reasons. Two interviewees from the same company (P10a and P10b) confirmed to not only offer Open Data but also to use Open Data such as OpenStreetMap for some of their projects, however, they also raise challenges relating to the validation of the data used.

Again, due to the small group sizes, meaningful comparisons between the three domains regarding the domain-specific uptake of Open Science resources can only be done for the sample in this interview study, thus, the generalisability of the results is very limited.

3.2.3. Absorptive Capacity & Business Model Archetypes

Absorptive capacity has been defined as "... a firm's ability to recognize the value of new information, assimilate it, and apply it to commercial ends" by ((Cohen and Levinthal 1990), 128). In the same direction goes an analysis by Deloitte, who call data "the new raw material of the twenty-first century" (Deloitte, LLP. 2012, 1). (Open) Data can be seen as a raw material that allows companies to create or develop new (business) opportunities based upon their business models, such as make more robust business decisions, get deeper insights about their customers or improve the products, services or applications. Therefore, we not only looked at which Open Science resources are already used in SMEs and industry, we also took a close look at the underlying business models and their relation to the usage of (open) data of the interviewees' companies. To do so, we shortly introduce five different business model archetypes that we use in our analysis, identified by Deloitte (Deloitte, LLP. 2012):

[55] P6: "[...] wir haben es uns jetzt nicht immer leisten können quasi bei allen letzten Publikationen, dass wir sie selber offen hergeben, aber jetzt gerade bei den wichtigeren haben wir das investiert einfach, weil es Öffentlichkeitsarbeit ist, wenn diese Publikation einfach frei verfügbar ist vom Verlag"

- Suppliers: organisations that publish their data via an open interface to allow others to use and reuse it.
- Aggregators: organisations that collect and aggregate Open Data and, sometimes, other proprietary data, typically on a particular sectoral theme, find correlations, identify efficiencies or visualise complex relationships.
- Developers: organisations and software entrepreneurs that design, build and sell web-based, tablet or smartphone applications for individual consumption.
- Enrichers: organisations (typically larger, established businesses) that use Open Data to enhance their existing products and services through better insight.
- Enablers: organisations that facilitate the supply or use of Open Data, but are not themselves users or re-users of Open Data.

Then, we asked the interviewees about the current business model of their companies and the corresponding archetypes, and which of them could be an option for them in the future.

Business Models and Archetypes

Company C1 is developing an application related to dizziness. Their goal is to develop this specific application and related services, "*[...] that tries to fulfil a medical purpose without in any way handing over your data to anyone*"[56], as stated by P1a. He also mentioned that they want to generate a certain amount of trust regarding their customers, and that their customers can be sure that data inserted in the application will not be used for other purposes. P1a stated that *"[...] the money doesn't come from the data, the money comes from this product*".[57] Regarding the archetypes, none of the five archetypes fits their current business model. In the future, P1b mentioned to also maybe create data sets related to dizziness, like for example "*[...] geographical data, they are quite interesting, because in our case a client could be a health insurance company and if the health insurance company comes from France, then they would perhaps like to know regionally where increased dizziness occurs*"[58] and such information could become a unique selling point (USP) in the future and could maybe refer to Aggregators.

Company C2 is developing and selling hardware in the form of a sensor for measuring PH, oxygen and temperature under different conditions. Their unique selling point is how the sensor is constructed and their goal is to remain the key technology provider for this special sensor. With regard to the business model archetypes, P2 stated that for his company "*[...] it would be maybe an Enabler, that's the closest fit, I think [...]*".[59] With regard to the future, P2 stated that they have already internally discussed providing measurement services, where sensor data could be stored on an internal cloud server and offer aggregated sensor data in a meaningful way to customers, maybe with monthly or annual subscriptions, which would go

[56] P1a: "das versucht einen medizinischen Zweck zu erfüllen ohne in irgendeiner Form Ihre Daten irgendwem auszuliefern sozusagen"

[57] P1a: "[...] das Geld kommt nicht durch die Daten, das Geld kommt eben durch dieses Produkt "

[58] P1b: "[...] geographische Erkenntnisse zum Beispiel, die sind durchaus interessant, weil bei uns könnte ein Kunde eine Krankenkasse sein und wenn die Krankenkasse aus Frankreich kommt, dann würden die vielleicht gerne regional wisse, wo tritt vermehrt Schwindel auf".

[59] P2: "[...] es wäre vielleicht ein Enabler, das passt, glaube ich, am ehesten [...]"

in the direction of being Aggregators. But at the moment, such a business model would go beyond their internal structure.

Company C3 is dealing with the maintenance, administration and organisation of a big database for livestock farming. Their business model is based on data. P3 highlights here especially the collaboration with partners in order to create an added value. With regard to the business model archetypes, P3 sees her company between Aggregators and Developers. On the one hand, they collect data about livestock farming and on the other hand, they offer this data for free to their member organisations and for research with exactly defined purposes. For the generation of this data, there is a lot of labour and money involved, thus, making it Open Data "*[...] would undermine the trust in our member organisations and customers*".[60]

Company C4 is developing a sensor for animal health. Their unique selling point is not the development of the sensor (as this is an easy thing to do as stated by P4), it is the analysis and interpretation of data collected with the sensor and to inform the farmers about relevant health issues related to their animals or, for instance, an imminent birth. With regard to the business model archetypes, P4 sees the company as an Aggregator with focus on proprietary data and for the future they will stay there. From time to time, they also function as an Enabler, where they provide data sets for scientific projects, but not in the form of Open Data.

Company C5 is developing a food supplement product. Their unique selling point is the extraction of a polyamine from wheat and selling it by offering different food supplement products. With regard to the business model archetypes, they see themselves as a mixture of Aggregators and Enrichers: Aggregators in that they collect and use Open Data for their product and to find out how their active substance works with other active substances; Enrichers in that they use open accessible data such as information from the *Deutsche Apothekerzeitung* (German pharmacists' magazine) in order to pursue the development of the market. Subsequently, they can decide how to react to the development by expanding their products and services.

Company C6 is developing a medical device, thus a software application, that supports the therapy of diabetes. Their unique selling point is selling this application to hospitals as well as to care homes. With regard to the business model archetypes, they see themselves as Developers as they are developing a web-based application and maybe as Enrichers that could maybe make use of Open Data in the future.

C7 is a company that is developing a product related to hearing implants which makes hearing possible again for people with hearing impairments. With regard to the business model archetypes, they see themselves as Developers and Enrichers by not only developing the hearing implants but also by making the data they receive available to generate benefits for their patients.

Company C8 is developing a software application for analysing data from wind turbines. For these analyses, they receive datasets from their customers, analyse the data and deliver the results back to their customers. The data they received will be deleted after the analysis is finished. Only lessons learned or insights gained from the data in relation to improving their algorithms for the data analysis stays in C8. Overall, the presented business model archetypes do not fit for C8, because their customers are very sensitive about their data.

[60] P3: "[...] damit würden wir das Vertrauen in unsere Mitgliedsorganisationen und Kunden untergraben [...]"

However, if he had to decide for one of them, P8 sees his company most likely as an Aggregator for customer data and he would not change this archetype in the future.

Company C10 is dealing with earth observation data in general. As they are a big research center, they are dealing with a lot of different types of data and have different types of business models, depending on the department and w.r.t their product catalogues and operational services. From the interviewees point of view, they see their department as a kind of Supplier, Aggregator, Enricher and Enabler. With regard to Developers, P10a stated that "*[...] at the moment we rely more on third-party providers, so it is our noble goal and we would like to do that, but we do not have the capacity*".[61]

Company C11 is developing an application to support elderly people at home, thus, the application is connected to an alarm receiving center as well as to the Red Cross. It offers reminders for medication reminders and drinking reminders. P11 sees herself as a Developer only.

Table 2 summarizes our findings with regard to the current business model types and possible future business model types of our interview partners. Summing up, our investigation shows that only one of them (C10) offers Open Data. Half of them (C3, C4, C5, C8, C10) collect and use proprietaire data, thus serving as Aggregator, while only three of them (C5, C7, C10) would use Open Data to enrich their services and products, and three (C2, C4, C10) facilitate the supply or use of Open Data, but do not use Open Data themselves. C10 is an exception here, in the sense that due to their size different business model archetypes could be referred to different departments. Two of the companies in our interview study (C6, C11) are developers only and do not use Open Data at all. With regard to possible future business model archetypes, only 3 of the companies (C1, C2, and C8) could think of collecting and aggregating Open Data, but with a big "Maybe". Summing up, although all interviewees are aware of Open Data and related opportunities, their current business models only rarely integrate Open Data in their daily business.

Company	Current Business Model Archetype	Future Business Model archetype
C1	No archetype fits	(Maybe) Aggregator: create data sets about the relation of geographical data and the occurence of dizziness
C2	Enabler	(Maybe) Aggregator: offer aggregated sensor data, but with monthly or annual descriptions
C3	Aggregator and Developer	No changes planned
C4	Aggregator, sometimes Enabler	No changes planned
C5	Aggregator, Enricher	No changes planned
C6	Developer	(Maybe) Enricher
C7	Developer, Enricher	No changes planned
C8	(Maybe) Aggregator	(Maybe) Aggregator - no changes planned
C10	Supplier, Aggregator, Enricher and Enabler	No changes planned
C11	Developer	No changes planned

Table 2: Current and future business model archetypes of the interview partners

[61] P10a: "Bei Developer da momentan bauen wir da eher auf Drittanbieter, also es ist so unser hehres Ziel und wir möchten das auch, aber die Kapazität haben wir nicht."

Analysis per Domain

Within the health domain, our interviewees stated to be Aggregator (1), Developer (3) or Enricher (1) and maybe to become an Enricher (2) in the future (multiple responses were possible). Being a Supplier or Enabler does not fit at all in the domain of health, as they are dealing with sensitive health-related data in different ways meaning that sharing health related data is not possible and also using Open Data for the development of health-related applications does not work especially with regards to compliance and liability. With regard to the domain of agriculture, both of them see themselves as Aggregators (2) that collect and aggregate open or proprietary data. One of them sees themselves as a Developer and the other as an Enabler. In the climate domain, one sees himself as Enabler (1) and the other as Developer (1). Interestingly C10 sees themselves as Supplier, Aggregator, Enricher and Enabler. This could be explained that C10 is a big company with a set of different departments covering one or the other business model archetype each.

Overall, with regard to the health domain, using and providing (anonymized) health-related (open) data raises several challenges and implications, thus the concept of Open Data seems not to fit due to compliance, liability, trust, or ethical reasons - just to mention only a few. In the domains of agriculture and climate it seems to individually depend on the business model as well as the products and services of the companies, however.

3.2.4. GDPR, Ethics and Data Ownership

During the interviews, topics including the role of GDPR, Ethics and data ownership were discussed. Especially in relation to the health domain, collected patient-related data and information are strictly protected by the rules of the GDPR. In this regard, especially P1b stated that one of their "*[...] most convincing point[s] is that we are writing medical software that takes data protection very seriously, in the sense that it is designed in such a way that we try to process everything more or less on premises in order to not have to collect anything*".[62] He is also aware that there are specific regulations in relation to statistical data in the GDPR, however, he states that "*[...] for me, as a non-expert on the GDPR or a non-lawyer, it is of course difficult to find the boundary that one dares to cross*".[63] He also emphasized that "*[...] the basic principle of purpose is a very strong basic principle [...] and that's why we don't collect or we haven't planned to collect any data that doesn't have exactly the purpose for the interpretation of dizziness [...]*".[64] Also P7, who is working on hearing implants, stated that they have a lot of very sensitive data about their customers and their implants, including the individual settings. In their case, this data "*[...] is managed meticulously and in a highly secure manner in accordance with the rules of data protection at our company, of course*".[65] But also in the domains of agriculture and climate, data protection is of crucial relevance. For example, P4, who is working on a sensor for animal health, stated that the customers - in his case farmers - are "*[...] very precise about what happens*

[62] P10b: "[...] überzeugendsten Points ist, dass wir eine medizinische Software schreiben, die Datenschutz sehr hoch schreibt, in dem Sinne, dass es vom Design her so ist, dass wir schon versuchen, möglichst alles 'on premise' mehr oder weniger zu verarbeiten und dann nix sammeln müssen."

[63] P1b: "[...] für mich als nicht DSGVO-Experte oder nicht Anwalt ist es natürlich schwierig dann die Abgrenzung zu finden, dass man sich da darüber traut [...]"

[64] P1b: "[...] das Grundprinzip der Zweckgebundenheit ist ein sehr starkes Grundprinzip [...] und deswegen sammeln wir oder haben wir auch nicht geplant irgendwelche Daten zu sammeln, die nicht genau diesen Zweck dieser Schwindelinterpretation verfolgen[..] "

[65] P7: "[...] werden ganz akribisch und hochsicher nach den Regeln des Datenschutzes natürlich bei uns in der Firma verwaltet."

to their data and that no one does anything wrong with it [...]".[66] And also for P8, who is working on data analysis of wind turbines, data ownership in his domain is a very sensitive topic and of crucial relevance as discussions about data ownership have been ongoing for years. On the one hand, the manufacturer of the wind turbines deliver the systems, including the data collection software and hardware. In their opinion they need to have access to the data in the first years, within the scope of the warranty requirements and maintenance. On the other hand, the operators of the wind farms state that they have bought the wind turbines so also the data belongs to them and they are allowed to handover this data to third-party providers like C8. Therefore, P8 stated "*We have to be very careful with this issue, [...] we have to have big terms and conditions documents signed so that all the legal things [are clarified], because there is extreme nervousness in this area*".[67]

3.2.5. Knowledge Risks

Another topic we addressed in the interviews was about knowledge risks. We asked the interviewees to what extent they are concerned that critical knowledge could flow out of their company through Open Science endeavours. In addition, we asked them how they are currently managing this risk in their respective companies.

As P1b stated, for their application related to dizziness, sharing and protecting data is a double-edged sword. On the one hand, he gets to know so many people through Open Source development. And he stated, "*[...] I also find science very fascinating and I like the values and the basic principles behind it [...] it's a lot of fun to share such things, because you somehow have a feeling of success, I've written this library and when you see how many people use your library, you're happy*".[68] On the other hand, he mentioned that he has to carefully consider what to share, "*[...] yes, I mean, knowledge gives you a head start and you would actually like to use this advantage and not necessarily send it to the potential competition by email. That is, here you somehow have interests that work against each other*".[69] In the end they try to find a good balance in sharing and protection of data, stating that "*[...] yes, one relies on Open Source and tries to give back improvements that do not directly interfere with the USP of the company in any way*".[70] This is also in line with P1a, who stated that "*[...] a characteristic that I've taken up as a data scientist is that I say I don't talk about projects on principle*".[71] Additionally he said that "*[...] when you do customer projects, where you do something for a customer, that is, you simply don't talk about projects and yes, and when we have appointments with potential partners or something, [...], you try to solve it via NDAs, but even so NDAs [...] are not the panacea*".[72]

[66] P4: "[...] Landwirte beispielsweise sind da sehr genau, was mit ihren Daten passiert und dass da eh nicht jemand Falscher was damit anfängt [...] "

[67] P8: "Wir müssen sehr vorsichtig sein mit diesem Thema, [...] wir müssen dann in großen Terms und Conditions-Dokumente unterschreiben lassen, damit die ganzen rechtlichen Dinge [geklärt sind], weil extreme Nervosität herrscht in diesem Bereich."

[68] P1a: "[...] ich finde auch die Wissenschaft sehr faszinierend und mir gefallen die Werte und die Grundprinzipien dahinter [...] es macht irrsinnig Spaß, solche Sachen zu teilen, weil man hat ja irgendwie so ein Erfolgsgefühl, ich habe diese Library jetzt wieder geschrieben und wenn man dann sieht, wie viele Leute die Library von einen verwenden, das freut einen."

[69] P1a: "[...] , ja ich meine, Erkenntnisse geben einem ja einen Vorsprung und diesen Vorsprung möchte man ja eigentlich gerne nutzen und nicht unbedingt der potenziellen Konkurrenz per Email schicken. Das heißt, hier hat man irgendwie Interessen, die gegeneinander wirken. "

[70] P1a: "[...] ja man stützt sich auf Open Source und man versucht Verbesserungen zurückzugeben, die jetzt nicht direkt irgendwie in den USP der Firma eingreifen."

[71] P1b: "[...] .eine Eigenschaft, die ich mir als Data Scientist angewöhnt habe, dass ich sage, ich rede prinzipiell nicht über Projekte

[72] P1b: "[...] wenn man Kundenprojekte macht, wo man für einen Kunden irgendwas macht, das heißt, man redet einfach nicht über Projekte und ja und wenn wir halt Termine haben mit potenziellen Partnern oder so was, [...], einerseits versucht man es über NDAs zu lösen, aber selbst, also NDAs ist [...] ist jetzt auch nicht das Allheilwunder."

Furthermore, he explained that they have to represent interests of two types that work against each other similar, to yin and yang: "*[...] one is that you don't really tell anyone about [our product] until it's finished, because then you don't make any enemies, and the other is that you have to talk to people, of course, so that you can get somewhere [...]*".[73] P2, who is developing sensors for measuring PH, oxygen and temperature, stated that the most important secret is the recipe of how their sensors are built: "*[...] these are recipes, I would say, which are then kept under lock and key. And that is really what is important. So it's also the components that are now in the sensors that are changed, that are different from what is published. But apart from three people, nobody knows that*".[74] P3, who is working on a big database for livestock farming, went in a similar direction, stating that "*[...] if someone invests money, it is clear that they also want to earn money with it, and then of course they look more at what is shared where, and then of course one is not so open, some algorithm, how one perhaps somehow derives a characteristic for the prediction of a disease, that of course one does not share that*".[75] Also P5, who is dealing with the development of a food supplement product, also has concerns that critical information could be lost. He stated:"*So, if I just think about the extraction, if one is able to find out which natural acids and bases we use [to get our food supplement], then that would of course be something that every competitor could take up and make it easier for them to bring their own product to the market*".[76] He also stated that they have very good employment contracts with their employees and that they take relevant precautionary measures, for example: "*So we try to keep our cloud system [...] secure and try to have antivirus programs up so no one can hack us*".[77] This is in line with P7, who is developing a product related to hearing implants, stating that their data about customers "*[...] are managed meticulously and in a highly secure manner in accordance with the rules of data protection at our company*".[78] He also stated that this applies also within the company: "*Only a very limited group of people has access to this very, very sensitive information. So access is very limited, but the data is of course used for common purposes such as risk management, i.e. the risk management process in our company, but the actual evaluation and the use of this data takes place in a very small circle*".[79] P8, who is working on the data analysis of wind turbines, is also very concerned about their data. Everything they do in their company is kept secure and confidential.

In contrast, some of our interview partners were not that concerned about the loss of data. They believe that data alone is not enough; there is a need for competence and experience in the corresponding domain before

[73] P1b: "[...] das eine ist, dass man eigentlich möglichst niemandem von einem erzählt, solange bis er nicht fertig ist, weil dann macht man sich auch keine Feinde und das andere ist, dass man natürlich mit Leuten reden muss, damit man weiterkommt[...]"

[74] P2: "[...] das sind Kochrezepte, sage ich mal so, die dann halt unter Verschluss bleiben. Und das ist wirklich das, was wichtig ist. Also es sind auch die Komponenten, die jetzt in den Sensoren sind, drin sind, die sind verändert, die sind anders als das was veröffentlicht ist. Aber das wissen halt außer drei Leute niemand."

[75] P3: "Wenn jemand Geld investiert, ist es klar, dass er damit natürlich auch Geld verdienen möchte und da wird natürlich dann immer mehr geschaut, was dann wo geteilt wird und man natürlich dann nicht so offen ist, irgendeinen Algorithmus, wie man vielleicht irgendwie ein Merkmal für die Vorhersage einer Erkrankung ableitet, dass man das natürlich nicht teilt."

[76] P5: "Also wenn ich jetzt nur an die Extraktion denke, wenn man da herausfinden würde, welche natürlichen Säuren und Basen wir einsetzen, [um unser Nahrungsergänzungsmittel zu bekommen], dann wäre das natürlich etwas, was jeder Konkurrent aufgreifen könnte und es ihm erleichtern würde ein eigenes Produkt auf den Markt zu bringen."

[77] P5: "Also wir versuchen unser Cloudsystem [...] sicher zu halten und versuchen Antivirus-Programme oben zu haben, damit niemand uns hacken kann."

[78] P7: "[...] werden ganz akribisch und hochsicher nach den Regeln des Datenschutzes natürlich bei uns in der Firma verwaltet"

[79] P7: "Also auf diese sehr, sehr sensiblen Informationen, da kommt nur ein sehr eingeschränkter Personenkreis zu. Also da ist der Zugriff sehr limitiert, aber die Daten werden durchaus natürlich für also gängige Zwecke wie Risikomanagement, also den Risikomanagementprozess bei uns in der Firma wird natürlich genützt, aber die tatsächliche Auswertung dieser Daten und Verwertung dieser Daten, die erfolgt in einem ganz kleinen Kreis."

someone can really benefit from the data. P4, who is developing a sensor for animal health, does not fear that critical knowledge could flow out of the company. He stated that they know so much about their data that "*[...] we know how difficult it is that we also believe it is not so quickly replicated*".[80] In addition, he stated: "*I rather see it as an opportunity, because even if someone, based on what we have, develops, for example, a new evaluation for these data series, this still has no added value as a standalone finding. So what will happen? We will integrate it into our app and if someone else has created value, then we will somehow find an agreement so that s/he can also benefit from it. So from that point of view, I can imagine that this is more of an opportunity*".[81] P6, who is working on an application for diabetes, is of a similar opinion like P5, in that sharing and protection is a kind of balancing act. Nevertheless, their asset is their overall system and integrating their algorithm somewhere does not create a medical product. Additionally, he stated that "*[...] we show screenshots and yes, because it's really the overall system that makes the difference and even if a large corporation decides that this would be a topic that they should also address, then they would have to invest considerable resources in order to imitate us and probably so many resources in the meantime that there is no decision-maker who can quickly decide that [...]*".[82] Also P10a and P10b are not really concerned about the loss of data. In their opinion data is important "*[...] but you need the competence to make a product out of it, I think this is the way to go [...]*"[83] as P10b stated.

Summing up our findings in this regard is that one (P1a, P1b, P5 and P6) has to carefully consider what to share and with whom to share, as they see it as a balancing act. For example (P1a, P1b), on the one hand you need to give back Open Source code to the corresponding community, but on the other hand you should not give away your knowledge that serves as an advantage in relation to your customers. P2, P3, P5, P7 and P8 strongly keep their unique selling point including their developed algorithms, recipes for building their sensors, or person-related information under lock and key. In contrast, P4, P6, P10a and P10b were not that concerned about the loss of data or to give away critical knowledge. They believe that data alone is not enough and that there is a need for competence and experience in the corresponding domain before someone can really benefit from the data they have.

[80] P4: "wir wissen schon so viel über die Daten, wir wissen, wie schwierig es ist, dass wir auch glauben, dass es nicht so geschwind nachgebaut ist"

[81] P4: "Ich sehe eher ein bisschen die Möglichkeit, weil selbst wenn jemand, basierend auf dem, was wir haben, beispielsweise eine neue Auswertung für diese Datenreihen entwickelt, hat das als Standalone-Erkenntnis noch keinen Mehrwert. Das heißt, was wird passieren? Wir werden das in unsere App integrieren und wenn da jemand anderer einen Wert geschaffen hat, dann werden wir das irgendwie, unsere Vereinbarung finden, dass der auch was davon hat oder die. Also so gesehen kann ich mir durchaus vorstellen, dass das eher eine Möglichkeit ist."

[82] P6 "[...] wir zeigen schon auch immer wieder Screenshots her und ja, weil einfach wirklich das Gesamtsystem das dann ausmacht und auch wenn jetzt ein Großkonzern quasi entscheidet, das wäre ein Thema, auf das wir auch zugehen sollten, dann muss er schon beträchtliche Mittel in die Hand nehmen, um uns das nachzumachen und wahrscheinlich inzwischen so viele Mittel, dass es keinen Entscheider gibt, der das jetzt schnell mal entscheidet[...]"

[83] P10b: "[...] sondern du brauchst schon die Kompetenz um ein Produkt daraus zu machen, ich glaube, dort geht auch der Weg h"

4. Questionnaire Study

In this chapter, we present the questionnaire study including its methodology, sample and results, organized as i) information seeking behaviour, ii) the uptake of Open Science Resources iii) absorptive capacity and business model archetypes, and iv) barriers and opportunities for the uptake. Where applicable, we analysed the results according to the three domains of health, climate and agriculture.

4.1. Methods

In order to create a meaningful, relevant and high-quality questionnaire, we developed the questions based on existing research (e.g. Flatten et al. 2011; Fawad Sharif et al. 2020) and the results of the interview study. First, we analysed the interviews according to the relevant topics (e.g. Information seeking behaviour, uptake of Open Science resources) and extracted topics that were of particular relevance for the interviewees and that we therefore wanted to address in the questionnaire as well. Second, we conducted a search for already existing questionnaires in this regard that could be included in our questionnaire. Finally, considering both types of information, we created our questionnaire taylored to answer our overall research question including all related topics.

4.1.1. Instrument

The questionnaire was developed by the project team and consists of the following sections:

1) information about the company where the respondent works (domain, location, etc.)
2) demographic information about the respondent (gender, age, highest education level, etc.)
3) information search behavior and related experiences of the respondent (information sources used, type of information needed, barriers relating to information search processes, etc.)
4) uptake of Open Science in the company (Open Science activities already performed by the company, Open Science activities interesting for the company, barriers to the uptake of Open Science, etc.)
5) absorptive capacity of the company (scale based on (Flatten et al. 2011)
6) business model archetypes (Deloitte, L. L. P. 2012)
7) knowledge risks related to Open Science activities (Fawad Sharif et al. 2020)
8) an open question to share additional thoughts on Open Science

The whole questionnaire can be found in the Annex: Questionnaire.

4.1.2. Data Collection

The questionnaire was created in LimeSurvey, one version in German and one in English. For recruiting participants from SMEs and industries across Europe, we used 2 different approaches: snowball sampling and Prolific. In both approaches, the invited participants were asked to fill in the questionnaire anonymously. No personal data was collected; additionally participants actively consented to participating in the survey study..

Approach 1: Snowball Sampling:

In order to invite participants to fill in our questionnaire, we prepared flyers in English and in German (see Annex: Flyers) that explain the purpose of our research. We disseminated these flyers through the ON-MERRIT social media channels as well as the project's website. Additionally, we distributed them proactively via email through the ON-MERRIT Consortium and the consortium's professional and private networks to reach SMEs and industries, and cluster organisations in the domains of health, climate and agriculture. Additionally, we invited the interview partners to participate and asked them to distribute the flyers in their network to cause a snowball sampling effect. The survey was open from mid January 2021 to the end of March 2021. Although we put a lot of effort into motivating people to complete the questionnaire, in the end only 10 fully completed and 2 partly completed questionnaires could be included in the analysis.

Approach 2: Prolific (research-related crowd-working platform based in the UK):

Due to the low number of completed questionnaires we received with our first approach, we decided to use Prolific - a research-related crowd-working platform - to recruit more participants. The Prolific-platform provides the opportunity to select participants based on various demographic information, including working domain and location. Regarding location, we focussed on participants who live and work in Europe. Regarding working domains, the domains health, climate and agriculture/forestry/fishery are selectable on Prolific. However, we had to widen the scope because of the extremely low number of potential respondents from the climate and the agriculture/forestry/fishery domain, respectively. We selected additionally the IT-domain (based on the insights from the interview study, we assumed that some people working in this domain deal with health, climate and/or agricultural data) and the option "other".

We started the survey on Prolific on April 15th 2021 at 11:30. Within less than two hours and after continuously reviewing the incoming completed questionnaires for plausibility (we had to exclude three questionnaires due to implausible answer patterns) we reached our target sample size of 100 fully completed, plausible questionnaires. In addition, we received 3 partly completed questionnaires. However, after a closer review of the answers, we had to exclude 7 of the fully completed questionnaires because the respondents did not work in SME/industry, but in the public domain (which we accidently included through allowing participants from the domain "other" to take part in the survey). Thus, from the Prolific study, 93 fully completed and 3 partly completed questionnaires could be included in the analyses (96 questionnaires in total).

4.1.3. Data Preparation

Prior to the analyses, some preparatory steps were undertaken. First, answers to open questions were clustered or subsumed in categories, where possible (e.g., recurring identical or very similar answers to questions of type "if other, please specify"). Second, internal consistency of the scales included in the questionnaire (absorptive capacity, knowledge risks) was checked and mean scores of the scales were computed. In our sample, the internal consistency (Cronbach's alpha) of the absorptive capacity subscales was alpha=.674 for the acquisition-subscale, alpha=.821 for the subscale "assimilation", alpha=.830 for the transformation-subscale, and alpha=.773 for the exploitation-subscale. Thus, internal consistency of the subscales "assimilation", "transformation", and "exploitation" was good/sufficient. The internal consistency of the subscale "acquisition" was slightly too low, however, due to the theoretical fit of the items we still computed a mean-score of the subscale and included it into further analysis. Although also for the knowledge risks scale and subscales a reliability analysis yielded sufficient internal consistency of alpha > .7, a meaningful analysis of this scale was not possible because many respondents had difficulty answering the questions, as illustrated by the feedback to the questionnaire, such as "*I had to click on 'neutral' for the last set of questions*

as we do not use Open Science and I have not heard of it prior to this questionnaire." Therefore, we excluded this scale from the analysis.

4.1.4. Participants

Background information - individual level

The sample consists of 108 respondents in total. In addition to the 103 fully completed questionnaires (93 from Prolific, 10 from snowball sampling), we included five partly completed questionnaires (3 from Prolific, 2 from snowball sampling; at least the section "information search behavior" was completed) in the analysis. The respondents were between 18 and 65 years old (*M*=29.52, *SD*=9.534) and almost two thirds (65.7%, *n*=71) were male (see Fig. 2).

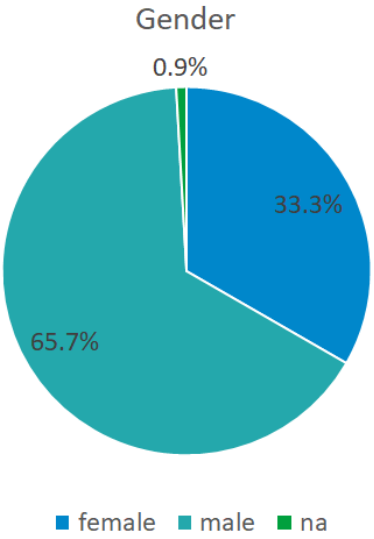


Figure 2: Gender distribution

The educational level of the respondents was rather high (see Fig. 3). Around two third (65.7%, *n*=71) had at least a bachelor's degree, including 8.3% (*n*=9) with a doctorate.

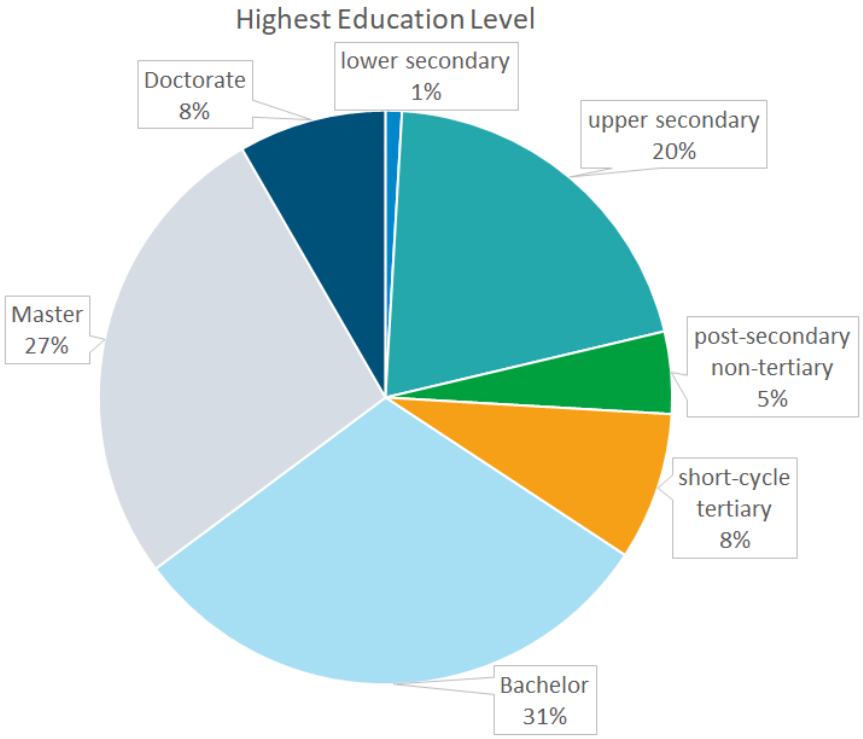


Figure 3: Distribution of the highest education level

On average, the respondents have been working in the current company for 3.79 years (*SD*=4.488, *Min*=0, *Max*=21) and have 5.6 years (*SD*=6.229, *Min*=0, *Max*=30) of working experience in the domain. The majority of the respondents (73.8%) do not have a managerial/senior position in the company.

Background information - company level

To gather information regarding the domain in which the respondents work, we included a multiple response question which included an "other" option to specify domains apart from health, climate and agriculture/forestry/fishery. Fig. 4 shows the results for this question, including categorized results for the "other" option. The IT and software domain was specified in this "other" field most often and therefore, a respective category was created. Other relatively frequently named domains were technology (without further specification) and manufacturing/industry. The new category "Domain: other" subsumes domains that were named less than four times, for example, finance, games or retail.

From our domains of interest (health, agriculture/forestry/fishery and climate) only the health domain was represented in larger numbers: Thirty-five of the respondents indicated that their company was situated in the health domain and only eight respondents work in the climate and agriculture/forestry/fishery domains, respectively.

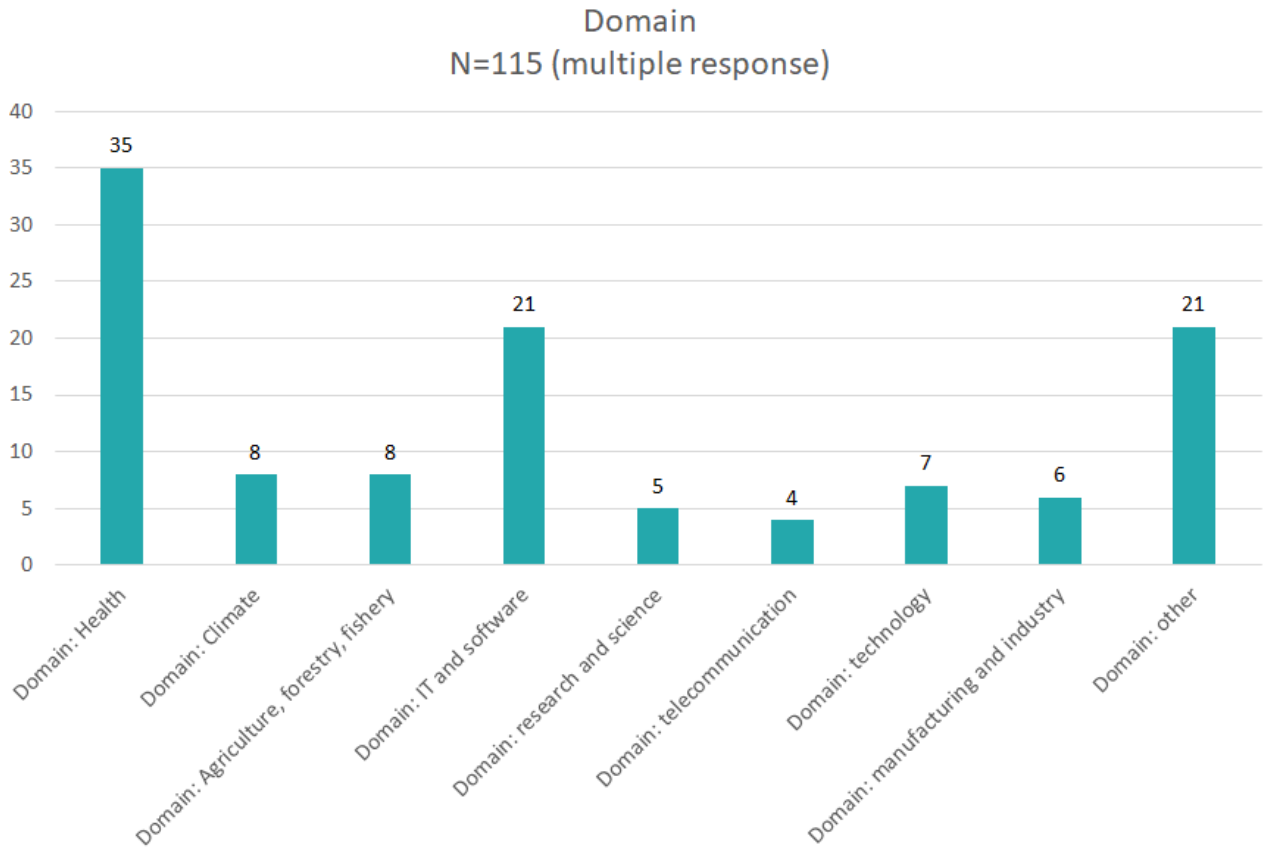


Figure 4: Respondents' company domains

Most of the companies where the respondents work are located in Europe, with Portugal (28%) and Poland (17%) being the most frequently named countries (see Fig. 5). This significant number of participants from Portugal and Poland might be due to an overrepresentation of participants from these countries on Prolific. However, Prolific does not provide detailed information on their participant pool apart from the information that their participants live and/or work in OECD countries (except Turkey, Lithuania, Costa Rica and Colombia) and that they can provide representative samples for the UK and the US.

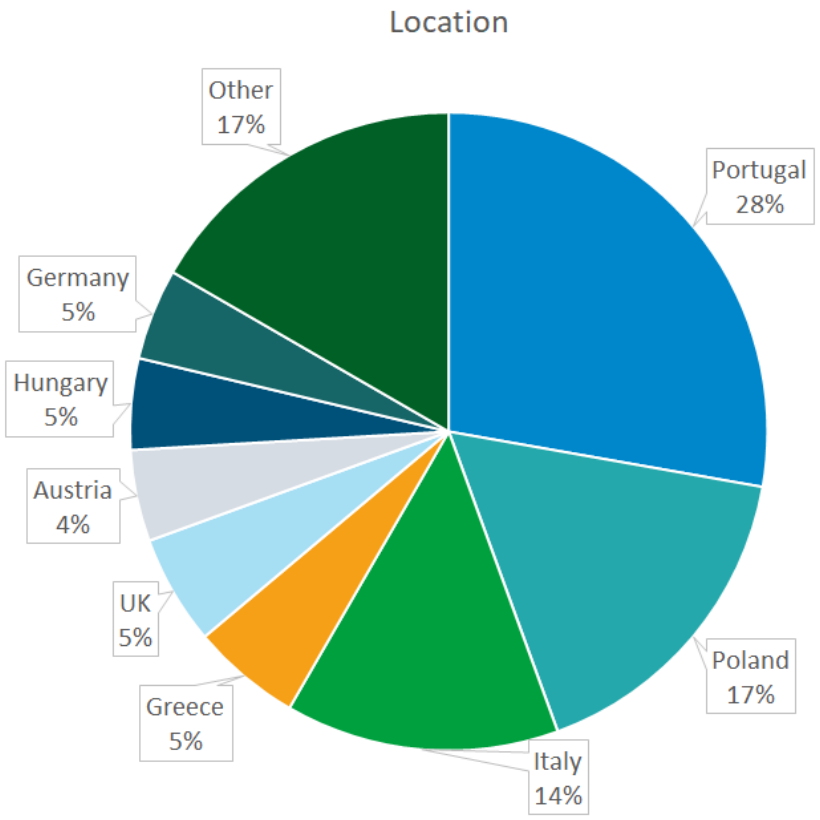


Figure 5: Respondents' company locations

Regarding company size, Fig. 6 shows that one third (33.3%) of the companies have less than 51 employees and 28.7% have more than 500 employees. The remaining companies are in between these poles.

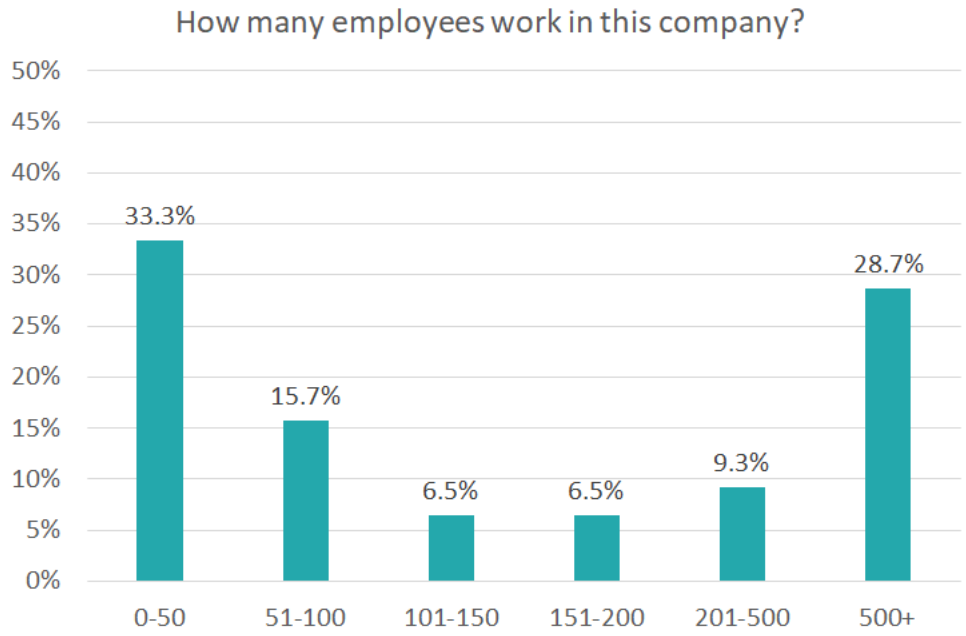


Figure 6: Respondents' company size

Summing up, our sample is predominantly male, the participants have a rather high educational level, most of them work in the health or in the IT domain and more than half of the companies are located in Portugal, Poland or Italy.

4.2. Survey Results

4.2.1. Information Seeking behaviour

In the questionnaire, we wanted to find out how the participants keep themselves and their company up to date with research. Therefore, the section regarding the information seeking behaviour of the participants was opened by a multiple-response question regarding the type of information participants require for their work. In total, 335 factors were reported by the 108 participants (see Fig. 7). Around one fifth of the respondents each selected one, two, or three different types of information they require for their work and around 40% indicated to need four or more different types of information for their work.

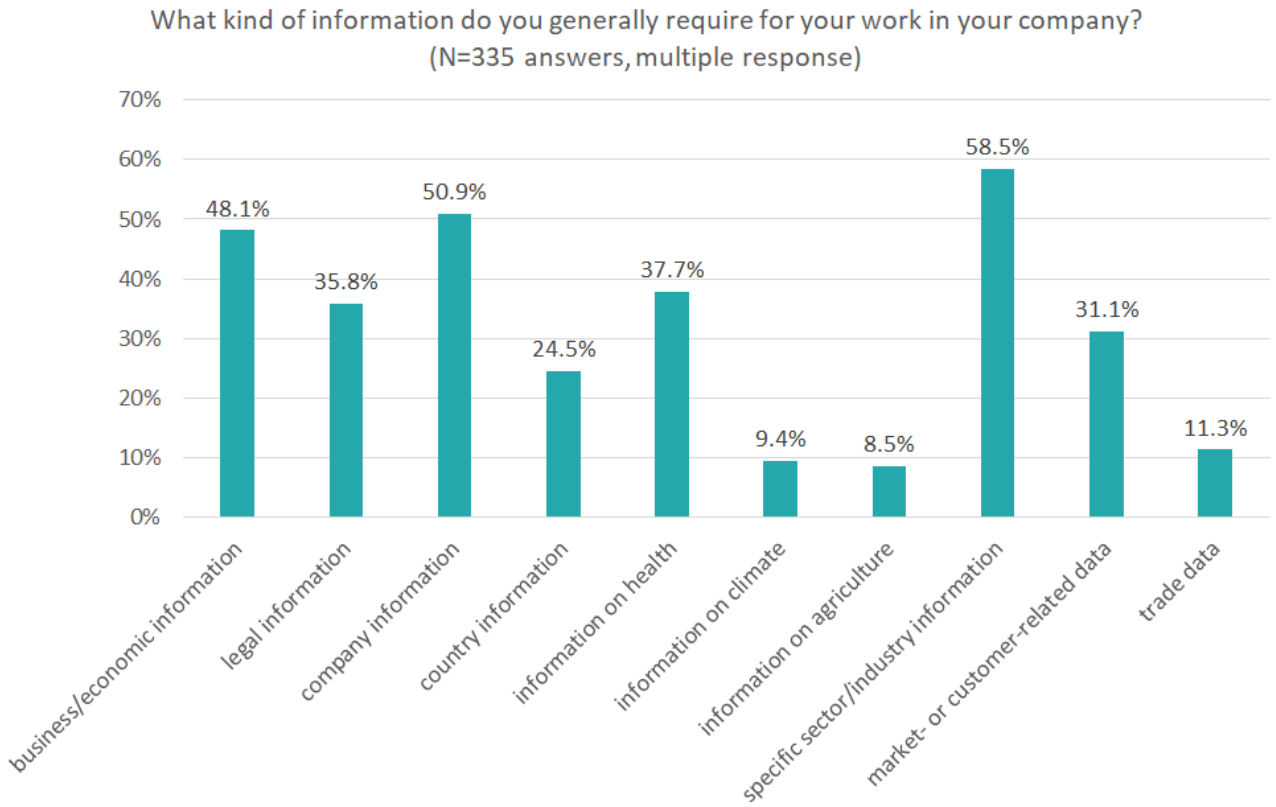


Figure 7: Information required for work

The most important type of information the respondents require for their work is specific sector or industry information (selected by 58.5%) followed by company information (50.9%) and business information (48.1%). Least important is information on agriculture (8.5%) and climate (9.4%) and trade data (11.3%).

Subsequently, the respondents were asked to indicate on a 5-point scale (1=never to 5= always) how often they use different information sources like printed newspapers/industry magazines or the internet. Results

(Fig. 8) show that the most important information source is the internet, i.e., search engines, websites and the like (*Mdn*=5, *IQR*=1). The second most important (even slightly more important than the Internet for the agriculture/forestry/fishery domain) are company internal materials and resources (*Mdn*=4, *IQR*=2), followed by professional networks, including supervisors, co-workers, etc. (*Mdn*=4, *IQR*=1). Patent databases, libraries/information centers and printed newspapers/industry magazines are among the least important information sources (all with a *Mdn*=2 and an *IQR*=2).

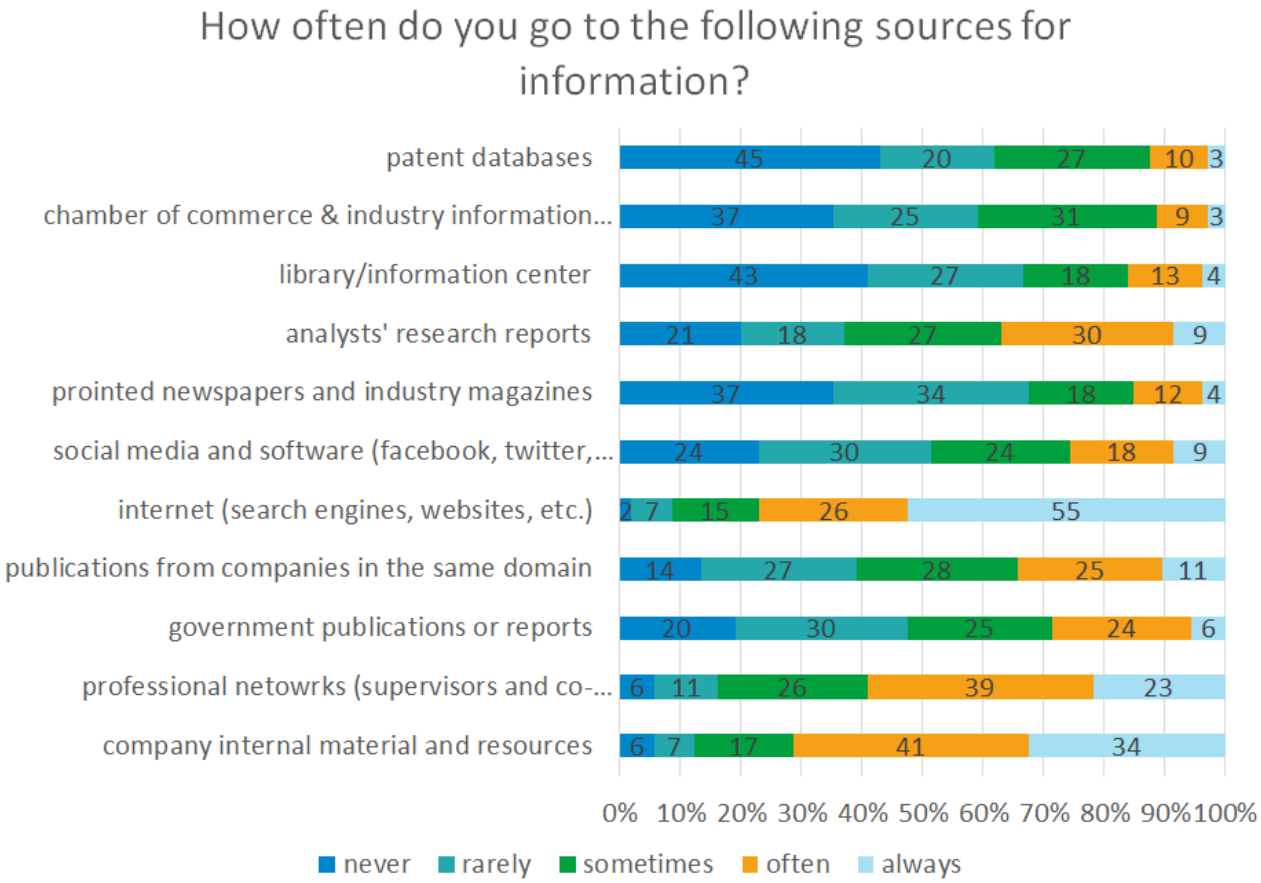


Figure 8: Relevance of different information sources

Searching for information takes up a significant amount of time in many of the respondents' daily work. As can be seen in Fig. 9, more than 70% of the participants spend on average at least three hours per week on information search. Around a quarter spend at least seven hours/week on searching information, while around 10% spend more than 10 hours a week.

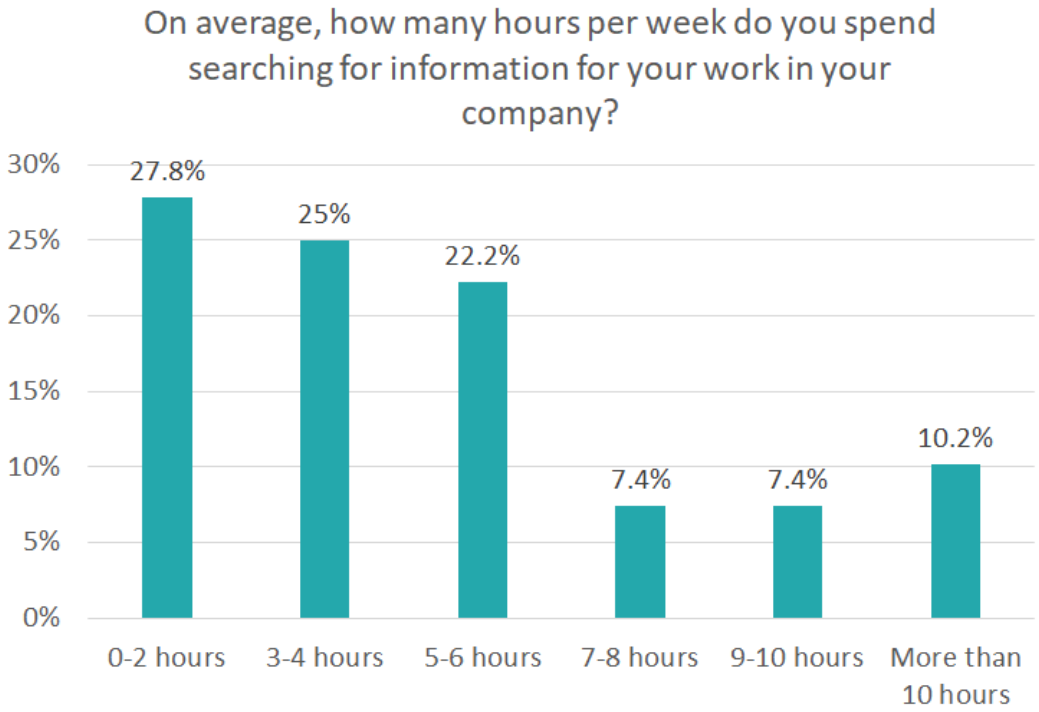


Figure 9: Hours per week spent searching for information

The information search process, however, does not always go smoothly. Respondents were asked to indicate how often they encounter specific barriers when searching for information. Their responses are depicted in Fig. 10.

In contrast, some of our interview partners were not that concerned about the loss of data. They believe that data alone is not enough; there is a need for competence and experience in the corresponding domain before someone can really benefit from the data.

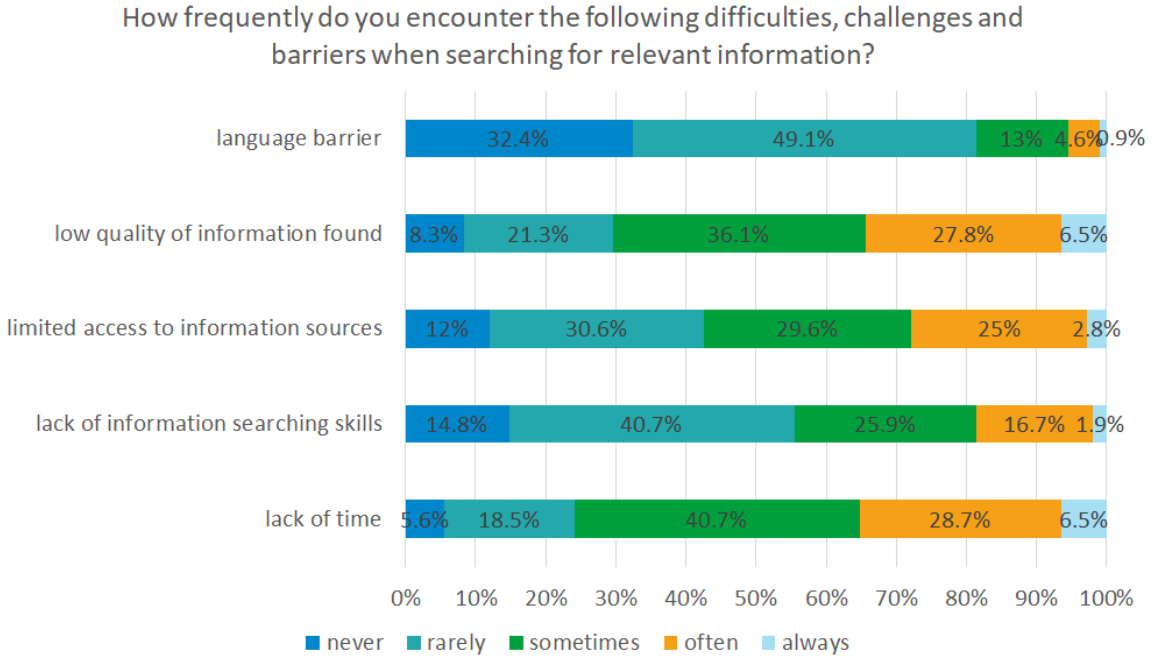


Figure 10: Barriers in relation to information search

The most significant barrier is a lack of time, followed by low quality of the information found and limited access (e.g., paywall) to information sources. Language barriers are not perceived as a relevant obstacle by most of the respondents. Additional barriers named by the respondents in freetext form concern - to name a few recurring answers - the outdatedness of information available, uncertainty about the reliability of information (sources) and a lack of skills to assess it, specificity/superficiality of information available, a lack of support by the management, and infrastructure issues like a bad internet connection.

4.2.2. The Uptake of Open Science Resources

Baseline: familiarity with the concept of Open Science

Next in the questionnaire, we wanted to find out which levels of uptake of Open Science resources already exist in the different domains. Therefore, after the section concerning their information search behaviour, the participants received a brief introduction to the concept of Open Science to prepare them to answer the questions in the subsequent section regarding the uptake of Open Science (resources). Because three participants dropped out after the information search behaviour section, the sample the following analyses are based on is slightly smaller (*N*=105). The section concerning the uptake of Open Science (resources) was opened with a simple question to capture participants' familiarity with the concept of Open Science (see Fig. 11).

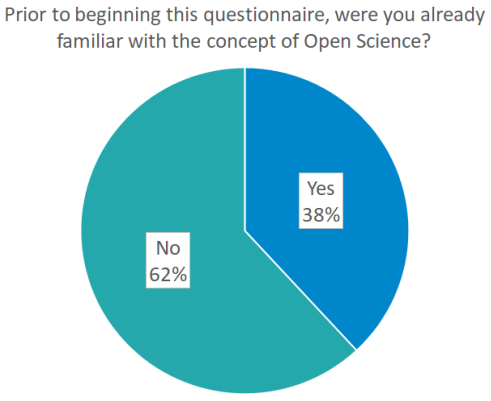


Figure 11: Familiarity with Open Science

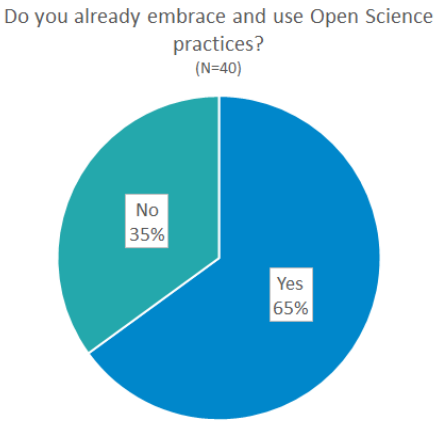


Figure 12: Use of Open Science practices

The majority of the respondents were not familiar with the concept of Open Science prior to the survey. Out of the 37% (n=40) who knew the concept before, around two third (65%) indicated that they already embrace Open Science practices (see Fig. 12). Participants who were familiar with the concepts were slightly older (M=31, SD= 10.27) than participants who did not know it (M=27.83, SD=7.865), they were more likely to have a managerial/senior position (44% vs. 34.7%) and at least a bachelor's degree (42.6% vs. 29.7%). Also, more men (42%) than women (31.4%) knew what Open Science was before taking part in the survey. However, none of these differences/relationships were statistically significant.

An analysis of respondents' familiarity with Open Science by domain revealed that participants from the health domain were more likely to be familiar with Open Science (45.2%) than participants from the climate (42.9%) and agriculture/forestry/fishery (33.3%) domains (see Fig. 13). Half of the respondents from the health domain who indicated that they are familiar with Open Science stated that they already embrace and use Open Science practices, while this was the case for 100% of the respondents from the climate and agriculture/forestry/fishery domains, respectively.

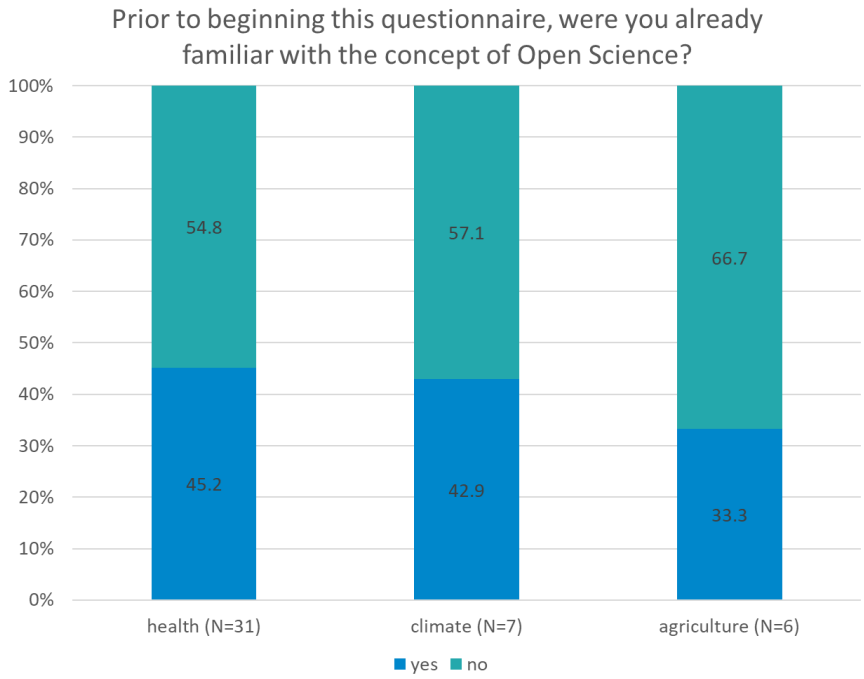


Figure 13: Familiarity with Open Science per domain (health, climate, agriculture)

Due to the very small group sizes of the climate and the agriculture/forestry/fishery domains, for further analyses we decided to focus on the health domain and compare it to the IT domain (next largest group) and the (although very heterogeneous) pooled remaining domains. Fig. 14 shows that in comparison to the IT domain and the pooled other domains, respondents working in the health domain were slightly more likely familiar with the concept of Open Science prior to the survey.

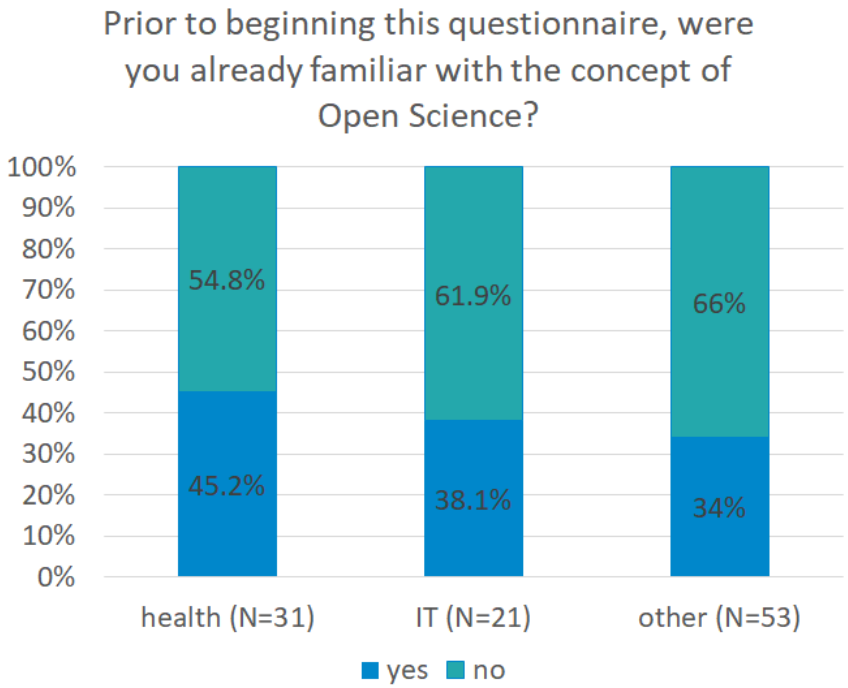


Figure 14: Familiarity with Open Science per domain (health, IT, other)

Additionally, we asked the respondents to give us some additional thoughts about the topic of Open Science. The following thoughts were given from respondents who were not familiar with the concept of Open Science before the questionnaire: "*I have never heard of it before [...]*" (Health Domain); "*I don't have yet a clear idea about this subject*" (Technology Domain); "*Frankly speaking, as previously stated I had no information regarding Open Science, I'm basically interested but I would need additional info to take a decision.*" (Health Domain); "*This topic is very interesting and I didn't know anything about it. I really believe that the company and the employers would benefit immensely from it. Access to knowledge is the best way for a company to progress and adapt to the future.*" (Agriculture/Forestry/Fishery)

From those who were already aware of the concept, we got some very positive comments: "*It is an implementation of a new and antagonistic view of the modus operandi in the previous scientific culture, closed and very focused on competition; in this way it is possible to remove synergies from the scientific development produced globally, accelerating new technological implementations and new and more impactful scientific knowledge.*" (Health Domain) "*Prevents a lot of research from being performed multiple times unknowingly and having none of it available to other researchers*." (Research/Science) and "*I think, Open Science would be a major boost to increase transparency in EU policy (given that proper security checks are in place) and in Energy. For the latter, it would probably stimulate minds to enable and foster new energy sources or takes on existing ones without requiring huge exploration and research costs (mainly data gathering) that sometimes kill a project.*" (Climate Domain) While the previous answer had a very positive connotation towards Open Science, one statement of a respondent criticises the implementation of Open Science practices in large companies. "*Necessary but uninteresting topic. Tends to cause a lot of confusion as it is something needed in a big and organized company such as mine but ends up being a mess of a situation due to lack of proper management.*" (IT Domain).

Current Open Science activities/actions of companies

Regarding the question of whether their company somehow conducts or contributes to research (e.g., collecting and/or processing of data and information, either for internal purposes or beyond), twelve respondents did not give an answer and out of the remaining 92, 62% stated that their company is somehow engaged in research (-like) activities. This includes 69.2% of the companies in the health domain, 50% of the companies from the IT domain and 62.5% of the companies from the other domains.

Thirty-five of the respondents indicated that their company is not at all involved in any data or information collection and/or processing activities, neither for internal nor for other purposes. Due to this question's function as a filter question, these respondents did not see the subsequent questions regarding the frequency in which Open Science activities are undertaken by the company. For those who did see these questions, we included a response option to each of the frequency-questions that allowed the participants to indicate that they lack information to answer the question ("I don't know/I don't have enough information"). Those who selected this option were not considered in the following analysis. The following analysis (Fig. 15), therefore, is based on a reduced sample of not more than 70 participants (actual sample size is indicated for each question).

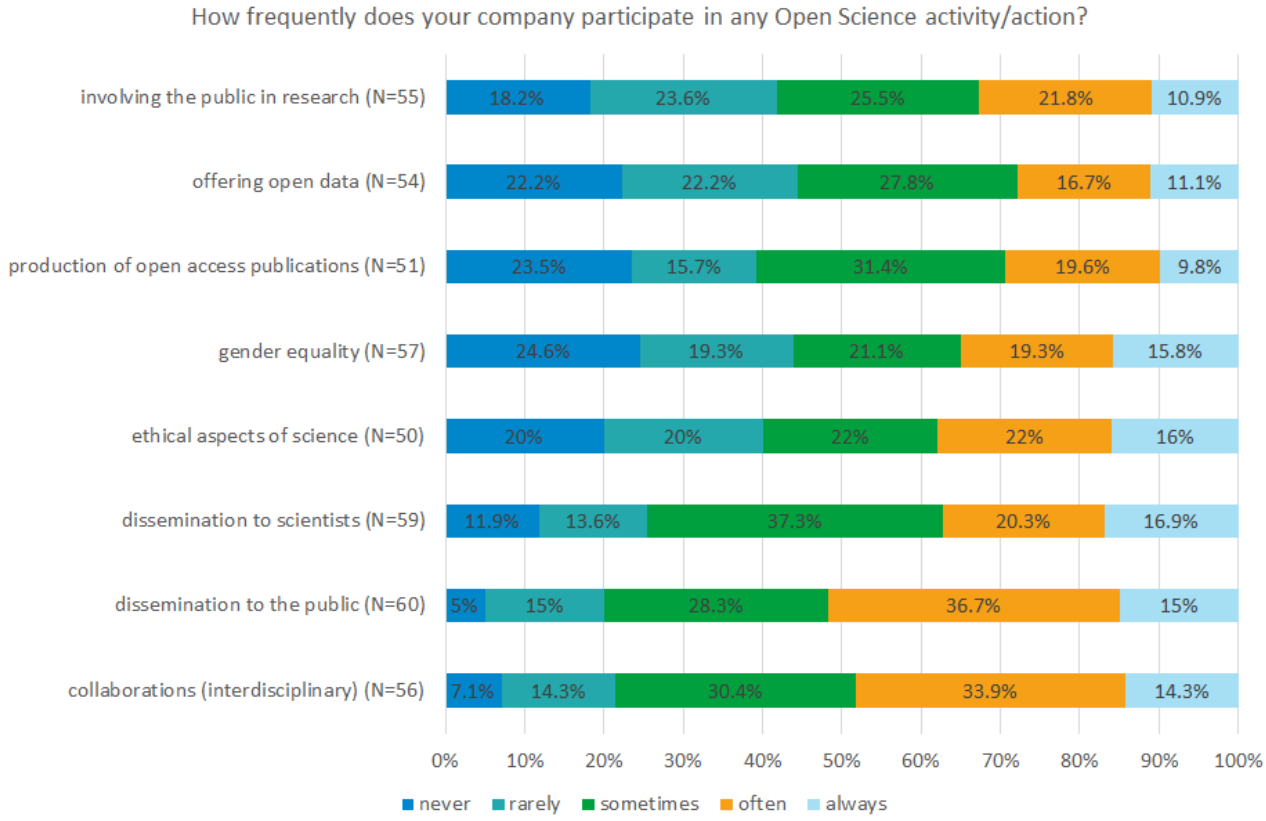


Figure 15: Companies' participation in any Open Science activity/action

Amongst the Open Science activities depicted in Fig. 15, dissemination to the public (e.g., social media activities, articles or talks targeting the public or media relationships, etc.) and collaborations across institutions and disciplines (e.g., interdisciplinary groups, projects, or meetings, etc.) were the most prevalent. In contrast, offering Open Data as well as the production of Open Access publications were the least popular activities.

While offering Open Data is not common practice in many of the companies where the respondents work, re-using data from external sources is. More than 90% of the companies (note that this question was presented to all participants, meaning that *n*=105) re-use data from at least one domain from external sources (see Fig. 16). More than half (55.3%) of the companies re-use data from more than one domain. Only 10 participants (9.5%) stated that their company does not re-use any kind of data from external sources.

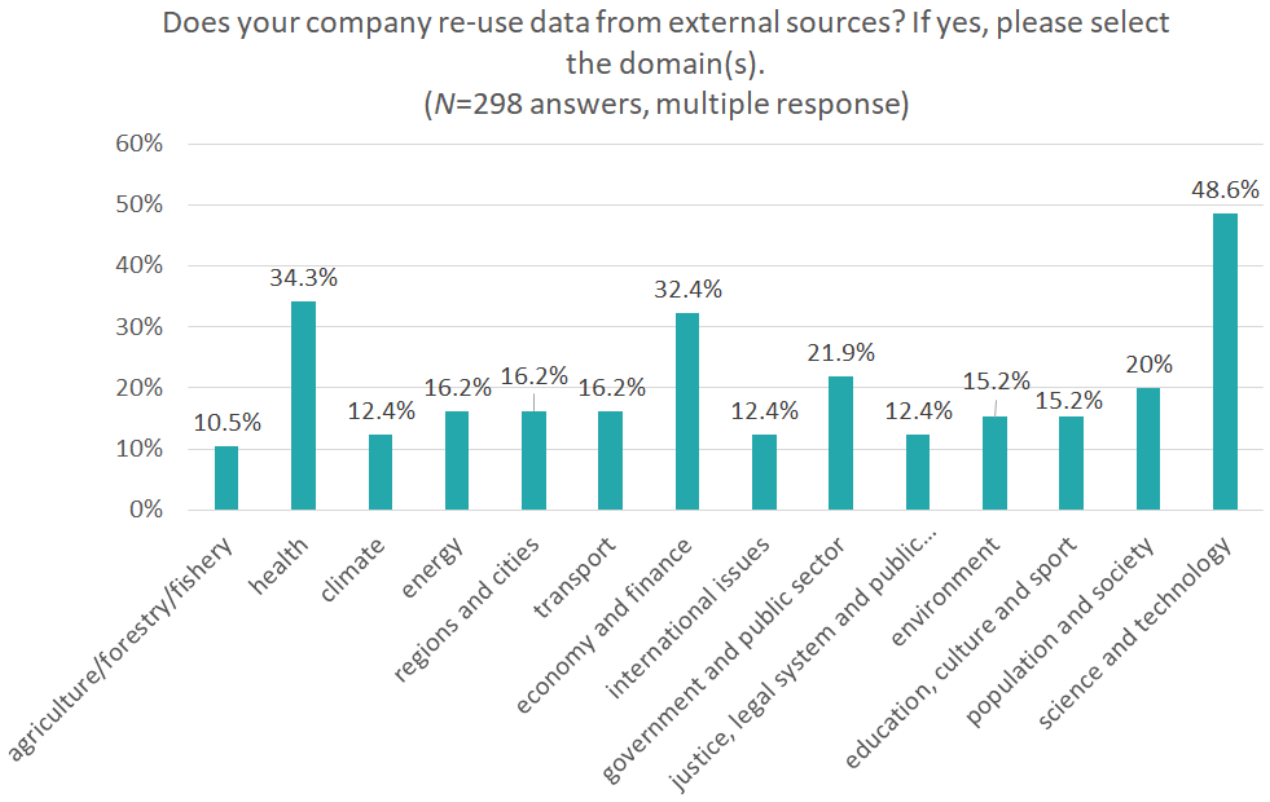


Figure 16: Companies' re-use of data from external sources

Data from the domain of science and technology is re-used by 48.6% of the companies making it the most frequently re-used external data, followed by health data (34.3%) and data from the domain of economy/finance (32.4%). A more detailed analysis regarding re-usage of data from science and technology revealed that respondents from the IT domain make up 31.4% of the respondents who indicated that they re-use data from science and technology and within the IT domain 76% indicated to re-use data from this area. Since our main focus is, amongst others, on companies from the health domain, the popularity of health data is no surprise and almost 70% of the respondents who stated that their company re-uses health data work in the health domain. Moreover, more than 80% of the companies in the health domain in our sample re-use health data from external sources.

Areas of Interest relating to Open Science activities/actions

A section of the questionnaire was dedicated to the readiness of companies for the exploitation of Open Science in terms of companies' interest in working with Open Science resources and the degree to which Open Science is seen as an opportunity by decision makers in SME and industry. As is shown in Fig. 17, there is a certain interest in working with Open Science resources. Open Data and Open Source software are

especially interesting for companies. In contrast, both reproducible research and ethics in research are not topics of interest for more than a third of the companies, respectively.

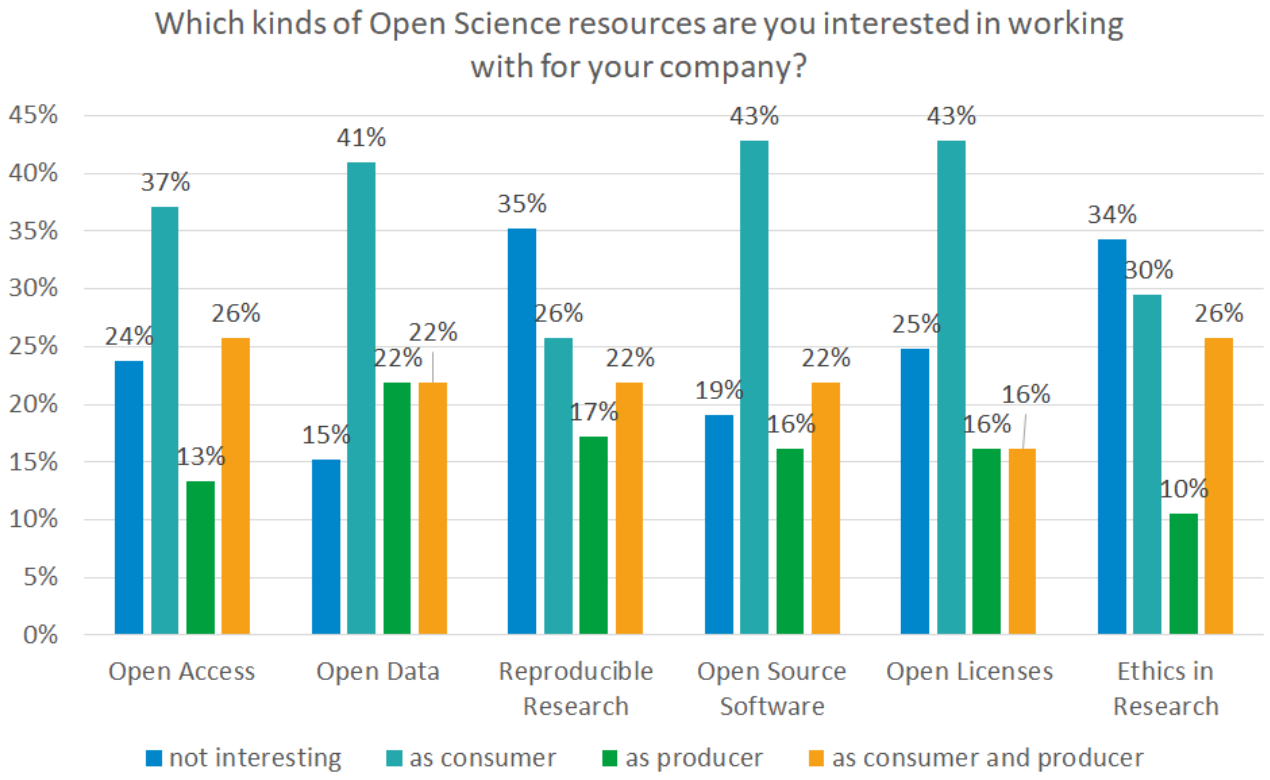


Figure 17: Respondents' interest in working with Open Science resources

Since Open Data is a crucial Open Science resource, particularly for companies, we added a question to the survey that should capture participants' interest in Open Data in more detail. This question has only been presented to those who had indicated to be interested in working with Open Data as consumer, producer or both, reducing the sample for this specific analysis to 88 participants. Fig. 18 shows these participants' answers.

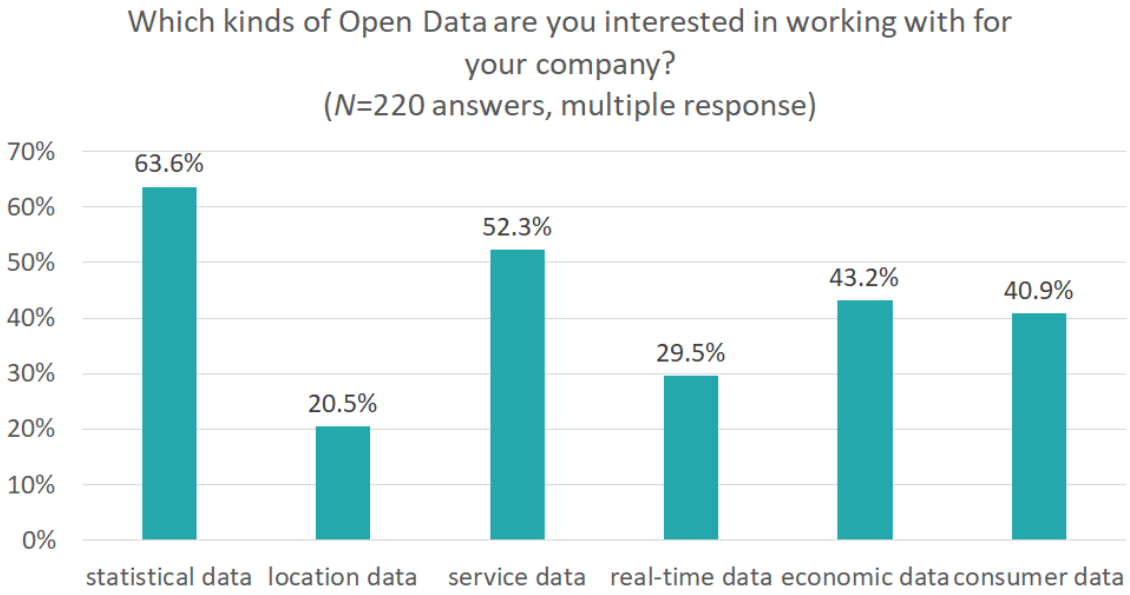


Figure 18: Respondents' interest in working with different types of Open Data

More than 60% of the respondents stated that they are interested in working with statistical data, while only a fifth of the respondents indicated interest in working with location data.

4.2.3. Absorptive Capacity and Business Model Archetypes

Absorptive capacity of a company can be a good indicator of the company's openness to Open Science activities (Huber, Wainwright, and Rentocchini 2020). In our questionnaire, the respondents were asked to indicate whether they agree or disagree (5-point Likert-scale; 0=strongly disagree to 4=strongly agree) with a set of statements related to the four dimensions of absorptive capacity proposed by (Flatten et al. 2011), namely knowledge acquisition, knowledge assimilation, knowledge transformation, and knowledge exploitation. Descriptive statistics for the four subscales are depicted in Table 3.

	Mean	SD
Acquisition	2.68	0.716
Assimilation	2.63	0.832
Transformation	2.95	0.634
Exploitation	2.82	0.788

Table 3: Mean values and standard deviation of the four dimensions of the absorptive capacity

Overall, the respondents rated the absorptive capacity of the companies they are working at as rather high, on average. A closer look at the perception of the companies' absorptive capacity in the domain of health compared to the IT domain and the pooled remaining domains showed that there is almost no difference between the health domain and the other domains (see Fig. 19). Accordingly, no statistically significant difference was found in a subsequent one-way ANOVA.

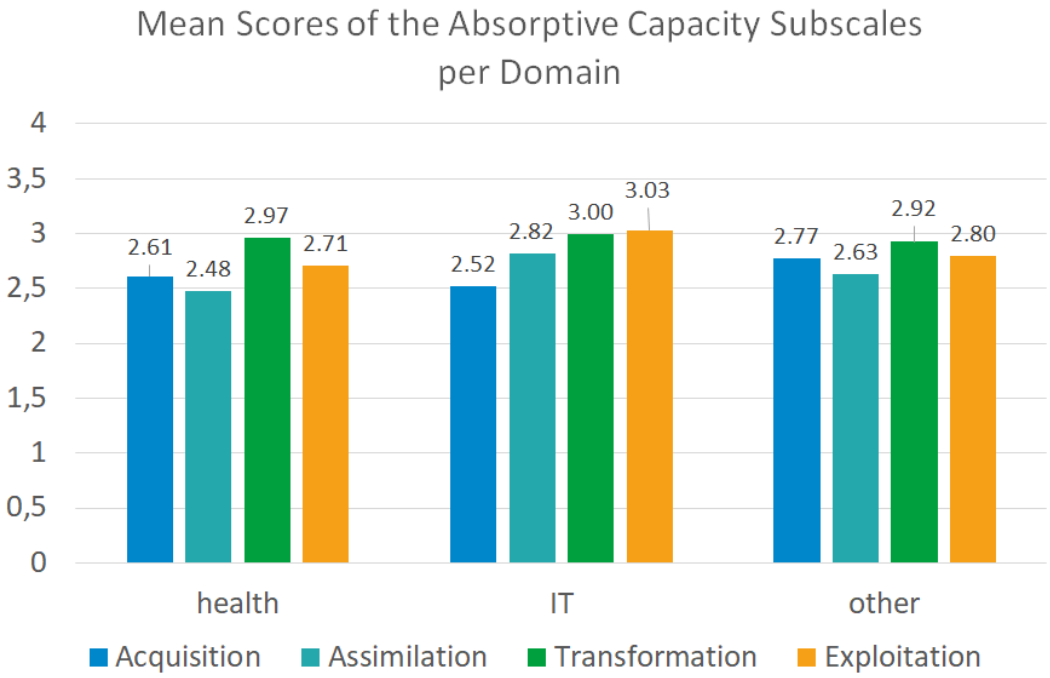


Figure 19: Mean scores of the absorptive capacity subscales per domain

Another construct that has been linked to the uptake of Open Science in companies, as discussed in the report on the interview study is the concept of Business Model Archetypes (Deloitte, L L P 2012). The participants, therefore, were asked to indicate (multiple response) which of the archetypes (supplier, aggregator, developer, enricher, enabler) fits their company at the time of the survey and in the future. Fourteen participants stated that none of the archetypes fit their company at that time. The distribution of the answers of the remaining participants can be seen in Fig. 20.

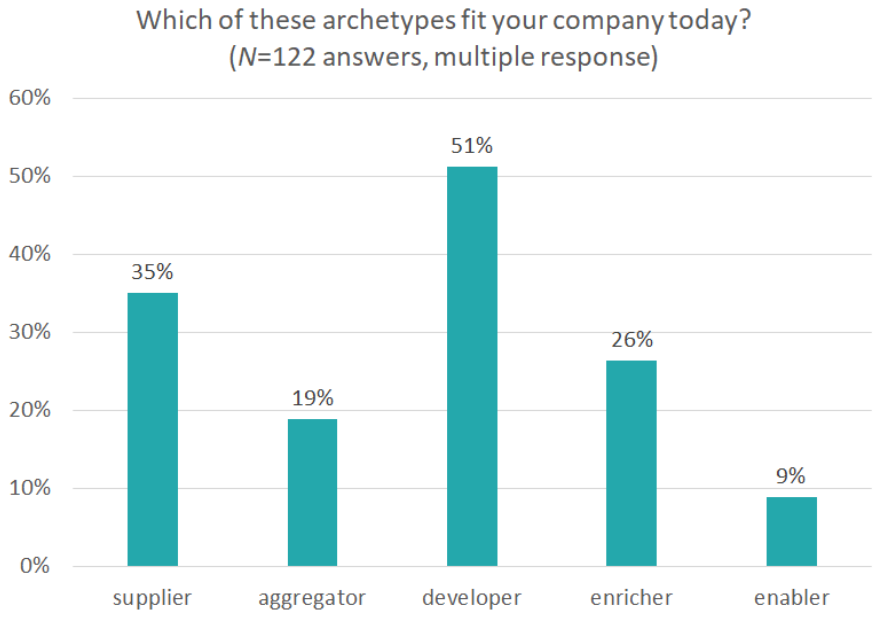


Figure 20: Companies' business model archetypes

More than half of the respondents perceive their company as “developer” and less than 10% think that their company fits the archetype “enabler”.

The second question related to Business Model Archetypes was which archetype might fit the respondents’ company in the future. The same 14 participants who indicated that none of the archetypes fit their company at the time of the survey also stated that no archetype would fit the company in the future. The remaining respondents’ answers are depicted in Fig. 21 and show a rather similar picture as the answers to the preceding question. However, there are certain important differences detectable, namely that in the future the archetype “developer” has less importance and companies might develop towards “enabler” and “enricher” in the view of the participants.

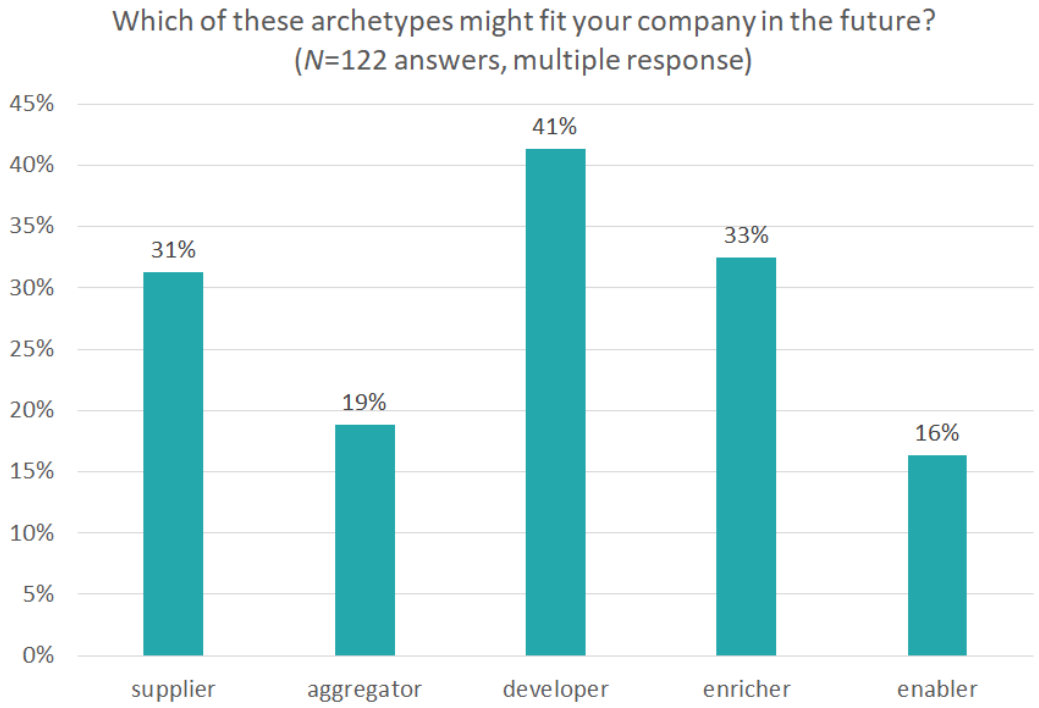


Figure 21: Companies’ business model archetypes in the future

4.2.4. Barriers and Drivers for the Uptake of Open Science in SME/industry

The participants were asked to rate a set of potential barriers to the uptake of Open Science in their everyday work, according to whether they perceive them as not relevant, as a minor barrier, as a significant barrier or as a very significant barrier. As can be seen in Fig. 22, a lack of clear steps to follow, e.g. How do I begin? How do I proceed?, is perceived by the majority of respondents as (very) significant barriers. Around 60% perceive a lack of clarity on where to find relevant data as a (very) significant barrier, and time constraints are also a noteworthy barrier to the uptake of Open Science activities in the participants’ everyday work.

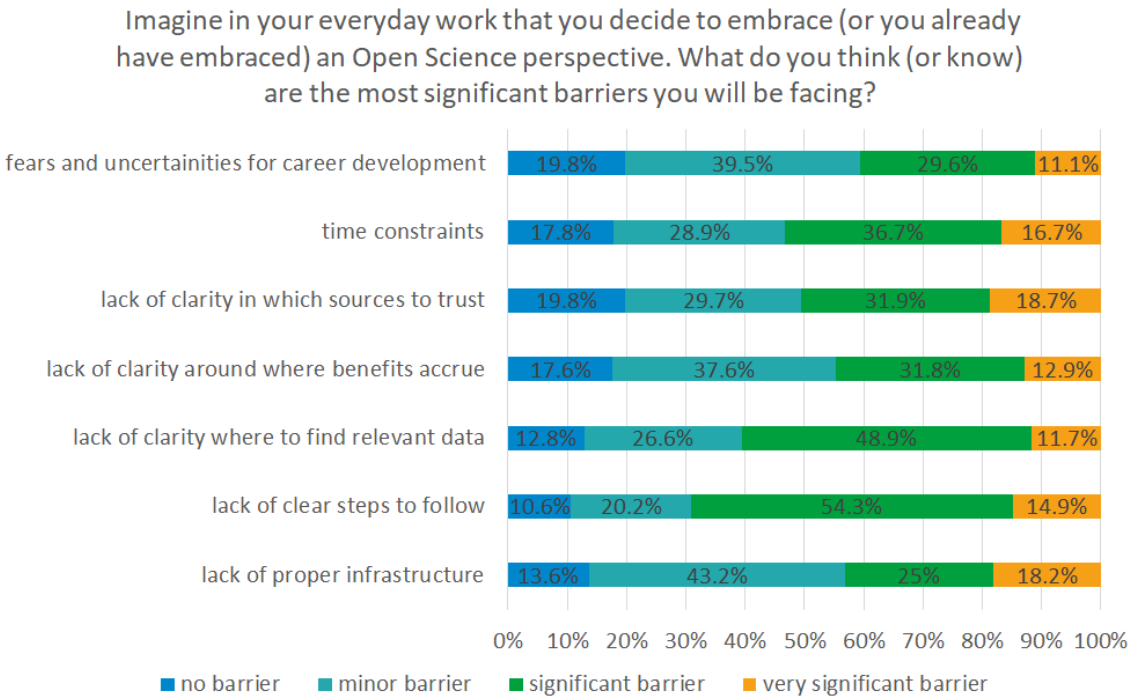


Figure 22: Barriers for the uptake of Open Science

Another question targeted a related topic, namely, whether the respondents' company offers them incentives or support related to Open Science. Respondents who indicated to not know or not having enough information to answer the question were excluded from the analysis. Fig. 23 shows that the majority of the firms (more than 60%) the remaining respondents work at, offer technical infrastructure to support Open Science activities. However, more than a third of the respondents (33.7%) would like to receive more support related to technical infrastructure. Financial support and rewards for Open Science related activities are provided by 42% of the firms. This kind of incentive, however, is not provided but also not needed in around a third of the companies, according to the participants' answers. A quarter of the participants indicated that financial support/rewards for Open Science activities are not offered by their companies, but they would like to get such incentives. Around half of the companies offer career perspectives/recognition and specialist/expert support as support/incentive for Open Science related activities, respectively. While a quarter of the respondents reported to receive adequate specialist/expert support, and almost a third indicated that they neither receive nor need such support, only slightly more than a quarter of respondents indicated that career perspectives/recognition are not offered and not needed and around a fifth reported that this incentive is provided adequately.

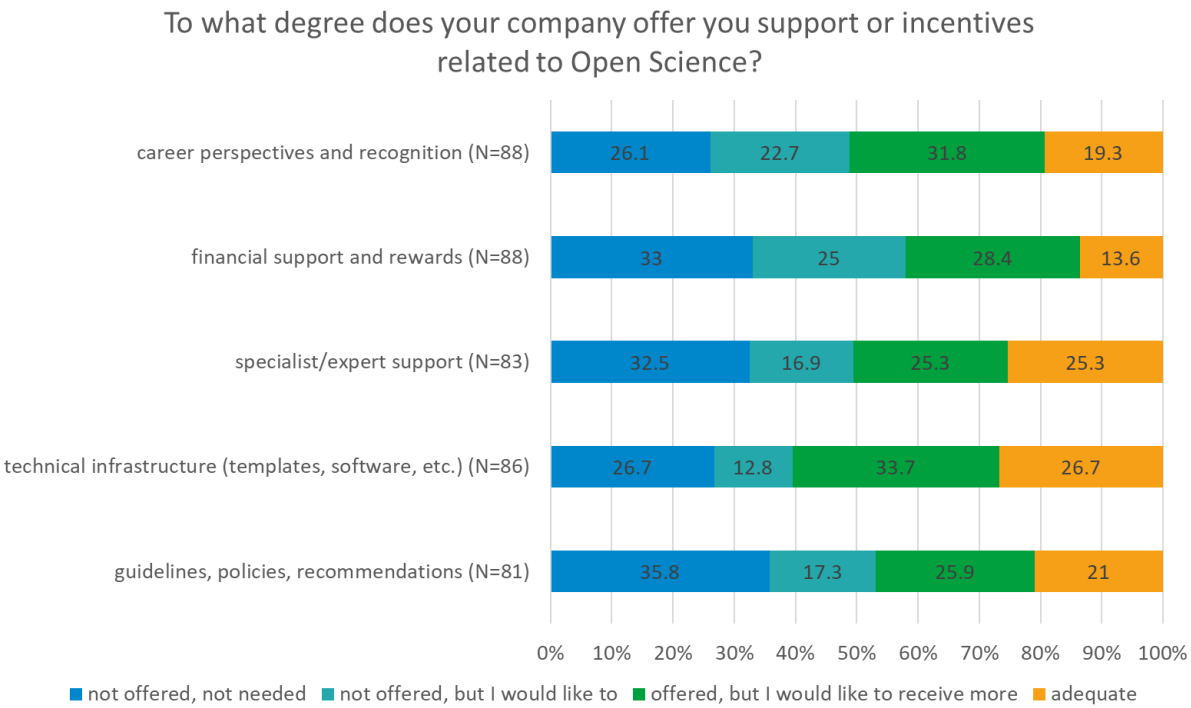


Figure 23: Companies' incentives or support for Open Science

Regarding possible drivers for the uptake of Open Science activities and resources in companies, some of the respondents' answers to open questions throughout the survey provided valuable insights. Some survey participants (n=4) mentioned specific education and training programs/activities for employees provided by the company as a relevant activity in the context of Open Science. One respondent from the health domain, for example, stated: *"We have education programmes in some projects to ensure that there is a bridge between experts and the general public"*.

Another important driver for Open Science in the private sector are inter- and transdisciplinary collaborations and events (6 times mentioned explicitly as a driver), especially collaborations with actors from academia but also activities and events involving the general public. The following thought of one of the respondents from the health domain on the impact of digital technologies on the relationship between science and the general public illustrates that also digital technologies are seen as potential drivers for Open Science: *"The use of digital social networks such as Twitter, Facebook, YouTube, and others is rapidly changing the relationship with the public, allowing scientists to communicate directly and without mediation with different actors."*

5. Discussion

In this section, we will discuss our findings along with the research questions posed at the beginning. We will first describe the results of the interview study, followed by the results of our survey study and then refer our findings back to literature.

5.1. Knowledge, Baseline and Understanding of Open Science?

In this section we address our first research question (RQ1): "What is the common knowledge, baseline and understanding of Open Science?"

All of our interview partners, located in Austria, were aware of the concept of Open Science in general and five of them really highlighted the value of the Open Science endeavours. This could be explained with the fact that we had a bias in relation to our interview partners as mostly all of them were somehow in contact with academia and therefore closely related to or aware of academic research. Most of them confirm to use either Open Access publications, Open Data or Open Source code. We found lower awareness through our EU-wide survey, which found that only 38% were familiar with the concept of Open Science. From these, about two thirds (*n*=26; around 24% of the entire sample) indicated that they already embrace Open Science practices. Their demographic background shows that they are slightly older, they are more likely to have a managerial/senior position and have at least a bachelor's degree, in comparison to those who were not aware of the concept of Open Science. Together, the results of the interview study and the questionnaire reveal that education level is correlated with awareness of Open Science. Those with a bachelor's degree or higher are more familiar with the concept of Open Science than others. This leads to the assumption that Open Science is not a topic in education before university/higher education which further could lead to a policy recommendation such as: In order to promote Open Science in SMEs and industries, make the concept of Open Science more popular, especially beyond the university context, thus already in secondary school or, where applicable, even earlier. Encouraging schools to, for instance, participate in citizen science projects could be one concrete approach to do this.

5.2. Search for Scientific Information within SMEs/Industry?

This section addresses the results related to the second research question: "How is the search for scientific information currently conducted within SMEs/industry?" Based on our literature review presented in D4.1, we discuss the search for information from three different perspectives, namely the sources of information, gender aspect regarding search, and barriers and challenges.

5.2.1. Sources of information

Sources of information: In the interviews we saw that the most important sources of information are literature/publications, social contacts, conferences, webtools, and patent search. In the survey, we found that the most important information source is the internet, second most important are company internal material and resources followed by professional networks, including supervisors, co-workers etc. Thus, the results of both studies correspond with each other and emphasize the literature/publications accessible via the internet and digital sources as well as social contacts as most important. The only difference we could show is that for the interview partners, patent search is more important than for the participants of the questionnaire study. This could be explained with the fact that our interview partners i) are mostly founders,

CEOs and managers, thus in leading positions and ii) that they are working in small companies or start-ups that need to establish themselves in the corresponding sector and need to keep themselves their direct competitors, their services, products, and patents under observation. In contrast, most of the questionnaire respondents were only employees without a leading position, thus, keeping an eye on competitors and patents is not included in their daily working tasks.

With regard to the three domains, we could show from the interview study that for the health and climate domain, literature/publications are the most important sources of information, while for the agricultural domain the most relevant source of information seem to be social contacts. Small differences in the preferred information source between the three domains were also found in the survey: while the internet (i.e., search engines, websites and the like) is the most important information source for participants working in the health and climate domains and company internal material is the second most important source for each, for the agriculture/forestry/fishery domain, it is the other way around. However, due to the low number of participants from the climate and agriculture/forestry/fishery domains in the survey and the overall low number of participants in the interview study, this can only be a first impression and needs further investigation.

Our findings with these two studies are also in line with the results of our literature review presented in D4.1 (Fessl et al. 2020) of ON-MERRIT. Our literature review showed that digital sources (e.g. the internet, web sites, blogs), analogue sources and social sources are the most relevant ones. The newer the publications reviewed were, the more important the digital resources are for searching for information (Li et al. 2019; Lundin and Eriksson 2018; Jones 2016; Le et al. 2016; Freund 2015; Starasts 2015).

Types of information: From the interview study, the most relevant type of information is literature that is related to their research topics or domain knowledge. This is followed by information they get through their social network/contacts, thus, it is the knowledge of people, business relationships or scientific experts, including new scientific research approaches, methods or results or estimations about the market, for example.

The questionnaire study showed that the most important type of information the respondents are looking for is specific sector or industry information followed by company information and business information. Least important are information on agriculture and climate and trade data. Bringing these results from both studies together, it seems that our interview participants are very interested in the state-of-the-art research in their domain related topic, while the respondents of the questionnaires seem to be more interested in business related facts about their companies and their corresponding domains. This could be explained with the fact that all our interview partners have a scientific background in that all of them have a master's degree and four hold a PhD degree. Moreover, most of them are still closely related to a university to stay up to date with research in their corresponding domains. In contrast, only 35% of the questionnaire respondents hold a masters or PhD degree, thus, most of the respondents are interested more in business facts than in ongoing research.

Hours per week spent for searching: From the interview study, we know that searching for relevant information belongs to the interviewees' daily business, however, the exact amount of hours they spent searching for information per week was not discussed during the interviews. Searching for information takes up a significant amount of time in the daily work of the questionnaire respondents. More than 70% of the

participants spend on average at least three hours per week on information search, around a quarter spends at least seven hours/week, while around 10% spend more than 10 hours/week on searching information. Thus, searching for information is nowadays an important component of work.

Summing up, we can answer our research question about how the search for scientific information is currently conducted within SMEs/industry, as follows: Literature/publications accessible via the internet and digital sources as well as social contacts are the most important sources of information. The most relevant types of information of our interview partners is the current state-of-the-art research in their domain while the survey respondents seem to be more interested in business related facts about their companies. Both studies confirm that searching for information is part of the companies' daily business.

5.2.2. The Role of Gender Regarding Information Seeking Behaviour?

We also investigated whether gender plays a role regarding information seeking behaviour, however, we found no obvious differences between men and women. Only two women participated in the interview study, therefore, we cannot derive any insights regarding the relation between information seeking behaviour and gender. And although 33% of the questionnaire participants were female, we could not find any statistically significant differences in the information seeking behaviour of men and women. Both our study results as well as our literature review presented in D4.1 (Fessl et al. 2020) showed that there is a research gap regarding gender and its impact on information seeking behaviour that is worth being investigated more in the future. However, we did not uncover any insights regarding gender and search behaviour. And in our literature review (ibid), we found only one paper dealing with this issue (Le et al. 2016).

5.2.3. Barriers and Challenges

From our interview study we derived the following barriers and challenges regarding information seeking: First, there is knowledge that is domain specific and cannot be "googled" as it consists of intrinsic knowledge of individuals, people or experts, which is sometimes not yet published, or which is not intended to be published at all. Second, although search engines are very powerful today, the large amount of data available is still seen as a challenge or barrier to find the relevant and needed information. Third, anonymized, statistical patient data is still scarce. Fourth, paywalls to access journals or conference papers is still a huge challenge in terms of costs, especially for SMEs. Fifth, interoperability between databases, trust and trustworthiness of resources, as well as defining the fitting keywords are challenges that participants experience.

From the questionnaire study, the most significant barrier is a lack of time, followed by low quality of the information found and limited access (e.g., paywall) to information sources. Language barriers are not perceived as a relevant obstacle by most of the respondents. Additional barriers that came up were the outdatedness of information available, reliability of information (sources) and a lack of skills to assess it, specificity/superficiality of information available, a lack of support by the management, and infrastructure issues like bad internet connection.

From both studies we found that the limited access to publications, credibility and trustworthiness of resources and basic information literacy components like finding meaningful and successful keywords are common barriers. Other barriers mentioned seem to be viewed different by participants of the interview study and participants of the survey which could be explained with two reasons: i) all interviewees are in

leadership positions while in the survey respondents were more employees without a leadership role and ii) all interviewees were academics while most of the survey respondents were not, thus they have different purposes to find different types of information.

The barriers and challenges uncovered in our interview and survey study are also mostly in line with our literature review in D4.1 (Fessl et al. 2020). For example, explicating information needs and formulating the fitting keywords was and still is a challenge as highlighted by Kraaijenbrink and Groen (2006). Also, which information sources and documents are used strongly depends on the information quality and the accessibility (Guo 2009; Kwasitsu 2003) as well as the trustworthiness and authoritativeness of the source (Hirsh and Dinkelacker 2004; 2003).

5.3. The Uptake of Open Science Resources in the Different Domains

In this section we address the following research question: “RQ3: Which levels of uptake already exist in the different domains and what differences can be observed across domains?”

The role of Absorptive Capacity and Business Model Archetypes:

Absorptive capacity has been defined as "a firm's ability to recognize the value of new information, assimilate it, and apply it to commercial ends" by (Cohen and Levinthal 1990), 128). From the interview study we could see that all interviewees see opportunities and impact for their company, their services and products when taking up different types of Open Science resources. Staying up to date in research through Open Access, developing more effective and efficient software applications with Open Source code, or comparing own sensor results with similar Open Data samples are just some examples where Open Science resources come into play. With regard to the business model archetypes, we see some differences in the data. In the health domain, they see themselves as Aggregators, Developers or Enrichers whereas being a Supplier or Enabler is not even considerable due to personal, sensitive health-related data and applications and including ethical, compliance and privacy reasons in this regard.

In the domain of agriculture, our interviewees saw themselves as Aggregators that collect and aggregate open or proprietary data. In the climate domain, we found an Enabler and a Developer, and one big company sees themselves as all archetypes but Developers. As we had only a low number of interview participants, and the representation of business models archetypes strongly depends on the underlying business model of the company, these findings need to be carefully considered.

In the questionnaire study, the respondents overall rated the absorptive capacity of their companies rather high, meaning they attributed good absorptive capacity capabilities to them. With regard to the business model archetypes, we could show that most respondents see their companies as Developers, which can be explained by the fact that most of them are located in the IT domain.

In our literature review conducted in D4.1 (Fessl et al. 2020) we found that being able to benefit from external Open Data for open innovation, companies need to acquire absorptive capacity (Zahra and George 2002; Cohen and Levinthal 1990) and capabilities (Huber, Wainwright, and Rentocchini 2020). Especially (Huber, Wainwright, and Rentocchini 2020) detected two research gaps in this regard, namely that scientific literature highlights the value of Open Data with respect to potential opportunities especially for the benefit of Open Data in open innovation practices and that SMEs struggle in the uptake of the benefits from Open

Data as they find it challenging to develop the necessary absorptive capacity. Aligning these insights with our findings from our two studies, we can see that SMEs and industries are already developing absorptive capacity, and especially the results of our interview study shows that the uptake of Open Science resources is already more or less taking place.

The uptake of Open Science resources: From our interview study we observed that all interviewees - independent of their domain - are aware of Open Science and highlight the value of it. With regard to the uptake of Open Science resources, we could show that Open Data, Open Access and also Open Source code already play an important role in the respective companies. Open Data, for example climate or location data (e.g. OpenStreetMap data) as well as economic data were explicitly mentioned as in use for the development of services or products. Open Source code - depending on the licences - is regularly used by the interviewees to further develop their applications, products and services. Domain specific literature is of crucial relevance for all of them to stay-up to date in the corresponding domain and to keep informed about the competitors' progress and development. In this context, it is especially noteworthy that Open Access only plays a minor role because it is not yet established as common practice in the domains of the interviewees. Analysing the differences in relation to our domains of interest showed that, especially in the health domain, free access to publications/literature seems to be more important than in the other two domains. One explanation could be the fact that our interview partners were somehow in contact with the Know-Center and therefore closely related to or aware of academic and corresponding.

The questionnaire results indicate that there is already an interest in the uptake of Open Science resources, and here especially with focus on Open Data and Open Source code/software, in terms of the reuse of data from external sources.

Summing up, we could show that Open Science resources find their way into SMEs and industries, however, this depends on the characteristics of the companies as well as on the education levels of the employees involved. We could show in the interview study who takes up Open Data, Open Source code and Open Access and for which purpose these Open Science resources are used. In contrast, in the questionnaire study we observed rather limited knowledge about the endeavours of Open Science in general, but could also show that, compared to the other domains, Open Science is more likely known in the health domain.

These findings are in line with the results of our literature review presented in D4.1 (Fessl et al. 2020), pointing towards conflicting evidence as to how important academic sources are for companies' innovation processes. Based on a survey conducted with companies in the US, large firms and start-ups are found to use public research more often than established SMEs (Cohen, Nelson, and Walsh 2002). This is somehow confirmed with our interview study, where especially the younger companies heavily rely on Open Source code for their application developments. Another study conducted in the UK found that only a limited number of firms use academic research as a direct source for innovation (Laursen and Salter 2004). While this is confirmed with our interview study, from our questionnaire study we cannot derive the purpose of usage of the Open Science resources. Additionally, Fell (2019) found indicative evidence that Open Science might have a positive economic impact through (a) efficiency gains by free access to Open Access journals or Open data and (b) enablement, thus, the development of new products or services. This can also be confirmed with our results.

5.4. Barriers Hindering the Uptake of Open Science Resources

In this section we address the following research question: "RQ4: What are challenges/barriers that hinder the uptake of Open Science resources". In both studies, we found some barriers and challenges that hinder the uptake of Open Science resources that will be presented in more detail below. From the interview study we got more detailed results than from the survey respondents.

Challenges/barriers for the uptake of Open Data: Our interviewees reported the following challenges to the uptake of Open Data: FIrst, especially for the interview partners from the health domain, one major obstacle is that health-related Open Data is scarce. On the one hand, publishing (even anonymized) health related data is a challenge due to privacy reasons and GDPR. On the other hand, trust, validation, and compliance play a major role in this regard. Thus, publishing health-related Open Data is a challenge per se. Second and applicable for all three domains, there exist licences that hinder the commercial use of available data, which makes it uninteresting for SMEs or industries to use, but is definitely in-line with the Open Science endeavours' goal per se. And third, another barrier is the reliability and validation of data.
From the questionnaires we found a lack of clear steps to follow, e.g. How do I begin? How do I proceed?, is perceived by the majority of respondents as barriers. Around 60% perceive a lack of clarity on where to find relevant data as a barrier, and time constraints are also a noteworthy barrier to the uptake of Open Science activities in the participants' everyday work.

Challenges/barriers for the uptake of Open Access: With regard to Open Access, high fees for publishing Open Access were mentioned as a big barrier from a producer perspective but having free access to publications was seen as huge advantage from a consumer perspective and paywalls and expensive licenses were strongly criticised from some interviewees. In general, as indicated by the results of the survey, Open Access (not interesting for almost a quarter of the respondents) plays a less important role for SMEs and industry than Open Source (interesting for more than 80% of the respondents) and Open Data (interesting for 85%), however, examining the reasons for this could be subject to future research.

Challenges and barriers for the uptake of Open Source code: First, there are licences that allow the usage of a source code in general but not for commercial purposes. This is a barrier for the uptake - but fully in line with the idea of the Open Science endeavours as a whole. Second, especially as an SME or start-up, one has to invest time to carefully consider what to share with the coding community and what not. Third, if a company uses Open Source code one cannot ensure and guarantee the reliability of the code, which could result in compliance and validation issues with regard to customers.

The questionnaire results confirmed more or less our findings from the interviews. We found significant barriers for the uptake, namely a lack of clarity where to find relevant information/data, which steps to follow and which sources to trust as well as time constraints related with an uptake.

Our literature review in D4.1 (Fessl et al. 2020) revealed similar findings. Paywalls play a crucial role as well. While larger companies tend to have higher rates of subscriptions to journals and have access to a higher number of scientific resources, SMEs and start-ups mostly cannot afford these fees although they rely on the access to basic research to enable their business (Fell 2019; Savage 2016). In addition, this finding is also in line with ElSabry and Sumikura (2020), who show that many SMEs suffer from high journal prices. Second, high levels of firm-internal R&D, rely on scientific resources to a considerable degree (Bergman 2010) and

the uptake of research papers and datasets in these industries goes hand in hand with companies also relying on knowledge sourcing by hiring graduates or directly collaborating with universities (Simeth and Raffo 2013; Veugelers and Cassiman 2005). And third, lack of skills and time prohibits companies from exploiting scientific resources, once they have been found. These insights are in line with our findings from the two conducted studies.

5.5. Stakeholder Map

In Fig. 24 we present our findings along a stakeholder map. To do so, we analyzed if an uptake and also provision of Open Science resources took place and which drivers might foster and which barriers might hinder the uptake. Therefore, we take into account the roles of our stakeholders, thus, our study participants. And we take into account the three target domains, health, climate and agriculture of our investigation.
In the health domain, we conducted interviews with CEOs, CTOs, and heads of department and in the questionnaire study we got feedback from 35 people who are working in this domain. With regard to Open Access, we could show that companies in the health domain serve as consumers and producers; consumers as they are taking up, e.g., literature and publications regarding their domain; producers - as most of them are closely connected with universities - they also often publish scientific work themselves. Major barriers for this group are paywalls and the associated high fees and expensive licenses. Referring to the uptake of Open Source code, they mentioned to use Open Source code for their application development (consumers) and return (parts of their) enhancements to the Open Source community (producers), as the Open Source code development community is based upon mutual give and take. Barriers mentioned consist of licenses that hinder the uptake of Open Source code for commercial purposes as well as compliance and validation issues as no one can ensure or guarantee the correctness of the code. Additionally, our interviewees stated that, in Austria, the uptake of open health-related data is not possible as there is no data available.

In the climate domain, we got input from CEOs, founders and heads of department through the interviews (we leave out the questionnaire results due to the low number of answers received from this domain). Open Access is similarly important for this domain like for the health domain, although they didn't use the term explicitly during the interview. Regarding Open Data, the interviewees reported experience as consumer and producer. Their experiences with the uptake of Open Source code and corresponding barriers are the same as in the health domain.

In the agricultural domain, we received information from a founder and a head of a department (we leave out the questionnaire results due to the low number of answers received from this domain). For the interviewees from this domain, Open Access seems to be a little less important than for the other domains, although publications also play a significant role as information sources and free access to publications is valued by the interviewees. With regard to Open Data, they confirm to take up Open Data from different sources and use them for their work (consumers). Finally, our interviewees do not use Open Source code.

With regard to the drivers, we identified 5 drivers for the uptake of Open Science resources, that will be explained below in more detail, namely i) the employment of people with a university background, ii) offering incentives and support for the uptake, iii) offering training to foster uptake, iv) learning from trans- and interdisciplinary collaboration (projects), and v) exploiting the wisdom of the crowd. All drivers need to be initiated and/or maintained by the companies themselves and can be equally referred to the uptake of Open Data, Open Access and Open Source code.

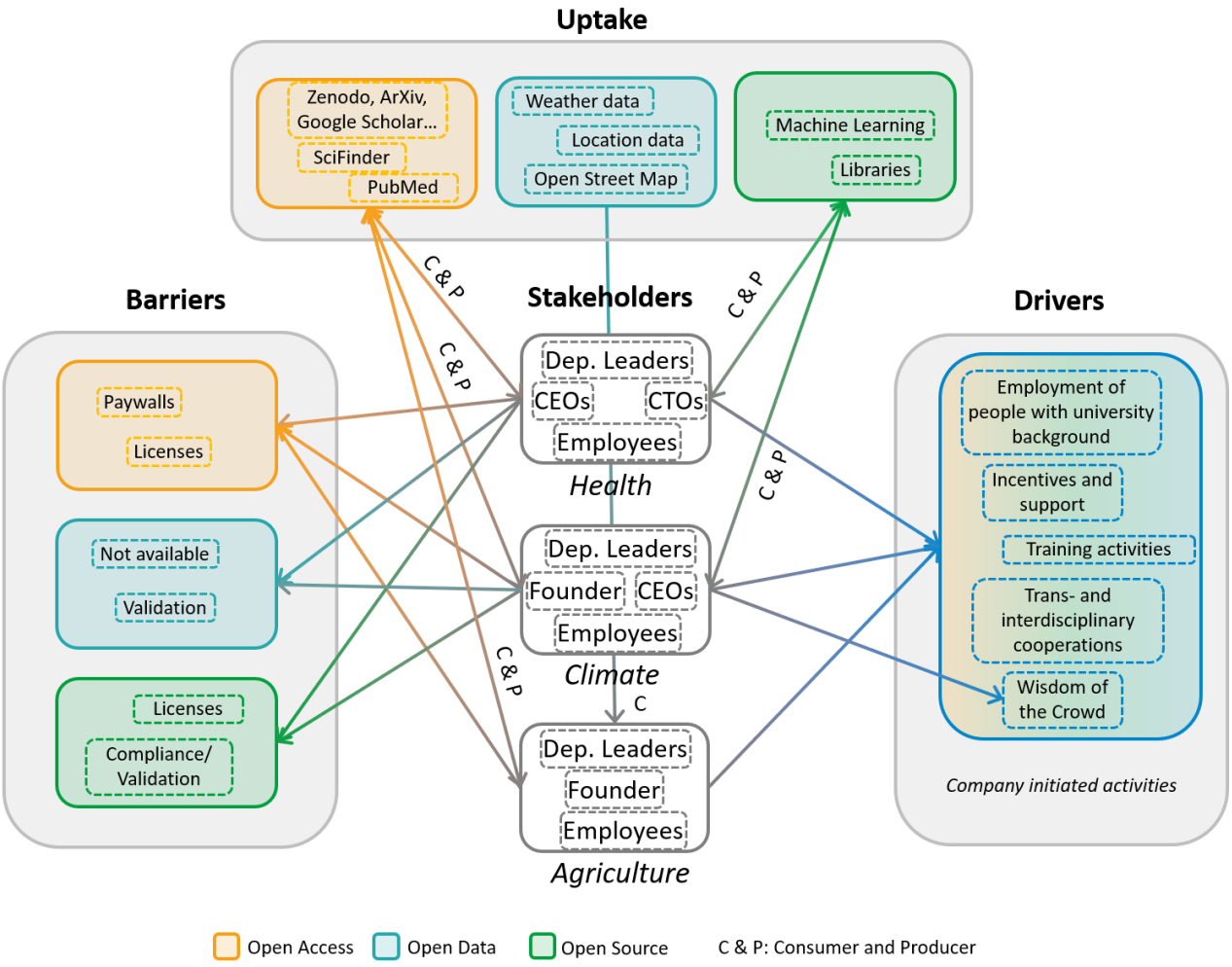


Figure 24: Stakeholder map

6. Conclusion

Within this deliverable report, we present the results of two studies on the drivers and barriers for the uptake of Open Science resources in SMEs and industries in the domains of health, climate and agriculture. We conducted an interview study and one questionnaire study. Our key findings include the following drivers that foster the uptake of Open Science resources:

- Employment of people with a university education: For our two studies and also in line with the literature review conducted, the uptake of Open Science is strongly dependent on the employees' education level. Therefore, to benefit from Open Science resources, companies are currently hiring people with a university degree, thus, with bachelors' degree or higher.
- Incentives and support: From the interviews we could derive that companies that already embrace Open Science activities allow their employees to invest a significant amount of time in Open Science activities like, for example, literature research, updating Open Source code, and in some cases, have developed specific procedures for conducting such activities. From the questionnaire study we found out that such incentives/support, which are given to employees when using Open Science resources or conducting Open Science activities, would be appreciated by the majority of the respondents and thus strengthen the uptake of Open Science in the private sector.
- Training activities: A further driver to increase the uptake of Open Science resources in companies is to offer training in this regard, to close possible gaps between academia and companies, but also between companies and the public. These trainings could show the potential of working with Open Science resources, instructions how to use them and to show them already existing possibilities and opportunities (e.g. Open Data platforms, Open Source communities, Online Social Networks for communicating research results). Additionally, they lower possible fears to try out something new and to strengthen the self-confidence of employees.
- Trans- and interdisciplinary cooperations: Learning from others especially when working together in projects or towards the same goal strongly impacts the uptake of Open Science resources.
- Exploit the wisdom of the crowd: Especially with regard to the uptake of Open Source code, the wisdom of the crowd could help to more quickly develop applications or services by SMEs (if the license allow commercial use): i) when having a bug within an algorithm the crowd is typically very quickly to find solutions, ii) maintenance of Open Source code is give per se, and iii) the code is continuously further developed.

Focusing on Open Data, barriers for the uptake are availability of data (especially in the domain of health), their licences (especially with regard to commercial use) and the reliability and validation of data. Based on these insights we were able to derive five drivers that foster the uptake of Open Science resources in SMEs and industries today: i) the employment of people with a university background, ii) offering incentives and support for the uptake, iii) offering trainings to foster the uptake, iv) learning from trans -and interdisciplinary cooperations and v) exploit the wisdom of the crowd.

Based on the current drivers and barriers found, that are in line with interview and questionnaire results, we suggest the following policy recommendations for action.

- Policy Recommendation 1: Persons with bachelors' degree or higher have at least a bit of a scientific background and they are more aware and familiar with the Open Science endeavours than others. Therefore, the following policy recommendation is suggested:
  "Make the Open Science endeavours, its opportunities and benefits more visible especially outside the university context."
  Our suggestion would be to make the concept of Open Science more popular especially beyond the university or higher educational context, thus, already in further educational settings, secondary schools or, where applicable, even earlier. Encouraging schools to, for instance, participate in citizen science projects could be one concrete approach to do this.
- Policy Recommendation 2: Open Access to state-of-the-art literature is relevant for all companies with R&D departments as well as for SMEs and new start-ups. However, Open Access venues and sources might not be known to all stakeholders in SMES and industries partly due to accessibility restrictions such as high fees for publishing Open Access. While big companies and higher education institutions can (and do) afford to subscribe to a higher number of journals, SMEs and start-ups typically cannot, although access to these information sources is relevant for advancing their services, products and innovation and to build up a stable and reliable business in their domain. Therefore, the following policy recommendation is suggested:
  "Increase the number of Open Access publications available for all interested stakeholders, especially in domains with no strong tradition in Open Access."
  Our suggestion is to first, reduce the fees for publishing Open Access in general and second, establish a tiered pricing system addressing the different stakeholders: fees for universities and big companies, fees for SMEs and start-ups, no fees for low income or lower middle-income countries. Furthermore, it would be of crucial relevance to offer people training, curricula or guidelines about "Open Access literacy" to ensure that people know where to find reliable OA information and how to publish OA in trustable venues.

Summing up, in two studies we could show that there are already a lot of things going on in relation to the uptake of Open Science resources in SMEs and industries, however, this uptake depends on the company, its leaders and employees and their (educational) background. Consequently, there is overall still a need to investigate how such an uptake can be enhanced and improved beyond the domains investigated in this report.

7. References

Axelsson, Karin, Ulf Melin, and Ida Lindgren. 2013. “Public E-Services for Agency Efficiency and Citizen Benefit — Findings from a Stakeholder Centered Analysis.” *Government Information Quarterly* 30 (1): 10–22. https://doi.org/10.1016/j.giq.2012.08.002.

Bergman, Edward M. 2010. “Knowledge Links between European Universities and Firms: A Review.” *Papers in Regional Science* 89 (2): 311–33. https://doi.org/10.1111/j.1435-5957.2010.00310.x.

Brugha, R. 2000. “Stakeholder Analysis: A Review.” *Health Policy and Planning* 15 (3): 239–46. https://doi.org/10.1093/heapol/15.3.239.

Chinyio, Ezekiel, and Paul Olomolaiye. 2009. *Construction Stakeholder Management*. John Wiley & Sons.

Cohen, Wesley M., and Daniel A. Levinthal. 1990. “Absorptive Capacity: A New Perspective on Learning and Innovation.” *Administrative Science Quarterly* 35 (1): 128. https://doi.org/10.2307/2393553.

Cohen, Wesley M., Richard R. Nelson, and John P. Walsh. 2002. “Links and Impacts: The Influence of Public Research on Industrial R&D.” *Management Science* 48 (1): 1–23.

Deloitte, L. L. P. 2012. “Open Growth: Stimulating Demand for Open Data in the UK. A Briefing Note from Deloitte Analytics. London: Deloitte Touche Tohmatsu Limited. Retrieved November 27, 2013." (2012).” London: Deloitte Touche Tohmatsu Limited, London.

ElSabry, ElHassan, and Koichi Sumikura. 2020. “Does Open Access to Academic Research Help Small, Science-Based Companies?” *Journal of Industry-University Collaboration* 2 (3): 95–109. https://doi.org/10.1108/JIUC-04-2020-0004.

Europäische Kommission, ed. 2016. *Open Innovation, Open Science, Open to the World: A Vision for Europe*. Luxembourg: Publications Office of the European Union. https://doi.org/10.2777/552370.

Fawad Sharif, Sayed Muhammad, Yang Naiding, Yan Xu, and Atiq ur Rehman. 2020. “The Effect of Contract Completeness on Knowledge Leakages in Collaborative Construction Projects: A Moderated Mediation Study.” *Journal of Knowledge Management* 24 (9): 2057–78. https://doi.org/10.1108/JKM-04-2020-0322.

Fell, Michael J. 2019. “The Economic Impacts of Open Science: A Rapid Evidence Assessment.” *Publications* 7 (3): 46. https://doi.org/10.3390/publications7030046.

Fessl, Angela, Thomas Klebel, Hannah Metzler, and Tony Ross-Hellauer. 2020. “ON-MERRIT D4.1 Information Seeking Behaviour and Open Science Uptake in Industry: A Literature Review,” May. https://doi.org/10.5281/zenodo.5018067.

Flatten, Tessa C., Andreas Engelen, Shaker A. Zahra, and Malte Brettel. 2011. “A Measure of Absorptive Capacity: Scale Development and Validation.” *European Management Journal* 29 (2): 98–116. https://doi.org/10.1016/j.emj.2010.11.002.

Freeman, R. Edward. 2010. *Strategic Management: A Stakeholder Approach*. Cambridge university press.

Freund, Luanne. 2015. “Contextualizing the Information-Seeking Behavior of Software Engineers.” *Journal of the Association for Information Science and Technology* 66 (8): 1594–1605. https://doi.org/10.1002/asi.23278.

Gonzalez-Zapata, Felipe, and Richard Heeks. 2015. “The Multiple Meanings of Open Government Data: Understanding Different Stakeholders and Their Perspectives.” *Government Information Quarterly* 32 (4): 441–52. https://doi.org/10.1016/j.giq.2015.09.001.

Guo, Bin. 2009. “Moderating Effects of Task Characteristics on Information Source Use: An Individual-Level Analysis of R&D Professionals in New Product Development.” *Journal of Information Science* 35 (5): 527–47. https://doi.org/10.1177/0165551509105196.

Hirsh, S., and J. Dinkelacker. 2003. “Impact of a Corporate Merger on the Information Seeking Behaviors of Research Practitioners.” In *Asist 2003: Proceedings of the 66th Asist Annual Meeting, Vol 40, 2003: Humanizing Information Technology: From Ideas to Bits and Back*, edited by J. B. Bryans, 40:78–84. Medford: Information Today Inc.

———. 2004. “Seeking Information in Order to Produce Information: An Empirical Study at Hewlett Packard Labs.” *Journal of the American Society for Information Science and Technology* 55 (9): 807–17. https://doi.org/10.1002/asi.20024.

Huber, Franz, Thomas Wainwright, and Francesco Rentocchini. 2020. “Open Data for Open Innovation: Managing Absorptive Capacity in SMEs.” *R & D Management* 50 (1): 31–46. https://doi.org/10.1111/radm.12347.

Jones, James W. 2016. “Wayfinding: A Grounded Theory Study of the Information-Seeking Behavior of Constructors.” *Grounded Theory Review* 15 (2): 59–75.

Kassen, Maxat. 2018. “Adopting and Managing Open Data: Stakeholder Perspectives, Challenges and Policy Recommendations.” *Aslib Journal of Information Management* 70 (5): 518–37. https://doi.org/10.1108/AJIM-11-2017-0250.

Kraaijenbrink, Jeroen, and Aard Groen. 2006. “Small Companies Seeking Information on the Internet Any Changes for Online Intermediaries?” *Industry and Higher Education* 20 (3): 195–205. https://doi.org/10.5367/000000006777690990.

Kwasitsu, L. 2003. “Information-Seeking Behavior of Design, Process, and Manufacturing Engineers.” *Library & Information Science Research* 25 (4): 459–76. https://doi.org/10.1016/S0740-8188(03)00054-9.

Laursen, Keld, and Ammon Salter. 2004. “Searching High and Low: What Types of Firms Use Universities as a Source of Innovation?” *Research Policy* 33 (8): 1201–15. https://doi.org/10.1016/j.respol.2004.07.004.

Le, Jette V., Line B. Pedersen, Helle Riisgaard, Jesper Lykkegaard, Jorgen Nexoe, Jeanette Lemmergaard, and Jens Sondergaard. 2016. “Variation in General Practitioners’ Information-Seeking Behaviour - a Cross-Sectional Study on the Influence of Gender, Age and Practice Form.” *Scandinavian Journal of Primary Health Care* 34 (4): 327–35. https://doi.org/10.1080/02813432.2016.1249057.

Lelea, Margareta A., Guyo Malicha Roba, Anja Christinck, and Brigitte Kaufmann. 2014. *Methodologies for Stakeholder Analysis: For Application in Transdisciplinary Research Projects Focusing on Actors in Food Supply Chains: Reload Reducing Losses Adding Value*. Witzenhausen: DITSL.

Li, Yuelin, Ying Li, Ying Pan, and Hongliang Han. 2019. “Work-Task Types, Stages, and Information-Seeking Behavior of Strategic Planners.” *Journal of Documentation* 75 (1): 2–23. https://doi.org/10.1108/JD-01-2018-0015.

Lundin, Jonatan, and Yvonne Eriksson. 2018. “An Investigation of Maintenance Technicians’ Information-Seeking Behavior in a Repair Center.” *Ieee Transactions on Professional Communication* 61 (3): 257–74. https://doi.org/10.1109/TPC.2018.2826087.

Merton, Robert K. 1968. “The Matthew Effect in Science: The Reward and Communication Systems of Science Are Considered.” *Science* 159 (3810): 56–63. https://doi.org/10.1126/science.159.3810.56.

Perkmann, Markus, Rossella Salandra, Valentina Tartari, Maureen McKelvey, and Alan Hughes. 2021. “Academic Engagement: A Review of the Literature 2011-2019.” *Research Policy* 50 (1): 104114. https://doi.org/10.1016/j.respol.2020.104114.

Perkmann, Markus, Valentina Tartari, Maureen McKelvey, Erkko Autio, Anders Broström, Pablo D’Este, Riccardo Fini, et al. 2013. “Academic Engagement and Commercialisation: A Review of the Literature on University–Industry Relations.” *Research Policy* 42 (2): 423–42. https://doi.org/10.1016/j.respol.2012.09.007.

Savage, Neil. 2016. “Competition: Unlikely Partnerships.” *Nature* 533 (7602): S56–58. https://doi.org/10.1038/533S56a.

Simeth, Markus, and Julio D. Raffo. 2013. “What Makes Companies Pursue an Open Science Strategy?” *Research Policy* 42 (9): 1531–43. https://doi.org/10.1016/j.respol.2013.05.007.

Slabá, Marie. 2016. “Stakeholder Profile and Stakeholder Mapping of SMEs.” *Littera Scripta* 9 (1): 16.

Starasts, Ann. 2015. “‘Unearthing Farmers’ Information Seeking Contexts and Challenges in Digital, Local and Industry Environments’.” *Library & Information Science Research* 37 (2): 156–63. https://doi.org/10.1016/j.lisr.2015.02.004.

Su, Chunke, and Noshir Contractor. 2011. “A Multidimensional Network Approach to Studying Team Members’ Information Seeking From Human and Digital Knowledge Sources in Consulting Firms.” *Journal of the American Society for Information Science and Technology* 62 (7): 1257–75. https://doi.org/10.1002/asi.21526.

Varvasovszky, Z. 2000. “A Stakeholder Analysis.” *Health Policy and Planning* 15 (3): 338–45.

https://doi.org/10.1093/heapol/15.3.338.
Veugelers, R., and B. Cassiman. 2005. “R&D Cooperation between Firms and Universities. Some Empirical Evidence from Belgian Manufacturing.” *International Journal of Industrial Organization* 23 (5–6): 355–79. https://doi.org/10.1016/j.ijindorg.2005.01.008.
Yitzhaki, M., and G. Hammershlag. 2004. “Accessibility and Use of Information Sources among Computer Scientists and Software Engineers in Israel: Academy versus Industry.” *Journal of the American Society for Information Science and Technology* 55 (9): 832–42. https://doi.org/10.1002/asi.20026.
Zahra, Shaker A., and Gerard George. 2002. “Absorptive Capacity: A Review, Reconceptualization, and Extension.” *The Academy of Management Review* 27 (2): 185. https://doi.org/10.2307/4134351.

8. Annex

8.1. Interview Guideline

Duration: 30 - 60 minutes

1. Introduction

On-MERRIT (https://on-merrit.eu) is an EU project, that investigates the impact, influence and uptake of Open Science not only in science, or politics but also in relation to companies and here especially SME's in the domains of climate, health and agriculture. Open Science means opening research results and scientific output in the era of digitization. Open Access means unlimited and free access to scientific information in the web. And open (research) data are data that have been created in research and are now freely available in the web.

With this interview study we would like to find out if and how Open Science / Open Access / Open Data is already taken up in SME's. Furthermore, we also would like to investigate which challenges and barriers exist that hinder or restrict the uptake of Open Data.

The results of this interview will be on the one hand made available to partners of the ON-MERRIT project and we also plan a publication. However, before we process the gained insights, we ensure that all data and information that you share with us, will be anonymized, so that no one will ever be able to draw conclusions to refer back to you or your company.

Consent for recording: All information you will give me during the interview will safely be stored (encrypted) and the results/answer given will only be used in an anonymous way (for a publication) and will not be given to any third person.

So, is it ok if I record this interview for later analysis w.r.t. research conducted?

2. Working Context

At the beginning of the interview, I am interested to learn more about your company, what your company does, which services and products does your company offer. And second, I would like to know what your role, your responsibility is and what tasks you do in the company. So, could you please tell me a bit more about you and your company?

- Describe shortly the purpose of your company
- Describe shortly your role in the company.
- Describe shortly your overall working activities/working tasks/ a typical working day.

3. Data/Information usage at work

- Which (type of) data/information do you work with in your job?

- What is the data/information you work with used for?
- Where do you get the data/information from?
    - Own or company data
    - Open Data/Science
    - ○ Others
- Which challenges/barriers do you have faced with the data you are working/using/applying with in general?

4. Information Seeking Behaviour

- How do you keep yourself up-to-date with research?
- How do you keep your company up-to-date with research?
- Which information sources/platforms/etc. do you / your company use?
- When you are looking for new data, research, etc.
    - How do you search?
    - Where do you search for which data/information? Which platforms, search engines…do you use?
    - How often do you search? Every day, once a week, …
- When you think of your information seeking behaviour:
    - Which challenges/barriers have you experienced with seeking for information? For example: Finding the right information? Challenges conducting a search e.g. defining the right keywords? (Lack of skills capacity in search/search expertise?) Challenges to identify and interpret the information/data found?

5. Open Science/Access/Data

- What do you know about Open Science, Open Access, and Open (research) Data? Have you heard about the terms?
- [Show SLIDE and give explanation]

Explanation: There are different definitions of Open Science including the corresponding topics like "Open Access", "Open (Research) Data" or "Open Source" that are often mentioned in combination.

From (scientific) literature there is no unique definition of Open Science. Part of these definitions see ***Open Science*** *as "Making knowledge freely available for everyone"; "Opening up the process of knowledge creation"; and "Making science accessible for citizens" – this is also how we use it in ON-MERRIT.*

From, for example, the Open Science Network Austria (OANA - oana.at), ***Open Science*** *describes the* ***opening of research processes and scientific output*** *in times of digitization. The term* ***Open Access*** *(OA) means* ***unrestricted and free access to scientific information*** *on the internet.* ***Open Research Data*** *are* ***data created in the course of scientific research*** *(e.g. through digitisation, source studies, experiments, measurements, investigations or surveys) and* ***provided openly****, i.e. made available worldwide on the internet.*

- **TODAY:** If you think about your company today:
- Have you ever used Open Access, open (research) data, …?
    - And if yes,

        - Which kind of Open Access/data? What for?
        - Which data sources/platforms do you use?
    - If not,
        - Why not? What are the main barriers to use Open Data/science?
        - How could they be overcome?
- When you are thinking of Open Data/science/research…
    - Which challenges/barriers do you have faced with this topic? For example: Lack of clarity where to find relevant data? Lack of clarity around where benefits accrue? Lack of clarity which sources/platforms to trust?

6. Absorptive Capacity

*Absorptive capacity has been defined as "a firm's ability to **recognize the value** of new information, **assimilate it**, and **apply** it to commercial ends"*

- **FUTURE:** In the future …
    - … can you imagine to use Open Science, data, access?
    - … which data could that be and what could it be used for?
    - … how could this data bring you/your company one step further?

7. Business Model

- TODAY: If you think about your company today…
- FUTURE: If you think about your company in the future…
    - … which type of business models 'archetypes' for exploiting open data would fit best to your company, and why:
- [Show SLIDE and give explanation]
        - **Suppliers**: organisations that publish their data via an open interface to allow others to use and reuse it.
        - **Aggregators**: organisations that collect and aggregate Open Data and, sometimes, other proprietary data, typically on a particular sectorial theme, find correlations, identify efficiencies or visualise complex relationships.
        - **Developers**: organisations and software entrepreneurs that design, build and sell web-based, tablet or smartphone applications for individual consumption.
        - **Enrichers**: organisations (typically larger, established businesses) that use Open Data to enhance their existing products and services through better insight.
        - **Enablers**: organisations that facilitate the supply or use of Open Data, but are not themselves users or re-users of Open Data.
- To what extent are you concerned that critical knowledge could flow out of your company through new business models or open science?
- How are you currently managing this risk or what would help you with it?

8. Demographic Questions

- Gender:
- Profession/Employment:

- Position:
- Years:
- Full-time/Part-time:
- Domain of your Company:
- Number of Employees

9. End of the interview

- Is there anything you would like to add to the interview?
- Do you have any final comments for us?

Thank you for participating.

8.2. Questionnaire

1. Introduction

With this questionnaire we are investigatíng if businesses take up freely available scientific resources (Open Science resources) such as data, information or publications and to what extent these resources are integrated into the working environment. Additionally, we hope to identify any barriers and challenges that might prevent this from happening. The main focus of this study is on small, medium and industry enterprises situated in the domains of agriculture, climate and health.
We invite employees of all levels to participate in this study. We would be grateful if you could also share this questionnaire with your professional network.
The study is conducted as part of the ON-MERRIT project . To learn more about ON-MERRIT, visit our website: https://on-merrit.eu/.

This project is funded by European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement number 824612.

Completing this questionnaire will take approximately 20 minutes.

2. Participants Information Sheet and Consent Form

Introduction
Please read the following information before deciding whether to participate in our study. Please ask the researcher any questions you may have about the process or project before completing the consent form (see contact details below).

Your participation in this study
You are invited to participate in this study on a voluntary basis and you are free to withdraw from the study at any time without providing any reason for doing so, even after completing the survey. If you agree to participate, you give us permission to:

- Collect information from you;
- Share information with project team members (no personal information will be shared);

- Conduct the study;
- Use this information for data analysis, in publications and presentations.

Confidentiality and publication of the study data
This survey does not ask you to provide any personally identifying information. Still, all the data you provide will be anonymised and treated confidentially and will be used for research purposes. It will not be used in a manner which would allow identification of your individual responses. This applies to all outputs that might stem from the project, including academic papers and other reports, conference presentations and published datasets. The raw data will be stored in the internal servers of Know-Center GmbH and will be protected by passwords that are only known to researchers conducting this study. All the raw data will be deleted five years after the completion of the project.
For further information or to withdraw from the study, contact DI Dr. Angela Fessl, Know-Center GmbH, afessl@know-center.at.

Consent Form [YES/NO]

- I confirm that I have read the participant information.
- I understand that my participation in this study is voluntary and that I am free to withdraw at any time. I understand that it will not be possible to remove my data from the project once it has been anonymised and forms part of the data set. I agree to take part on this basis.
- I agree that any data collected may be published in anonymous form in books, reports or journals and shared in presentations.
- I agree to take part in this study.

3. Company/Working Context

In this section we want to learn about the company for which you work..

3.1 In which domain is your company situated? [CHECKBOX]

- Health
- Climate
- Agriculture
- Other [OPEN TEXT]

3.2 What is the name of your company?
Hint: This is a voluntary question. If you answer this question, we would be able to get more information about your company and the products and services you are offering. This helps us to get deeper insights about the uptake of Open Science resources in the industry. However, this data will not be used to uncover your identity nor will it be published.

3.3 Please enter the company url:
Hint: This is a voluntary question. If you answer this question, we would be able to get more information about your company and the products and services you are offering. This helps us to get deeper insights about the uptake of Open Science resources in the industry. However, this data will not be used to uncover your identity nor will it be published.

3.4 In which city and country is your primary company office located (where you are employed) e.g. Graz, Austria or Berlin, Germany? (If you are working in your home office, please still refer to the company's office)

3.5 What is the primary language spoken in your workplace? [TEXT]

3.6 How many employees work in this company?

- 0-50
- 51-100
- 101-150
- 151-200
- 201-250
- 250-500
- 500+

3.7 With whom are you in business? [CHECKBOX]

- Business to Business
- Business to Customer
- Business to Government
- Other []

4. Demographics

In this section we want to learn about your background.

4.1 Gender: Man, Woman, Other, Prefer not to say [RADIOBUTTONS]

4.2 Age: [NUMBER FIELD]

4.3 Please select your highest education level: [RADIOBUTTONS]

- Primary education
- Lower secondary education
- Upper secondary education
- Post-secondary non-tertiary education (e.g. VET Schools, schools of healthcare and nursing)
- Short-cycle tertiary education (e.g. master schools, colleges, vocational training schools)
- Bachelor or equivalent
- Master or equivalent
- Doctorate or equivalent
- Other [FreeText]

4.4 What is your position in the company (e.g. CEO, CTO, project manager, …)? [FREE TEXT]

4.5 For how many years have you worked in this company? [NUMBER FIELD]

4.6 For how many years have you worked in this industry/domain? [NUMBER FIELD]

5. Information and Search Behaviour

In this section, we want to learn about the tactics you use to find relevant information or data in the course of your work.

5.1 What type of information do you generally require for your work in your company? Please check all that apply. [CHECKBOX]

- Business/economic information;
- Legal information;
- Company information;
- Country information;
- Information on health
- Information on climate
- Information on agriculture
- Information about projects;
- Specific sector/industry information;
- Market- or consumer-related data (e.g. trends, target groups)
- Trade data
- Other [FREETEXT]

5.2 How often do you go to the following sources for information (Levels: always, often, sometimes, rarely, never;)?

- Company internal material and resources
- Professional networks (supervisors and co-workers, etc);
- Government publications or reports;
- Publications from companies in the same domain;
- Internet (search engines, web sites, etc),
- social media and software (Facebook, Twitter, LinkedIn, Blog, Wiki etc.);
- Printed newspapers and industry magazines;
- Analysts' research reports
- Library/information center (outside the company);
- Chamber of commerce & industry information sources;
- Patent databases
- Other [FREETEXT]

5.3 On average, how many hours per week do you spend searching for information for your work in your company?

- 0-2 hours;
- 3-4 hours;
- 5-6 hours;
- 7-8 hours;
- 9-10 hours;
- More than 10 hours

5.4 How frequently do you encounter the following difficulties, challenges, and barriers when searching for relevant information? (Levels: always, often, sometimes, rarely, never; or something similar.)

- Lack of time
- Lack of information searching skills (e.g. defining or formulating keywords, not knowing where to search)
- Limited access to information sources (e.g. subscription/payment is needed for some sources)
- Low quality of the information found
- Language barrier
- Other [FREE TEXT]

5. Understanding and Uptake of Open Science

Open Science (OS) aims at removing barriers to access and re-use of research outputs. It is designed to make the scientific process and the knowledge more transparent, reproducible, accessible, shared and collaborative. As an umbrella of activities it includes Open Access, Open Data and Open Source, among others. In this section, we want to learn about your awareness of and attitudes towards Open Science, as well as those within your company, and if and how you or your company uses Open Science practices, and about any barriers or challenges that may prevent you or your company from doing so.

6.1 Prior to beginning this questionnaire, were you already familiar with the concept of Open Science? [YES/NO]
6.2. If you answered yes, do you already embrace and use OS practices? [YES/NO]
6.3 [Condition: IF 6.3 No -> 6.4 does not appear]
Does your company conduct or contribute to research (e.g. collecting and processing of data and information)? [RADIOBUTTONS]

- Yes, we collect, analyse and disseminate data and information for research purposes.
- Yes, we collect and (at least partly) analyse data and information and we transfer our results to other institutions for dissemination.
- Yes we collect, generate, and pre-process data and information and transfer them to other institutions for aggregation and dissemination.
- Yes, we aggregate and disseminate data and information, but we do not collect or generate them ourselves.
- Yes, we collect and analyse data and information but only for internal purposes.
- Yes, we analyse data and information for our customers, but only for commercial purposes.
- No, we do not contribute to research, however we observe current scientific development in our domain.
- No, research is not directly relevant to our work.
- Yes, we contribute to research in a different way: _____

6.4 How frequently does your company participate in any Open Science activity/action? [RADIOBUTTONS] (Levels: always, often, sometimes, rarely, never; I don't know/ I don't have enough information)?

- Collaborations across institutions and disciplines (interdisciplinary groups, projects or meetings, collaborative initiatives, etc.)
- Dissemination to the public and outreach (social networks, articles or talks to the lay public, relationship with the media, etc.)
- Dissemination to scientists (conferences and seminars, courses, articles, etc.)
- Ethical aspects of science and research integrity (participation in ethics committees, bioethical research, training, awareness activities, etc.)
- Gender equality (gender or sex is taken into account in your research, promotion of women's visibility in science, training, mentoring, etc.)
- Production of Open Access publications (Open Access journals, economic support to publish in Open Access, open peer review, etc.)
- Offering Open Data (use of public data infrastructure to deposit and/or access data, participation in Open Data management, training, etc.)
- Participation of the public and/or different stakeholders in your research (dialogues with the public, science cafes, citizen science initiatives, patient associations, etc.)

- Other [FREETEXT]

6.5 Does your company re-use data from one or more domains of external sources and if yes, please select the corresponding domains [CHECKBOX]:

- Agriculture, forestry and fishing
- Health
- Climate
- Energy
- Regions and cities
- Transport
- Economy & Finance
- International Issues
- Government & Public Sector
- Justice, Legal System & Public Safety
- Environment
- Education, Culture & Sport
- Population & Society
- Science & Technology
- No, we do not use any external data
- Other [FREETEXT].

6.6 Does your company provide you with adequate training related to Open Science? [RADIOBUTTONS] (Levels: Yes, I receive adequate training; No, I need more training; These topics are not relevant for my specific professional tasks; I don't know / I don't have enough information)

- Research and data management (data storage, sharing, FAIR - "Findable, Accessible, Interoperable, and Reusable" - approaches)
- Research integrity (animal research, data analysis and interpretation, research with human samples/subjects, good practice in the lab, etc.)
- Research publishing and dissemination (Open Access, pre-prints, peer review)
- Collaborating and networking (how to improve collaboration through Open Science)
- Communicating science to the general public (different audiences, practical guides to getting started, online and offline options)
- Involving the general public in research (citizen science: data gathering, data analysis, use of results)
- Evaluation of research projects and researchers
- Assessment of public impact

6.7 To what degree does your company offer you support or incentives related to Open Science? [RADIOBUTTONS]

(Level: I receive adequate support or incentives; I would like to receive more support or incentives; I do not receive any support or incentives but would like to; I do not receive any support or incentives; ; I don't know/I don't have enough information)

- Guidelines, policies, recommendations for how to use Open Science resources
- Technical infrastructure (templates, software, storage, databases, publication and/or data repositories, etc.)
- Specialist support (experts on different aspects of Open Science, research data committees, courses, workshops, etc.)
- Financial support and rewards
- Careers perspectives and recognition

6.8 [BARRIER]: Imagine in your everyday work that you decide to embrace (or you already have embraced) an Open Science perspective. What do you think (or know) are the most significant barriers you will be facing? [RADIOBUTTONS]
(Levels: Very significant barrier, significant barrier, minor barrier, no barrier, I don't know / I don't have enough information.)

- Lack of proper infrastructure. How/where do I store Open Data?
- Lack of clear steps to follow. How do I begin? How do I proceed?
- Lack of clarity where to find relevant data
- Lack of clarity around where benefits accrue
- Lack of clarity in which sources/platforms to trust
- Time constraints. I don't have time to practise Open Science; it is too time-consuming.
- Fears and uncertainties for career development. Will my Open Science practice be valued at the institutional level or during my career? Does it mean I will receive more funding or merit?
- Other [FREETEXT]

6.9 Which kinds of Open Science resources are you interested in working with for your company? [CHECKBOX]
(Levels: As consumer (1), As producer (2), Not interested (0))

- Open Access (e.g. publications)
- Open Data
- Reproducible Research
- Open Source Software
- Open licenses
- Ethics in research/ Research integrity
- None
- Other: [FREETEXT]

6.10 [CONDITION to 6.8 "Open Data"] Which kinds of Open Data are you interested in working with for your company? [CHECKBOX]

- Statistical data (e.g. performance statistics, climate statistics)
- Location data (e.g. location of weather services)
- Service data (e.g. details of services provided)
- Real-time data (e.g. weather date, rain data, …)
- Economic data (e.g. stock market trends, market industry data, …)
- Consumer data (e.g. trends, …)
- I'm not sure yet
- None
- Other: [FREE TEXT]

6.11 What are the management's motives for working with or using Open Data for your company? [CHECKBOX]
(Levels: Very important, Important, Neutral, Unimportant, I don't know)

- Curiosity
- Building an innovative site/service
- Making a profit
- Meeting the request of a client/manager
- Making a difference to my local community
- Organisational development

- Developing the semantic web/linked data web
- Being recognised as the creator of something useful or insightful
- Providing a better service to citizens or customers
- Providing a platform for other people to build upon
- Solving a specific problem

6.12 Overall, if you had to summarise your view on Open Science with regard to your company, what would you say? [RADIOBUTTONS]

- Open Science is an exciting opportunity for my company, mostly with benefits
- Open Science is an opportunity for my company, with the benefits outweighing the drawbacks
- Open Science is mostly positive for my company, it has benefits but also important drawbacks
- Open Science is an unimportant bureaucratic burden for my company
- Open Science is a worrying new perspective for my company
- Open Science is a real threat to my company
- Open Science is not relevant for my company.

7. Indicators for the uptake and integration of new information

A company's ability to recognize the value of new information, assimilate it, and apply it to commercial ends is typically relevant for a company to stay competitive. In this section we are interested in learning how you and your company deal with the uptake and integration of new information.
(ACAP SCALE – 5 – point Likert Scale: strongly agree, agree, neutral, disagree, strongly disagree)

Acquisition: Please specify the extent to which your company uses external resources to obtain information (e.g., personal networks, consultants, seminars, internet, database, professional journals, academic publications, market research, regulations, and laws concerning environment/technique/health/security):

7.1 The search for relevant information concerning our industry is every-day business.
7.2 Management motivates employees to use information sources within our industry.
7.3 Management expects that employees deal with information beyond our industry.

Assimilation: Please rate the extent to which the following statements fit the communication structure in your company:

7.4 Ideas and concepts are communicated across departments.
7.5 Management emphasizes cross-departmental support to solve problems.
7.6 There is a quick information flow, e.g., if a business unit obtains important information it communicates this information promptly to all other business units or departments.
7.7 Management demands periodic cross-departmental meetings to interchange new developments, problems, and achievements.

Transformation: Please specify the extent to which the following statements align with knowledge processing in your company:

7.8 We as employees have the ability to structure and to use collected knowledge.
7.9 We as employees are used to absorb new knowledge as well as to prepare it for further purposes and to make it available.
7.10 We as employees successfully link existing knowledge with new insights.

7.11 We as employees are able to apply new knowledge in our practical work.

Exploitation: Please specify to what extent the following statements fit the commercial exploitation of new knowledge in your company (NB: Please think about all company divisions such as research & development, production, marketing, and accounting):

7.12 Management supports the development of prototypes.

7.13 My company regularly reconsiders technologies and adapts them according to new knowledge.

7.14 My company has the ability to work more effectively by adopting new technologies.

8. Business Model Archetypes

The focus of this section is on Open Data. There exist several business model archetypes which represent business activities in relation to data. Please think about your company and which of these archetypes fit today and which of them might fit in the future. The five archetypes are:

- *Suppliers: Organisations that publish their data via an open interface to allow others to use and reuse it.*
- *Aggregators: Organisations that collect and aggregate Open Data and, sometimes, other proprietary data, typically on a particular sectorial theme; find correlations, identify efficiencies or visualise complex relationships.*
- *Developers: Organisations and software entrepreneurs that design, build and sell web-based, tablet or smartphone applications for individual consumption.*
- *Enrichers: Organisations (typically larger, established businesses) that use Open Data to enhance their existing products and services through better insight.*
- *Enablers: Organisations that facilitate the supply or use of Open Data, but are not themselves users or re-users of Open Data.*

8.1 Which of these archetypes fit your company today? [CHECKBOX]

- Suppliers
- Aggregators
- Developers
- Enrichers
- Enablers
- None

8.2 Please explain why/how your selection(s) above best fits your company today. [FREE TEXT]

8.3 Which of these archetypes might fit your company in the future? [CHECKBOX]

- Suppliers
- Aggregators
- Developers
- Enrichers
- Enablers
- None

8.4 Please explain why/how your selection(s) above best fits your company in the future. [FREE TEXT]

9. Knowledge Risks

The following questions refer to risks related to the loss of proprietary knowledge in Open Science collaborations. Please think about your organization and consider which of these statements you agree or disagree with in terms of knowledge leakage, contract completeness and trust.

9.1 Knowledge leakage Scale from … [RADIOBUTTONS]
(Levels: strongly agree, agree, neutral, disagree, strongly disagree)

- Our proprietary knowledge and core technologies have risks of being appropriated by Open Science collaborators.
- Our private knowledge could leak out since we do not pay much attention to protecting it in the Open Science platform.
- Our private knowledge is imitated by Open Science collaborators behind our back.
- Open Science collaborators usually encroach on our commercial secrets deliberately.
- Open Science collaborators have illegally internalized our private knowledge and technologies.
- Our core knowledge has sometimes been unconsciously transferred to partners through daily communication and interaction.
- Our core knowledge has sometimes been accidentally transferred to the partner due to unwanted facilities in the Open Science platform(such as unrestrained collaborative environments created or easy access)
- Our core knowledge has sometimes been involuntarily transferred to the partner by offering detailed information.

9.2 Contract completeness Scale [RADIOBUTTONS]
(Levels: strongly agree, agree, neutral, disagree, strongly disagree)

- Roles and responsibilities in the Open Science platform are clearly defined by the agreement.
- Schedules and milestones are detailed in the agreement.
- The agreement lists safeguards (such as confidentiality).
- Intellectual property rights are detailed in the agreement.

9.3 Trust Scale [RADIOBUTTONS]
(Levels: strongly agree, agree, neutral, disagree, strongly disagree)

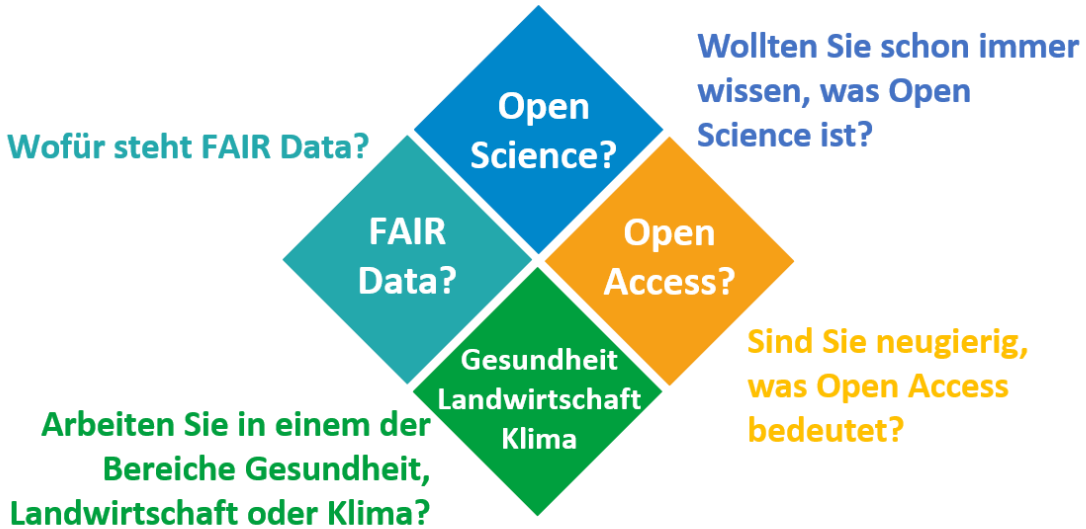
- We can rely on our partner to abide by the Open Science agreement.
- There is a high level of trust in the working relationship with the Open Science agreement.
- We trust that our partner's decisions will be beneficial to the Open Science platform.
- We trust that our partner's decisions will be beneficial to our firm.

8.3. Flyers for Promoting the Questionnaire

KNOW Center

on merrit
www.on-merrit.eu

UMFRAGE ZUR VERWENDUNG VON FREI VERFÜGBAREN WISSENSCHAFTLICHEN RESSOURCEN IN KMUS UND INDUSTRIEBETRIEBEN



Dann sind wir auf der Suche nach Ihnen!

Mit dieser Umfrage möchten wir herausfinden, ob und inwieweit Klein-, Mittel- und Industriebetriebe frei verfügbare wissenschaftliche Ressourcen (= Open Science Resources) wie beispielsweise Daten, Informationen und Veröffentlichungen verwenden oder nicht verwenden.

Wir laden daher MitarbeiterInnen aller Ebenen aus KMUs und Industrieunternehmen aus den Bereichen Gesundheit, Klima und Landwirtschaft ein, an der Umfrage teilzunehmen. Deshalb bitten wir Sie, sich 20 Minuten Zeit zu nehmen, denn jeder ausgefüllte Fragebogen zählt!

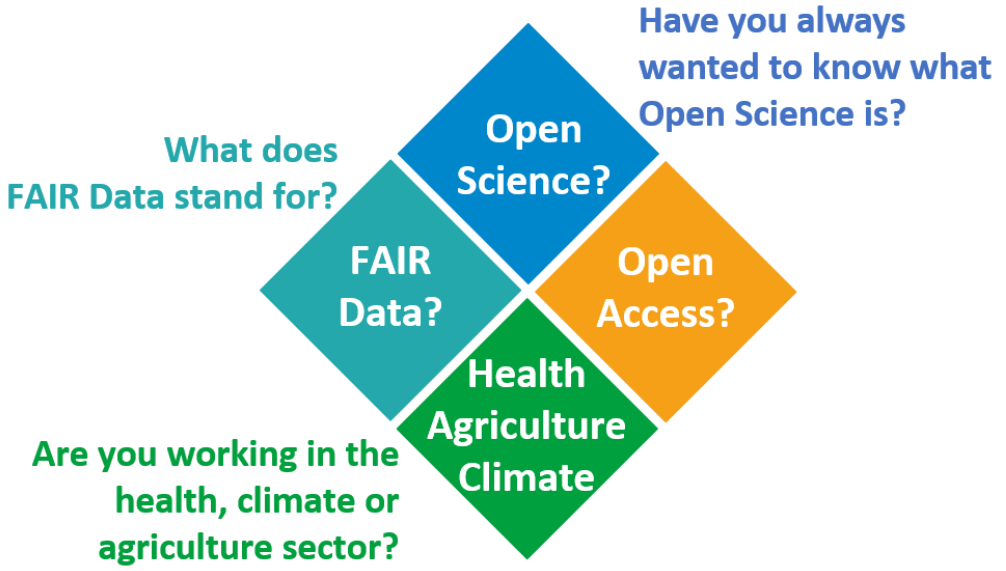
Des Weiteren würden wir uns freuen, wenn Sie die Umfrage in Ihren Netzwerken und Sozialen Medien verbreiten.

You find the survey at: https://bit.ly/3sfgI0C

Für Fragen kontaktieren Sie bitte DI Dr. Angela Fessl (afessl[at]know-center.at)

The research leading to these results has received funding from the European Union's Horizon 2020 Research and Innovation Programme, under Grant Agreement no 824612.

SURVEY ON THE USAGE OF FREELY AVAILABLE SCIENTIFIC RESOURCES IN SMES AND INDUSTRY.



Then we are looking for you!

With this survey we would like to find out whether and to what extent small and medium-sized enterprises and industry use or do not use freely available scientific resources (= Open Science Resources) such as data, information and publications.

We invite employees of all levels working in SMEs and industry in the domains of health, climate and agriculture to take part in the survey.

We kindly ask you to spend 20 minutes - each completed survey counts!

We would also be happy if you spread this survey in your networks.

You find the survey at: https://bit.ly/3ibqLPy

For questions, please contact: DI Dr. Angela Fessl (afessl[at]know-center.at)

The research leading to these results has received funding from the European Union's Horizon 2020 Research and Innovation Programme, under Grant Agreement no 824612.