Enabling Data-Driven Mobility Research: Design Principles and Design Features for an open Platform approach

Daniel Sivizaca Conde
Freie Universität Berlin
daniel.conde@fu-berlin.de

David Rößler
Freie Universität Berlin
david.roessler@fu-berlin.de

Natalia Kliewer
Freie Universität Berlin
natalia.kliewer@fu-berlin.de

Lars Stegemann
Freie Universität Berlin
lars.stegemann@fu-berlin.de

Abstract

In the field of mobility research, up-to-date data is needed in order to explore current problems and challenges. Proponents of Open Science argue that research data should be Findable, Accessible, Interoperable, and Reusable (FAIR). Adopting the Design Science Research methodology, we derive design requirements, design principles and design features for a suitable platform to realize the FAIR principles. This is based on a literature review, existing solutions, and interviews with mobility researchers (with qualitative and quantitative research backgrounds). In a practical way, and after an evaluation of our prototype, we show how a platform should be designed to allow users, regardless of their level of experience, to participate. We contribute to the discussion in the field of Open Science on how artifacts should be designed according to an open design, which also takes into account the different needs of researchers.

Keywords: Open Science, Open Data Platform, Mobility Platform, Design Science Research (DSR)

1. Introduction

"A spirit of openness is gaining traction in the science community". There is an increasing trend toward open conduct, publication, and communication of research (Gewin, 2016). Such approaches are commonly referred to as open science (OS), meaning the process of creating transparency and accessibility of scientific knowledge and methodologies to others. To accomplish this, open practices such as "open access", "open data" or "open source" are encouraged by OS (Munafò et al., 2017).

Researchers are nevertheless facing challenges like a lack or the inaccessibility of research data. Various research areas are confronted with these challenges that are also found in mobility and transportation research (Audirac et al., 2022; Chen et al., 2021; GarcíaPeñalvo et al., 2010; Munafò et al., 2017). Researchers often use primary data in their own research and also generate secondary data themselves for the mobility context. In the process, certain facets of the data are often missing, such as passenger flows in research on delays in transportation networks (Kliewer and Suhl, 2011). Such missing data aspects can sometimes be obtained through other researchers' data generation efforts. Experience shows that the acquisition of mobility and complementary data can be both tedious and lengthy. By mobility data, we mean spatio-temporal data. Especially in interdisciplinary mobility research, spatio-temporal data often need to be related to complementary data from other domains. Such data can be, e.g., weather data or infection data. Even when this data is available, some scientists lack the technical skills needed to analyze the data, process it and put it into context with complementary data (Brandão, 2015). For these researchers, it could be attractive to base initial findings on available data. It would be useful to have a platform solution which provides mobility data, complementary data, and findings from this data for various research needs. This could avoid data acquisition efforts for future research projects and make research findings accessible to all. The basic idea of the data science platform Kaggle, was an inspiration for this project. On the Kaggle platform, companies and institutions voluntarily provide data sets. The data providers also provide questions that the platform participants can answer by analyzing these data sets. Other platform participants can be, for example, researchers who are interested in the data sets and also able to answer the questions raised. A first literature search has shown that there are solutions

URI: https://hdl.handle.net/10125/102696
978-0-9981331-6-4

where mobility data is provided openly and centrally, but they rarely provide complementary data or individual analysis. To address this challenge, we propose a design science research (DSR) project following the guidelines by Peffers et al. (2007), Sonnenberg (2012) and Meth et al. (2015). This research aims to develop a digital platform that provides and interconnects mobility data, complementary data, and supports researchers with methodological problems to process this data and to generate first findings. In the long run, interactions such as data provision and analysis execution should be performed by researchers. The idea of interaction should be emphasized. A platform therefore is considered the appropriate implementation. Our research question raised is, ***RQ:*** *how to design an open science platform that provides open mobility and complementary data and research findings?* To answer this question, we follow a design science research approach. We started by conducting interviews with mobility researchers to identify the specific problem. We then validated the problems by drawing on the literature and by searching for existing solutions. Based on the validated problem statements, we derive initial design requirements (DRs) in an argumentative way. From the literature, we can add more DRs. For the conceptualization we derive nine design principles (DPs). For the subsequent construction of the artifact, design features (DFs) were derived from the DPs. In conclusion, the final evaluation of the artifact is intended to validate the DFs.

The remainder of this paper is organized as follows: In the following chapter, we present the theoretical background to OS and open data (OD) subject and present the existing solutions. In chapter 3 we explain our DSR approach and our corresponding evaluation activities. The following chapters are based on Peffers et al. (2007) DSR phases. In chapter 4, we present the results for the problem and objective identification phase. This leads to the description of the design and development phase in chapter 5. The evaluation of the artifact is described in chapter 6. The paper concludes with a discussion of the results, and a short summary.

2. Related work

2.1. Open Science & Open Science Data Platforms

In order to improve the problems described, we follow the discussion of Open Science in the literature (e.g. Doyle and Luczak-Roesch, 2020) and the call for artefacts in this area (Doyle and Luczak-Roesch, 2019). Our focus is especially on open accessibility and on sharing of mobility data and the analysis of research. Following Open Science practices (e.g., Doyle et al., 2021, Doyle and Luczak-Roesch, 2020, Doyle et al., 2019), by making the data, the analysis and the findings of the research available. This increases transparency and traceability, e.g., through open reviewing of the entire research work (Hossain et al., 2016). With regard to the design of an open IT artifact (Doyle et al., 2019) for the exchange of mobility data, the FAIR principle provides a suitable guideline for the design of the artifact (e.g., Strawn, 2019). In this context, the acronym FAIR stands for findability, accessibility, interoperability, and reusability of data (Wilkinson et al., 2016), to enable Open Science by design (e.g., Strawn, 2019, European Commission and Innovation, 2018). In some ways, the FAIR principles are considered the "gold standard" for open data in the literature (e.g., Strawn, 2019). However, there are also doubts about an overstated discussion of Open Science, in particular Open Data according to FAIR principles, namely that the individual benefit to researchers is not sufficiently considered in this context (Allen and Mehler, 2019, Staunton et al., 2021). There is also lack tools to help researchers make data available in a FAIR manner. One additional issue to the discussion on open and FAIR Data is the aspect of simple data outflow without the community benefiting from contributions. The design of the artifact's interface plays a special role in this context (e.g., Ghazawneh and Henfridsson, 2010). These and other aspects and requirements should be considered when designing the artifact.

2.2. Existing platform solutions for mobility data

To identify the current state of research, we searched for existing mobility data platforms. The results were evaluated according to whether complementary data is provided alongside mobility data, and whether the findings of individual analysis are provided.
Such open science platforms already exist. For example, HUBzero (Mclennan and Kennell, 2010) enables scientists to share data online, work collaboratively and perform analysis. This solution is open source and can be implemented across disciplines, addressing a wide range of researchers.

We were also able to identify existing solutions for the area of mobility research. The majority of the solutions presented are limited to the provision of data directly related to mobility, or focus on a specific subarea of mobility research (Du et al., 2017, Vaisman and Zimányi, 2019, Pan, 2022,). We identified a lack of concepts that would allow non-mobility related complementary data to be provided in addition to mobility data. Furthermore, only a few solutions could be identified that provide different data and enable to carry out analysis on this data,

in order to make these findings available in a processed form (Jo et al., 2019, Vítor et al., 2021). In none of the identified solutions was it possible to carry out individual analysis on request.

In addition to these conceptual solution descriptions, there are various existing mobility data platforms that focus on specific countries, such as “Mobilitäts Daten Marktplatz MDM - MDM Portal” (n.d.). This platform is limited to the provision of mobility data with reference to Germany. The platforms “mCLOUD” (n.d.) and “Mobility Data Space - Mobility Data Space – die Data Sharing Community” (n.d.) also have a German focus, but offer complementary data such as climate and infrastructure data. Both platforms provide free and open access to the data. In contrast to the open data idea, in the future “Mobility Data Space - Mobility Data Space – die Data Sharing Community” (n.d.) will commercialize access to the data. An open science platform with a focus on geodata is MyGeoHub (Kalyanam et al., 2019). This solution is also built on the HUBzero (Mclennan and Kennell, 2010) platform and enables collaborative work between researchers. This solution lacks the provision of complementary data and the ability to submit analysis requests. While these solutions (currently) fulfill the requirement of open provision of data, they make no provision for sharing the findings of individual data analysis.

3. Methodology

In this paper we present the results of the development of a mobility data and knowledge platform prototype, which represents an IT artifact. According to March and Smith (1995) these can be classified into constructs (vocabulary and symbols), models (abstractions and representations), methods (algorithms and practices), and instantiations (implemented and prototype systems). Among other aims, DSR should generate scientific knowledge that shows how something has to be done to achieve a goal. For this reason, design knowledge includes DPs. These are prescriptive statements that show, among other things, how a construct must be designed to achieve a particular goal. Therefore, our concrete contribution is the definition of DPs and DFs that are used to instantiate our IT artifact. Design Science has yielded many different Design Science Research Methodologiess (DSRMs) (cf., Hevner and Park, 2004; Peffers et al., 2007). In the widely used DSRMs by Peffers et al. (2007) six concrete activities along the DSR process are defined which structured our project and the structure of this work.

During the first phase, *Problem identification and motivation*, we conducted four interviews with mobility researchers to identify possible problems in the context of mobility data and the availability of analysis findings. To cover both aspects, the interviews were conducted with researchers dealing with quantitative data analysis, and with those who are primarily interested in the processed findings of these analysis. To evaluate the problem statements, we have scanned the literature for the problems described and for existing solutions. This evaluation represents an ex-ante evaluation and resulted in justified problem statements. In the second phase, *Objectives of a solution*, we defined requirements to solve the problems, following the three steps presented in Meth et al. (2015), where DRs are initially defined, and DPs are derived that are later instantiated by DFs. For this purpose, we derived initial DRs from the "justified problem statements". By defining them, we were able to specify the scope of the proposed solution. Having analyzed the platform literature, we were able to validate the initial DRs and add additional ones.

Based on these DRs we derived DPs in phase three, *Design and Development*. For the construction of DPs, we followed the approach presented in Gregor et al. (2020). Then we instantiated the DPs from which concrete DFs were derived. Based on these DFs, we then implemented a working prototype. To enable the value creation of the hub, we collected mobility data and complementary data from various sources and made them available on the platform. In order to highlight the added value to researchers, we conducted exemplary analysis based on the data and made them available on the platform. We presented several thematically related analysis in aggregated form in a dossier section.
The fourth phases *demonstration (4)* and *evaluation (5)* were performed in one combined step. Following Sonnenberg (2012) Evaluation pattern, we conducted an ex-ante evaluation in the previous phases. This evaluation represented our ex-post evaluation. The evaluation consisted of organized interviews with two groups of mobility researchers. The first group consisted of two mobility researchers who apply quantitative methods and thus are presumably interested in raw data. The two researchers in the second group were concerned with the political and social science aspects of mobility. It can be assumed that the second group has a greater interest in processed data analysis and insights. We used the evaluation framework proposed by Venable et al. (2012), and adopted utility, quality, and usability as our central evaluation criteria (Hevner and Park, 2004). The focus groups had to operate the prototype. The actual evaluation of the prototype was done by answering a questionnaire. The communication step of the DSR process takes place through the publication of research findings. The reviewer feedback can also be represented

as an evaluation which we use to iteratively improve the prototype.

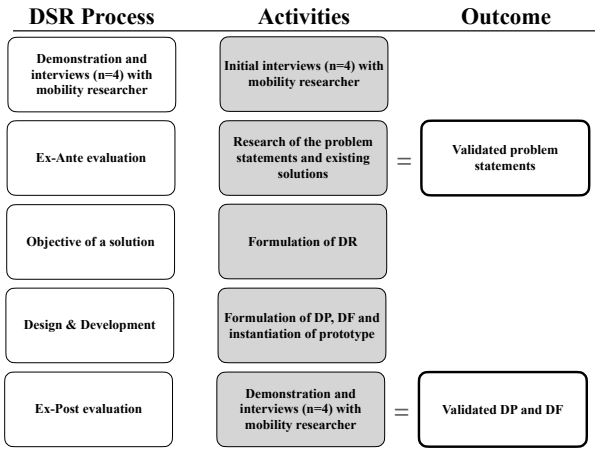

Figure 1: Our DSR process

4. Problem identification and objectives of a solution

4.1. Problem statements

Prior to the start of the DSR project, the problem of the lack of data and knowledge availability was expressed in discussions with mobility researchers. At the beginning of the project, we conducted interviews with two quantitative and two qualitative mobility researchers to identify the specific problem. The interviews showed, among other things, that mobility researchers who work quantitatively and conduct data analysis themselves often face a "lack of data". In practice, a lack of concrete mobility data is common, as is complementary data. According to the interviews, corresponding data often has to be generated first or, if it already exists, access to it is not known. For these scientists, the "lack of data" means time and financial effort that could be avoidable in some cases. A fragmented data landscape was cited as one reason for the lack of data. The following is a selection of researchers' quotes that relate to the lack of data. *"I spend too much time acquiring or generating data over and over again."*, *"Data acquisition is particularly time-consuming when the data comes from different sources."*.

In order to validate these problem statements, we searched for studies that investigate the described problems. The results showed that the described problems are indeed discussed in the literature and represent challenges for mobility researchers. The generalizability of the results of their study is limited by the lack of mobility data for 222 counties. Chen et al. (2021) and Heo et al. (2020) also faced limitations in COVID-19 mobility research due to the scarcity of mobility data.

In addition to researchers who analyze data quantitatively, researchers who study mobility from a political and social science perspective were also interviewed. According to the interviews, their research is less quantitative in nature, and they often lack the technical capabilities for quantitative data analysis. Despite their way of working, there is still a need for these researchers to obtain findings from the analysis of mobility data and complementary data. This problem is illustrated by the following quotes. *"I am often interested in possible insights from data but I am not familiar with data and statistical analysis."*, *"In my domain, more and more data is being generated, but only a few researchers know how to analyse this data properly."*. The literature has reflected this problem to some extent and shows that for some researchers, the process of data analysis in particular is a challenge (Savage, 2000). Scientists often lack the methodological skills but also the guidance to analyze data, especially larger data sets (Dierckx de Casterlé et al., 2012; Hunter et al., 2002; McCance (DPhil MSc et al., 2001). Froggatt (2001) argued that researchers inexperienced in data analysis often fail to exploit the full potential of the data and do not get beyond descriptive analysis. The analysis of the existing solutions showed that most of the solutions focus on the provision of mobility data, but do not support researchers who lack data analysis skills. Based on the interviews and the literature examined, it can be summarized that mobility researchers face two central problems: *(1) lack of mobility and complementary data*, and *(2) struggling with data analysis*. Based on these validated problem statements, the project goal of creating a mobility and knowledge platform for researchers emerged.

4.2. Design requirements

At the beginning of the "Objectives of a solution" phase, according to Meth et al. (2015), we derived initial DRs from the validated problem statements. An overview of the formulated DRs is shown in Table 1. **DR1** and **DR2** could be derived directly from the "Lack of mobility and complementary data". Since the lack of research data has been identified as a major problem, these DRs represent key demands on the platform. **DR3**, **DR5** and **DR11** could be deduced directly from the problem of the data analysis, since these DRs are essential to solve this problem. Since the planned solution is to create a platform, and not all DRs can be derived directly

from the problem statements, we were looking for literature that deals with the platform topic. The literature emphasized that the design of an open platform should facilitate its own sustainable growth. To enable this, our artifact should allow data donation from third parties **DR4**, incentivize researchers**DR9** to share their research questions and analysis findings with the platform, and prevent unidirectional data drainage **DR8**. The results also indicated that it is of central importance to grant appropriate rights to the platform participants. In our case, the platform owner should be able to curate the platform content and the data contributor should be able to control access to the contributed data. From this, **DR6** and **DR7** could be derived. The interviews and the literary source also revealed the relevance of an appealing appearance and straightforward usability. These requirements resulted in the **DR10**.

5. Designing the mobility data platform

5.1. Design: Design Principles

At the beginning of the design phase, we developed DPs that form the basis for the subsequent instantiation. Based on the previously identified DRs, we derived initial DPs formulated according to the Gregor et al., 2020 guidance. According to this approach, DPs comprise the intended *aim* (*A*), the *context* (*C*), the *mechanism* (*M*) to achieve the goal and the justification of the DP in the form of a *rationale* (*R*). The actors consist of an *implementer* (*I*) who instantiates the DP, a *user* (*U*), and an *enactor* (*E*) who is responsible for the execution of the mechanism (Gregor et al., 2020). Since mobility research provided the context for all the DPs, the implementer is the platform designer, and the rationale results from the argumentative construction of the DRs, these components are no longer mentioned in the following. Table 1 shows from which DR the respective DP is derived, and table 2 gives a short description of all nine DPs.

(DP1)**: Provide openly accessible mobility and complementary data A to researchers U through the platform owner E who collect mobility data and complementary data from different data sources and make them available on one platform M.* The implementation of the central requirements DR1 and DR2 is intended to address the lack of available data in research practice and to save resources. Next, based on DR3, we formulated the further central DP ***(DP2)**: Enable mobility researchers without a technical background (U) to gain insights from data A through the platform owner E who analyze the datasets and then provide them on a central platform M.* The analysis of data stored on the platform and the subsequent provision of the findings, in addition to the provision of data, represented one of the central requirements of the planned solution. In the next step we formed the third DP from the DR4: ***(DP3)**: Enable third parties U to contribute data (A) by providing the facility to third parties to upload their own datasets to the platform and annotate them with metadata M.* This mechanism is expected to contribute to the long-term growth of the platform. Next, we addressed DR5 through the fourth DP ***(DP4)**: Enable mobility researchers U to participate on the platform with their own research questions A by providing a facility enabling them to submit their own research questions with the corresponding data sets on the platform M.* The ability to submit research questions and indicate which datasets they relate to is intended to appeal primarily to non-technical mobility researchers. Thus, this DP is related to DP2. Subsequently, we formulated the fifth DP based on DR6: ***(DP5)**: Enable platform owners U to control the platform content A by providing them with the facility to add, change and remove content and allow to extend the rights to others as well M.* This ensures that there is always a control authority that can define and execute platform governance mechanisms. Next, we formulated ***(DP6)**: Enable data contributors U to control the self-contributed data A by providing the possibility to remove data contributed at any time M.* This DP was derived from DR7 and is intended to ensure that data donors have control over their own data at all times, regardless of other governance mechanisms. The ability to take data off the platform in case of doubt could also make data donors more willing to contribute to the platform. We then derived the seventh DP from DR8 and DR9: ***(DP7)**: Enable the incentivization of mobility researchers U to share analysis findings on the platform and keep them on the platform to avoid knowledge drain A by ensuring that the analysis and associated research questions are made available on the platform under the condition that all researchers have access to them M.* This mechanism is intended to ensure that when a user submits a research question, and it is answered, not only they personally, but everyone benefits from the knowledge. Next, we formulated DP8.(DP8)**: Enable mobility researchers U to find data and analysis intuitively and easily A by presenting data and analysis in a structured way and make them searchable and filterable M.* This DP refers mainly to the usability and the platform experience in general and addresses DR10. Through a modern interface through which data and analysis can be found intuitively, users should be supported in their actual work. Finally, we derived the ninth DP from DR11: ***(DP9)**: Enable mobility researchers U to conduct their own exploratory*

#	Design Requirement	DP	DF
1	Provide openly accessible mobility data	DP1	DF1
2	Provide openly accessible complementary data		
3	Provide prepared analysis results of mobility and complementary data	DP2	DF2
4	Enable data donation from third parties	DP3	DF3, DF5
5	Enable researchers to participate/contribute on the platform with own research questions	DP4	DF4, DF5
6	Enable (pro-)active data curation	DP5	DF5
7	Enable researchers to control access and usage of their provided data	DP6	DF5
8	Prevent unidirectional exploitative data drainage	DP7	DF2,DF4,DF6
9	Incentivize users to share research questions and analytical results		
10	Provide an intuitive and fast way to find data and analyses	DP8	DF2, DF7, DF9
11	Enable researchers to conduct their own exploratory data analysis	DP8	DF7, DF8

Table 1: DRs with the associated DPs and DFs

data analysis A by providing the facility to display data descriptions, descriptive statistics, data visualizations and performing data aggregations M. This should enable non-technical mobility researchers to obtain a first overview of the data and to perform smaller analysis without directly submitting a research question. At the current state, the artifact is a functional prototype that can be accessed online for evaluation purposes. It is already possible to upload data with annotations and tags via a corresponding form. Furthermore, mobility data can be searched for and downloaded. Moreover, data sets can be selected via a form and a corresponding research question can be submitted. At the beginning of the platform launch, data analysis will be carried out by the platform team. In the long term, researchers active on the platform can work on research questions and make the results available on the platform. The following link links to a public Github repository where the source code of the prototype and exemplary screenshots are provided. *https://github.com/daveknave/mobilityhub.berlin.hicss2023*.

5.2. Development: Instantiation of design through Design Features

The next step of the design phase was the construction of the mobility data and knowledge platform. To instantiate the previously defined DPs, nine DFs were derived from it, which are shown in table 1. For the construction of an artifact, DFs are more suitable since they are much more concrete and closer to the actual artifact than DPs Meth et al., 2015.

A *content management system* (*CMS*) represents our first DF (**DF1**). A *CMS* enables the simple management of web-based information without having to manually enter each piece of information into an HTML page. Through its implementation, we can provide data (DP1) and analysis findings (DP2) on the platform. Furthermore, data donors can use it to manage their contributed data (DP6). To make the analysis accessible to the platform users (DP2), we provide a user analysis area **DF2**, where the analysis stored in the *CMS* are visualized. In this way, we implement DP2, DP7 and DP8. To enable third parties to contribute their own data to the platform (DP3), the platform provides a data submission form (**DF3**), where data donors can annotate the data with meta-information and upload it through the underlying *CMS* in a common interoperable standard. Mobility researchers who want to have research questions answered by specific data can submit them through a research question submission form (**DF4**). This form gives access to several data sets on the platform and allows to select and submit the corresponding research question. This addresses DP4 and DP7. In order to enable and coordinate the individual access options available to platform owners, mobility researchers and data contributors, we have implemented a user management system (**DF5**). The user management system is realized through the *CMS* and enables the creation, editing and removal of users, and assigning users individual rights. Data donors can manage their contributed data in their user area, and mobility researchers can get an overview of their research questions. DP3, DP4, DP5 and DP6 are realized through this DF. According to DP9, mobility researchers should be able to carry out simple exploratory data analysis

on the platform. The aggregation of this data (**DF6**) is a central feature. For example, time series data can be aggregated temporally, variables can be grouped categorically in tabular format, and it is possible even to integrate different data sets into a combined data set. In order to gain a deeper understanding of the data after aggregation, a data analysis functionality (**DF7**) is provided. It can be used, e.g., to create descriptive statistics, output feature characteristics, and generate visualizations. This means that mobility researchers without the relevant technical skills can also carry out simple analysis. This feature addresses DP8 and DP9. As already described, one condition for answering individual research questions is that all platform participants have access to the processed findings (DP7). To ensure this, a researcher must agree to the terms and conditions (**DF8**) before submitting their research question. In order for mobility researchers to intuitively and quickly find the desired data and analysis on the platform (DP8), these are tagged (**DF9**). The tags allow cross-connections to be made between data and analysis in which this data is used. Conversely, cross-connections can also be established between analysis and the data sets.

#	Design Principle
1	Mobility and complementary data
2	Prepared analysis results of mobility and complementary data
3	Data contribution from third parties
4	Participate with own research questions
5	(Pro-)active data curation
6	Data control for contributor
7	User bonding and incentivization for participation
8	Intuitive interface for data exploration
9	Exploratory data analysis

Table 2: Short version of the DPs

6. Evaluation

Through demonstrations of the prototype and interviews with mobility researchers, we were able to evaluate the DPs and DFs, in particular alignment with the *FAIR* principles. All interviewed mobility researchers who regularly analyze and interpret data themselves fed back that our openly *accessible* provision of mobility and complementary data represents added value for them (i.e. by DP1, DP3). It was reported that the acquisition and *finding* of research-relevant data often takes a lot of time, especially when the data comes from different sources. Accordingly, there was mainly positive feedback for the provision of mobility and complementary data in one central place. In addition to access to data, all interviewed researchers much valued the possibility to contribute data sets themselves. They support the OS and OD approach, with one person recommending to adapt the solution in that direction. Data donors should have the possibility to provide their data on the platform with a DOI, and make it citable. Thus, people would be more motivated to contribute their own data to the platform. Another aspect that one researcher felt could be improved was the search of data and analysis. At the time, the platform content was mainly *findable* and searchable by categories and tags, which can lead to a large number of results (i.e. by DP8). According to the researcher, he usually searches for data on websites using a free-text search and recommends this for our platform as well. Both researchers found the operation of the platform basically very intuitive and visually appealing. They only wished that the data categories and data sets could be displayed in a much smaller format in order to provide a better overview.

The researchers from the social and political field of mobility research stated that our platform can solve the challenge of insufficient data analysis capabilities. Furthermore, they stated that the central provision of data on one platform would support them in their work. From this group of researchers, we received particularly positive feedback for the provision of complementary data, which has a positive impact on data *reuse* from a data generativity perspective (i.e. by DP7, DP9). One person stated that their research often requires contextualising data such as car registration numbers and historical weather data. Rather than a free-text search, they requested a more precise way of filtering data. From an *interoperability* point of view, it is important that only data based on current industry standards are offered (i.e. DF3).

Through the evaluation interviews, we were able to gather valuable feedback that helps us to further develop our DPs and DFs. Based on this, we will iteratively adapt the prototype. To improve the identification of relevant data and analysis, we will increase the filtering options and implement a free-text search. Furthermore, we will develop a feature for the platform to provide donated data with a DOI and make it citable. This could motivate more people to donate data to our platform and can thereby create network effects.

7. Discussion

This work has both practical implications for mobility research and research implications. Our central research contribution is the derivation of DPs that guide the design and construction of an Open Data Platform. With our

DPs and concretized DFs, we contribute to the discussion on OS, and address single challenges in the provision of open data (e.g., Hossain et al., 2016) through a platform approach with our artifact. While the specific formulation of DPs is focused on mobility researchers, the DPs can be generalized and applied to a wide variety of contexts, thus enabling open data and knowledge sharing among researchers. The same applies to the DFs. These DFs can be considered as a guideline to instantiate the DPs in the development of data platforms. By evaluating the DFs, we contribute to the existing body of knowledge of "descriptive knowledge". The potential usage of our DPs and DFs thus contributes to the expansion of the Open Science approach more widely and related concepts such as Open Data and Open Knowledge accordingly to the FAIR principles. However, the evaluation based on the FAIR principles also opens up a discussion for further improvements, especially in the outlook for the further development of the platform. In the evaluation, the acquisition of data for research in particular was identified as a very painful point. Thereby, metadata describing the data becomes very crucial to anticipate the potentials and possibilities for research with a dataset in a transparent way in advance. In industries such as medicine, meta-data platforms exist to bundle research associations and their platforms in order to substantially improve findability and acquisition of data.

In addition to our research contribution, our ultimate goal of providing a prototype is to support mobility researchers from all sorts of backgrounds in their practice. From our interviews and the searched literature (e.g., Audirac et al., 2022; Chen et al., 2021; Kliewer and Suhl, 2011), it became clear that the acquisition and generation of data requires precious researchers' time and efforts that could in part be avoided. The central provision of data on a platform is intended to counter this problem. Researchers who previously had to collect mobility and complementary data from different sources can thus save time and resources. Since the corresponding data often has to be acquired anew in practice, such an open-data platform can also save costs. Furthermore, research can be advanced by discovering new data on the platform and putting it into context with other data. Furthermore, it was identified at the beginning of the DSR project that there are mobility researchers who are interested in the findings from data but are not able to analyze them themselves (e.g., Brandão, 2015). This problem is addressed by providing individual analysis on the platform. Mobility researchers are thus given the opportunity to submit individual research questions with the corresponding data to the platform. Researchers associated with the platform can create the corresponding analysis and make the findings available to everyone.

For this work and the underlying DSR project, we encountered some limitations. We designed the DPs to be easily transferable to different domains. However, other research domains may have specific challenges regarding the nature of the data and its legal usability. Thus, the DPs may need to be more tailored to the targeted domain. Another limitation arises from the formulation of the DPs. We have derived these argumentatively, which may not be the optimal approach. Possible privacy implications are another limitation of our work. The extent to which we are allowed to bundle publicly available data in our platform is not always clear. For this aspect, we plan to consult a data protection expert. A limitation concerning the evaluation process is the small interview sample and the fact that two interviewees were involved in the ex-post and ex-ante evaluation. In the spirit of Sonnenberg, 2012 evaluation pattern, we are planning another evaluation with more participants for the next stage of expansion.

8. Conclusion

OS and related concepts such as OD and "open artefact" play an increasingly important role in supporting researchers in their work. Mobility research in particular, which is characterized by an ever-increasing amount of data, can benefit from this. According to the literature and our initial interviews, however, mobility researchers still report a lack of such data, or do not have the technical skills to analyze data themselves. Therefore, this paper describes our DSR project, the design and development of a mobility data and knowledge sharing platform. At the beginning we raise the question of how such a platform has to be designed to help mobility researchers to solve the described problems. We addressed this question by formulating DRs at the beginning and deriving DPs from it. Based on this, we were then able to formulate DFs that were used to develop a platform prototype, and to evaluate the design of the prototype through several expert interviews and demonstrations. The result of the evaluation will be used in the further course of the project to adapt existing DPs and DFs and to add new ones. One contribution of our work is to propose a design for a mobility and knowledge platform. The DPs and DFs developed for this purpose can also be transferred to other domains. By instantiating it, we make a practical contribution, as mobility researchers are thus supported in their research practice. By making the data and findings openly available, we contribute to the dissemination of OS.

References

Allen, C., & Mehler, D. M. (2019). Open science challenges, benefits and tips in early career and beyond. *PLoS Biology*, *17*(5). https://doi.org/10.1371/journal.pbio.3000246

Audirac, M., Tec, M., Meyers, L. A., Fox, S., & Zigler, C. (2022). Impact of the Timing of Stay-at-Home Orders and Mobility Reductions on First-Wave COVID-19 Deaths in US Counties. *American Journal of Epidemiology*, *191*(5), 900–907. https://doi.org/10.1093/aje/kwac027

Brandão, C. (2015). P. Bazeley and K. Jackson, Qualitative Data Analysis with NVivo (2nd ed.) *Qualitative Research in Psychology*, *12*(4), 492–494. https://doi.org/10.1080/14780887.2014.992750

Chen, X., Zhang, A., Wang, H., Gallaher, A., & Zhu, X. (2021). Compliance and containment in social distancing: mathematical modeling of COVID-19 across townships. *International Journal of Geographical Information Science*, *35*(3), 446–465. https://doi.org/10.1080/13658816.2021.1873999

Dierckx de Casterlé, B., Gastmans, C., Bryon, E., & Denier, Y. (2012). QUAGOL: a guide for qualitative data analysis. *International journal of nursing studies*, *49*(3), 360–371. https://doi.org/10.1016/j.ijnurstu.2011.09.012

Doyle, C., Chiu, Y.-T., Nagle, T., & Luczak-Roesch, M. (2021). Introduction to the Open Science Practices in Information Systems ResearchMinitrack. *Proceedings of the 54th Hawaii International Conference on System Sciences*.

Doyle, C., & Luczak-Roesch, M. (2019). *Extending the Boundaries of Design Science Theory and Practice - 14th International Conference on Design Science Research in Information Systems and Technology, {DESRIST} 2019, Worcester, MA, USA, June 4-6, 2019, Proceedings* (Vol. 11491). Springer International Publishing. https://doi.org/10.1007/978-3-030-19504-5

Doyle, C., & Luczak-Roesch, M. (2020). This paper is an artefact: On open science practices in design science research using registered reports. *Proceedings of the Annual Hawaii International Conference on System Sciences*, *2020-Janua*, 5026–5035. https://doi.org/10.24251/hicss.2020.619

Doyle, C., Luczak-Roesch, M., & Mittal, A. (2019). We Need the Open Artefact: Design Science as a Pathway to Open Science in Information Systems Research. In B. Tulu, S. Djamasbi, & G. Leroy (Eds.), *Extending the boundaries of design science theory and practice* (pp. 46–60). Springer International Publishing.

Du, Y., Zhao, C., Li, F., & Yang, X. (2017). An open data platform for traffic parameters measurement via multirotor unmanned aerial vehicles video. *Journal of Advanced Transportation*, *2017*. https://doi.org/10.1155/2017/8324301

European Commission & Innovation, D.-G. f. R. a. (2018). *Turning FAIR into reality : final report and action plan from the European Commission expert group on FAIR data*. Publications Office. https://doi.org/doi/10.2777/1524

Froggatt, K. A. (2001). The analysis of qualitative data: processes and pitfalls. *Palliative Medicine*, *15*(5), 433–438. https://doi.org/10.1191/026921601680419492

GarcíaPeñalvo, F. J., García Figuerola, C., & Merlo, J. A. (2010). Open knowledge: challenges and facts. *Online Information Review*, *34*(4), 520–539. https://doi.org/10.1108/14684521011072963

Gewin, V. (2016). Data sharing: An open mind on open data. *Nature*, *529*(7584), 117–119. https://doi.org/10.1038/nj7584-117a

Ghazawneh, A., & Henfridsson, O. (2010). Governing Third-Party Development through Platform boundary Resources. *ICIS*.

Gregor, S., Chandra Kruse, L., & Seidel, S. (2020). Research perspectives: The anatomy of a design principle. *Journal of the Association for Information Systems*, *21*(6), 1622–1652. https://doi.org/10.17705/1jais.00649

Heo, S., Lim, C. C., & Bell, M. L. (2020). Relationships between Local Green Space and Human Mobility Patterns during COVID-19 for Maryland and California, USA. *SUSTAINABILITY*, *12*(22). https://doi.org/10.3390/su12229401

Hevner, A., & Park, J. (2004). *Design Science in Information Systems Research* (tech. rep.). https://www.researchgate.net/publication/201168946

Hossain, M. A., Dwivedi, Y. K., & Rana, N. P. (2016). State-of-the-art in open data research: Insights from existing literature and a research agenda. *Journal of Organizational Computing and Electronic Commerce*, *26*(1-2), 14–40. https://doi.org/10.1080/10919392.2015.1124007

Hunter, A., Lusardi, P., Zucker, D., Jacelon, C., & Chandler, G. (2002). Making Meaning: The Creative Component in Qualitative Research.

Qualitative Health Research, *12*(3), 388–398. https://doi.org/10.1177/104973202129119964

Jo, J., Joo, I.-H., & Lee, K.-W. (2019). Constructing National Geospatial Big Data Platform: Current Status and Future Direction. *2019 IEEE 5th World Forum on Internet of Things (WF-IoT)*. https://ieeexplore.ieee.org/document/8767322/

Kalyanam, R., Zhao, L., Song, C., Biehl, L., Kearney, D., Kim, I. L., Shin, J., Villoria, N., & Merwade, V. (2019). MyGeoHub—A sustainable and evolving geospatial science gateway. *Future Generation Computer Systems*, *94*, 820–832. https://doi.org/10.1016/j.future.2018.02.005

Kliewer, N., & Suhl, L. (2011). A Note on the Online Nature of the Railway Delay. *Networks*, *57*(1), 28–37. https://doi.org/10.1002/net

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, *15*(4), 251–266. https://doi.org/10.1016/0167-9236(94)00041-2

McCance (DPhil MSc, B. H. T., Mckenna, H., & Boore, J. (2001). Exploring caring using narrative methodology: An analysis of the approach. *Journal of advanced nursing*, *33*, 350–356. https://doi.org/10.1046/j.1365-2648.2001.01671.x

Mclennan, M., & Kennell, R. (2010). HUBzero: A Platform for Dissemination and Collaboration in Computational Science and Engineering. *Computing in Science and Engineering*, *12*(2), 48–53.

mCLOUD. (n.d.). https://www.mcloud.de/

Meth, H., Mueller, B., & Maedche, A. (2015). Designing a requirement mining system. *Journal of the Association for Information Systems*, *16*(9), 799–837. https://doi.org/10.17705/1jais.00408

Mobilitäts Daten Marktplatz MDM - MDM Portal. (n.d.). https://www.mdm-portal.de/

Mobility Data Space - Mobility Data Space – die Data Sharing Community. (n.d.). https://mobility-dataspace.eu/de

Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., Percie Du Sert, N., Simonsohn, U., Wagenmakers, E. J., Ware, J. J., & Ioannidis, J. P. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 1–9. https://doi.org/10.1038/s41562-016-0021

Pan, H. (2022). Design and Research of Intelligent Traffic Cloud Platform Based on Flexible Customization of User Identity. *Tehnicki vjesnik - Technical Gazette*, *29*(3), 759–764. https://doi.org/10.17559/tv-20200220091709

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, *24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Savage, J. (2000). One voice, different tunes: Issues raised by dual analysis of a segment of qualitative data. *Journal of advanced nursing*, *31*, 1493–1500. https://doi.org/10.1046/j.1365-2648.2000.01432.x

Sonnenberg, C. (2012). Design Science Research in Information Systems. Advances in Theory and Practice. *7286*(May). https://doi.org/10.1007/978-3-642-29863-9

Staunton, C., Barragán, C. A., Canali, S., Ho, C., Leonelli, S., Mayernik, M., Prainsack, B., & Wonkham, A. (2021). Open science, data sharing and solidarity: who benefits? *History and Philosophy of the Life Sciences*, *43*(4). https://doi.org/10.1007/s40656-021-00468-6

Strawn, G. (2019). Open science, business analytics, and fair digital objects. *Proceedings - International Computer Software and Applications Conference*, *2*, 658–663. https://doi.org/10.1109/COMPSAC.2019.10283

Vaisman, A., & Zimányi, E. (2019). Mobility data warehouses. *ISPRS International Journal of Geo-Information*, *8*(4), 1–22. https://doi.org/10.3390/ijgi8040170

Venable, J., Pries-heje, J., & Baskerville, R. (2012). LNCS 7286 - A Comprehensive Framework for Evaluation in Design Science Research, 423–438.

Vítor, G., Rito, P., & Sargento, S. (2021). Smart City Data Platform for Real-Time Processing and Data Sharing. *2021 IEEE Symposium on Computers and Communications (ISCC)*. https://ieeexplore.ieee.org/document/9631427/

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). Comment: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*. https://doi.org/10.1038/sdata.2016.18