Barriers to Adopting Data Pools for Product Information Sharing – A Literature Review

Thomas Schäffer¹ and Dirk Stelzer²

¹ Hochschule Heilbronn, Studiengang Wirtschaftsinformatik, Heilbronn, Germany
   Thomas.Schaeffer@hs-heilbronn.de

² Technische Universität Ilmenau, Fachgebiet Informations- und Wissensmanagement, Ilmenau, Germany
   Dirk.Stelzer@tu-ilmenau.de

Abstract. Data pools, i.e., master data repositories that support companies in exchanging information on trade items with business partners in a standard format, have been rated as ideal tools for product information sharing. However, it is widely accepted that adoption of data pools has fallen short of expectations. There is an obvious mismatch between the benefits and advantages associated with the use of data pools and their actual diffusion. This paper presents a theory-based literature review into barriers to adopting data pools. The findings may support providers of data pools to improve their services and adopters of data pools to lay the foundations for an efficient use of these electronic intermediaries for product information sharing.

Keywords: Data Pools, Data Quality, Global Data Synchronization, Master Data, Product Information Sharing, Technology Adoption and Diffusion

1 Introduction

Product information sharing denotes the inter-organizational synchronization of master data relating to products [1]. The term master data refers to critical business objects of an organization that rarely undergo changes [2]. Master data synchronization means achieving consistent information values for items or products within and between organizations [3]. Product information sharing is a problematic, error-prone, labor-intensive, and costly process in most industries [4]. Companies may adopt a variety of standards, technologies, tools, and intermediaries to support product information sharing [1, 5, 6]. One of the most powerful tools for enhancing inter-organizational synchronization of master data is a data pool. A data pool is a repository that supports trading partners in obtaining, maintaining, and exchanging information on trade items in a standard format through electronic means [4]. Data pools combine standards and technologies to exchange product information from manufacturers’ to retailers’ internal databases through the use of electronic catalogs. These catalogs contain data describing articles and they also coordinate data
exchanges [5]. Instead of multiple bilateral exchanges a data pool acts like a clearing center. Manufacturers transfer data to the pool only once and retailers may collect the data from the pool [7]. Data pools assume the role of B2B intermediaries in data synchronization, but they do not focus on matching buyers and sellers [8]. The Global Data Synchronization Network (GDSN) [9] is a global network of specialized data pools. However, companies may also use other data pools that are not part of GDSN.

Gopal and McMillian characterize GDSN as one building block of successful collaboration in the supply chain [3]. De Cobière and Rowe [5] have analyzed various forms of interconnections to achieve data synchronization between companies. They come to the conclusion that data synchronization as supported by GDSN is an “ideal” form of product information sharing. However, it is widely accepted that the potential of data pools has not been fully realized, and that the adoption of data pools has fallen short of expectations [1, 4, 6, 7].

There is an obvious mismatch between the benefits and advantages associated with the use of data pools and their actual diffusion. Accordingly, several scholars ask for more intensive research into the issues of data pool adoption [1, 4, 7].

Some scholars have explored adoption and usage of data pools [1, 4, 5, 7, 10, 11]. They have identified issues and limitations of data pool adoption and usage in manufacturing and retailing companies. However, to the best of our knowledge, no author has yet attempted to provide a comprehensive overview of issues that companies face when adopting data pools for product information sharing.

The objective of our research is to identify reasons for the low adoption rate from previous research and to draw conclusions for providers and users of data pools and for future research. We aim to answer two research questions:

- What are the reasons for the low adoption rate of data pools?
- What conclusions can be drawn from identifying barriers to adopting data pools?

This paper is organized as follows. First, we give a short overview of our methodology. Next, we describe findings of our literature review. We then discuss limitations of our research. We conclude with a summary of our results and suggestions for providers and users of data pools, and for future research.

2 Method

2.1 Literature Review

We conducted a literature review to identify prior research into adoption barriers and issues when using data pools for product information sharing. We followed the guidelines provided by Webster and Watson [12] and vom Brocke et al. [13] to identify relevant publications. As a first step, we examined IS journals and IS conference proceedings using the AIS Electronic Library, ScienceDirect, Google Scholar, and SpringerLink. We conducted electronic searches in titles on the following search term: [“data pool” OR “Datenpool” OR (“electronic” AND (“catalog” OR “catalogue”))]. Furthermore we did a full text search on the following
search term: ["master data pool" OR "Stammdatenpool"]]. These searches identified a total of 793 publications. After analyzing each article’s abstract, keywords, or the full article when necessary, we excluded 757 articles that were duplicates or did not appear to be concerned with or relevant to our research focus. As a third step, we performed forward and backward searches in relevant articles to identify further sources. A total of 45 publications were read in full and coded. We excluded all papers that did not seem to be peer reviewed publications or that only stated the keywords mentioned in the search term without elaborating on these concepts. We also excluded all papers that did not mention issues, limitations, and weaknesses of data pools, or adoption barriers, respectively. Out of the 45 coded articles, 23 include passages of interest. A complete list of all papers and a documentation of the coding procedure is available from the authors upon request. Table 1 in section 3 gives an overview of the 23 publications that elaborate on barriers to adopting data pools. In the remainder of this paper we focus on these papers.

2.2 Selection of a Theoretical Basis

In the course of our literature review we have identified several candidates that may serve as theoretical or conceptual basis for coding and analyzing the papers. In the following section we briefly discuss the most widely used theories and models.

Coordination theory is the most commonly used theory in research on product information sharing [1, 5, 6]. However, coordination theory is not a comprehensive theoretical basis for assessing barriers to the adoption of technologies or tools.

The transaction cost approach may help to assess economic aspects of adopting data pools for product information sharing. However, apart from a few rare exceptions, e.g., [3, 8, 14, 15], prior research has not attempted to quantify transaction costs or benefits related to data pool adoption and usage. Thus, literature does not yet provide enough empirical data for such research.

Madlberger [7] has explored adoption barriers of data pools from the perspective of diffusion and adoption of standards. She identified network effects, standardization costs, excess inertia, and path dependencies as potential barriers or facilitators of adopting data pools. However, as data pools are a combination of standards, technology, and data, diffusion and adoption of standards alone falls short of a comprehensive analysis of adoption barriers of data pools [1, 5].

The technology acceptance model is another candidate for a theoretical basis. However, the level of analysis of this model is the individual. Decisions to adopt data pools or to refrain from doing so, however, are made on a group, organizational, or even industry level [7]. Consequently, this model is inadequate for our research.

The diffusion of innovations theory [16] explains how innovations are adopted by social systems, i.e., individuals, groups, organizations, industries, or societies. An innovation is an idea, practice, or object that is perceived as new. Thus, adopting a data pool in a company can be regarded as an innovation [7, 10]. The most important characteristics of innovations that explain the rate of adoption are: relative advantage, compatibility, complexity, trialability, and observability. The diffusion of innovations theory seems to be well suited for our research objective.
Information systems (IS) success models – the most famous models were published by DeLone and McLean [17, 18] – conceptualize system quality, information quality, and service quality as important antecedents to a system's usage, organizational impact, or its net benefits. Data pools can be regarded as inter-organizational IS [4, 11]. Given the fact that quality is a key characteristic in many publications on the adoption of data pools for product information sharing [6, 7, 11, 19], IS success models may also be well suited for our research.

Usually, diffusion of innovations theory and IS success models are used for explaining the adoption of innovations or success of IS. In the present paper, however, we strive for identifying barriers to adoption, i.e., we do not search for enablers but for inhibitors of adoption. Consequently, the characteristics of innovations and the success factors must be interpreted contrariwise.

Diffusion of innovations theory focuses on the following characteristics of innovations that influence their adoption [16]: Relative advantage, the degree to which an innovation is perceived as better than the idea it supersedes; compatibility, the degree to which an innovation is perceived as being consistent with the existing values, past experiences, and needs of potential adopters; complexity, the degree to which an innovation is perceived as difficult to understand and use, trialability, the degree to which an innovation may be experimented with on a limited basis; and observability, the degree to which the results of an innovation are visible to others.

From the IS success models we deduced the following antecedents to system usage: System quality denotes desirable characteristics of an IS, e.g., functionality, ease of use, or system flexibility and reliability [17]. Information quality describes desirable characteristics of the system outputs, e.g., relevance, understandability, accuracy, completeness, understandability, timeliness, usability [17]. Service quality denotes quality of the service or support that users receive from the IT support personnel. Key service quality characteristics are responsiveness, accuracy, reliability, technical competence, and empathy of the personnel staff [17].

2.3 Coding Scheme

We used these criteria or, more precisely, their opposite, for identifying barriers to adopting data pools in a first iteration of our literature analysis. After having analyzed all 23 papers, we found that none of the two approaches is suitable to categorize each of the barriers mentioned in the literature. Furthermore, some of the criteria do not seem to be relevant for the adoption of data pools as they are not mentioned in the literature. We did not identify a single reference to issues relating to trialability, observability, or issues with service quality of data pools. We therefore deleted these factors from the second iteration of our analysis. Furthermore, we discovered that issues relating to system and information quality are inherent characteristics of data pools. Another characteristic that directly relates to data pools and that is mentioned frequently in the literature is cost, i.e., fees for access, licenses, and usage that users of data pools have to pay. We therefore decided to complement system and information quality by cost of data pools and label these three criteria as characteristics of data pools. They are perceived as barriers to adoption.
Apart from factors directly relating to characteristics of data pools, some authors [3, 10] suggest to distinguish two more perspectives on adopting data pools.

- Inter-organizational factors denote environment characteristics that influence the adoption of data pools. When analyzing the literature for inter-organizational factors, we identified the following categories:
  - As data pools are subject to network effects they require a critical mass of data providers and data recipients to become attractive [1, 7].
  - The decision to adopt a data pool is depending on the perceived commitment of relevant business partners to use the same data pool and on the trust that business partners make available high quality data [4].
  - External IT providers have supported many companies in inter-organizational synchronization of master data before they decided to adopt a data pool. Therefore, external IT providers may act as gatekeepers that may facilitate or inhibit data pool adoption [20].

- Intra-organizational factors denote factors relating to users of data pools, i.e., data suppliers and data recipients. We identified three categories of such factors:
  - Internal capabilities (e.g., knowledge of personnel, processes or process areas, technology and infrastructure, and organization structures) of data providers and data recipients may impede or prevent the adoption of data pools [1, 6].
  - As the adoption of a data pool requires considerable internal adaptations and investment, a high management commitment is indispensable [4].
  - The adoption of a data pool requires considerable efforts for internal adaptation, e.g., setup costs for electronic integration. Perceived high internal costs may become an inhibitor of data pool adoption [1, 3, 6, 21].

Figure 1 gives an overview of the methodology used in our research.
3 Findings

In the following section, we briefly present barriers to adopting data pools for product information sharing that we identified when analyzing the literature. We use the coding scheme presented in the previous section as a structure for our findings.

3.1 Characteristics of Data Pools

Data pools are combinations of standards and technologies to exchange product information [5]. Consequently, some barriers to adoption relate to standards used by data pools and others concern data pool technologies.

System Quality

Data pools do not always seem to apply adequate mechanisms for ensuring the quality of product master data included in the pool [22]. They trust that data providers supply data of the desired quality. However, this is not always the case. This leads to insufficient data quality in some cases [8, 21].

Both manufacturers and retailers miss feedback loops to improve data quality. They articulate the need for providing and receiving feedback on the quality of data available in the data pool. Currently, data pools do not offer such mechanisms [11].

Several companies use more than one data pool. As each data pool has a unique user interface, users complain about complicated usage of data pools [6].

Process standards supporting usage of data pools do not seem to be sufficient yet. For example, the standards do not include definitions of responsibilities for plausibility checks or liability agreements for incorrect data [23].

Data pools do not necessarily provide options for definition and handling of product subsets, product families, or product bundles [24, 25]. Moreover, GDSN does not provide options for identifying specific physical products. This complicates supply chain management activities, e.g., product recalls or drug tracking [26]. These shortcomings make it difficult to use data pools in day-to-day operations.

Information Quality

Most data pools restrict their offerings to a specific industry or product category. GDSN, for example, specializes on fast-moving consumer goods [8]. Retailers that also procure apparel or automotive parts usually have to obtain master data from more than one data pool. In other words, as data pools do not cover all product data required by many companies, completeness of information is an issue [27, 28].

Several users complain about a limited number of attributes [6]. Data pools do not provide all attributes needed for executing business processes. Prices and discounts, for example, are often negotiated on a bilateral basis. Consequently, these data have to be exchanged directly between business partners [5, 7, 20, 23]. This requires maintaining additional communication structures. Given that most companies must establish and support dyadic communication channels in addition to exchanging data with data pools, many companies question the cost-effectiveness of using data pools [4, 8, 23].
Some papers report on complaints about missing or wrong identification numbers for logistic units, e.g., packaging units or pallets [29]. This is a result of inadequate validation mechanisms of data providers as well as of data pools [22].

Another weakness is closely related to characteristics of standards. GS1, for example, prescribes several mandatory fields, e.g., height, width, and depth of products [27, 30]. Although this information is relevant for most products, this does not apply, for example, to nets of onions. However, suppliers of onions are forced to enter corresponding data [11]. This demotivates data providers and lowers their readiness to provide high quality data. This corresponds to a finding reported by Legner and Scherm [1] who pointed out that data synchronization needs to be enhanced by providing context information. Otherwise data may often be misinterpreted.

Inadequate timeliness is another issue in many data pools [11]. In case of new product launches delays in the provision of data may impede the efficient execution of business processes.

Moreover, data availability is only guaranteed for a limited time in most data pools. Relevant data are not always available throughout the product life cycle [24].

**Cost of Data Pools**

Prices charged by providers of data pools, i.e., fees for access, licenses, and usage, are perceived as too high by many suppliers and recipients of data [1, 3, 4, 6, 7]. This problem is aggravated by the lack of transparency of data pools’ license models [14].

### 3.2 Inter-Organizational Factors

**Critical Mass of Data Pool Users**

Like most intermediaries in electronic business, data pools are subject to network effects, excess inertia, and path dependencies [4, 7, 30]. Data pools will not be able to attract enough data recipients unless they have already attracted enough data providers, and vice versa. However, several of the issues and drawbacks of data pools presented in the last section do not make data pools appear particularly attractive. This, in turn, prevents more companies from participating in data pools.

**Commitment and Trust**

Business relationships built upon a data pool demand a long-term commitment from all relevant business partners. Additionally, product information sharing via data pools cannot be achieved without a strong commitment to provide high quality data [27]. However, many companies are not sure whether all trading partners are sufficiently dedicated to provide data on an adequate quality level and to maintain long term business relations with a data pool [4, 30].

**External IT providers**

Boukef Charki et al. point to the power of external IT providers that may act as gate keepers for adopting data pools [20]. If these IT providers perceive data pools as threats to their business models they might try to prevent their clients from participating in data pools.
3.3 Intra-Organizational Factors

Internal Capabilities

Insufficient internal capabilities may prevent data suppliers and data recipients from efficiently participating in product information sharing via data pools. Major problem areas are personnel, business processes, technical infrastructure, and organizational structures.

Small organizations regularly lack IT skills, knowledge, and capabilities for inter-organizational data synchronization. In larger organizations IT personnel is often little aware of the product information sharing process in their own organizations [11]. Master data experts are often not directly involved in this process. Thus, there is no central unit to make a well-informed decision about adopting data pools [1, 6, 8, 30].

In many companies, responsibilities for master data management, data quality, and data governance are scattered over several lines of business. Other companies have not defined adequate strategies for master data management and inter-organizational data synchronization [1, 6, 15, 30, 31]. This again, prevents many companies from adequately assessing potential benefits of using data pools.

Using data pools requires adjustments of internal IS. This is not only a lengthy and complex procedure but also requires technical skills and capabilities. However, staff members responsible for product information sharing often neither have adequate IT skills, nor sufficient support from their IT departments to manage necessary adjustments [4, 6, 15].

Inadequate organizational structures often impede efficient product information sharing. In many companies, sales and procurement staff are responsible for exchanging product data with trading partners [23]. Often, they are not sufficiently motivated to improve product information sharing and they usually do not have necessary technical capabilities to adequately assess the potential benefits of using data pools [1, 6, 23].

Management Commitment

Many senior managers do not seem to be ready to invest in data pools because they do not fully understand the necessity of automated product information sharing and the economic benefits of using data pools [4, 15, 30].

Internal Cost

Organizations that consider adopting a data pool are confronted with significant setup costs for electronic integration and additional cost for internal adaption of personnel capabilities, business processes, and organizational structures [15, 30]. In some cases they are afraid of the additional effort required, the lengthy implementation procedure, and the perceived high costs for installing, updating, and maintaining new systems to compile and exchange data with a data pool [15, 32].

Consequently, several scholars have identified lack of internal readiness as a major inhibitor of adopting data pools for synchronizing master data [1, 3, 4, 15]. Table 1 gives an overview of barriers to adopting data pools as identified in the literature.
Table 1. Overview of barriers to adopting data pools

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Becker et al. 2007 [25]</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Becker et al. 2008 [31]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boukef Charki et al. 2011 [20]</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen &amp; Prater 2013 [15]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Corbière &amp; Rowe 2011 [32]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Corbière &amp; Rowe 2013 [5]</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gopal &amp; McMillian 2005 [3]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Húner et al. 2011 [29]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Karpischek et al. 2012 [28]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Karpischek et al. 2014 [27]</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Legner &amp; Schemm 2008 [1]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Madlberger 2011 [7]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nakatani et al. 2006 [4]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schäffer &amp; Stelzer 2017 [6]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schemm &amp; Legner 2008 [8]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schemm 2009 [14]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schemm et al. 2006 [21]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schemm et al. 2008 [23]</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winkelmann et al. 2008 [24]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhou et al. 2006 [26]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhou et al. 2011 [22]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4 Discussion

To the best of our knowledge this is the first review of research into issues that companies face when adopting data pools for product information sharing. Our research has identified a significant number of limitations and drawbacks that can be understood as barriers to adoption. All in all, these barriers may explain why data pools are used so little for product information sharing.

Our review has shown that many companies evaluate system and information quality and the price/performance ratio of data pools as unfavorable. Moreover, many users perceive relative disadvantages of a data pool when they compare it to more traditional forms of data synchronization, e.g., bilateral exchanges of product master...
data via EDI connections or spreadsheet files [1, 5–7]. Moreover, companies that have not already implemented master data management and internal data synchronization systems perceive data pools as not being consistent with their internal structures, systems, and processes.

Our review has revealed that none of the theories and conceptual approaches presented in section 2.2 fully covers all barriers that were identified in the literature. IS success models support the identification of unfavorable properties of data pools. However, these models are not appropriate when it comes to identifying and analyzing barriers that are not inherent characteristics of data pools. Diffusion of innovations theory puts the focus on users of data pools, i.e., providers and recipients of master data. However, characteristics put forward by this theory are too abstract for a detailed identification of barriers to adopting data pools. Based on the literature we have developed a more specific scheme that helped us to identify and to classify barriers to adopting data pools for product information sharing.

As several of the papers included in our review are more than ten years old, some of the issues with data pools may have been superseded in the meantime. Some of the barriers shown in table 1 are mentioned in a few papers only. It might be helpful to explore whether these properties are relevant for individual companies in specific situations only or whether they keep many companies from adopting data pools. Most of the papers included in our review focus on GDSN data pools. Our review does not distinguish between these data pools and others that are not part of GDSN. More empirical research is needed that explores current issues relating to data pools.

5 Conclusions

From the findings of our review we draw conclusions for providers and users of data pools and for future research.

Implications for Providers of Data Pools

Providers of data pools should strive for improving the price/performance ratio of their services, either by improving system and information quality or by lowering fees for access, licenses, and usage of data pools.

Several users criticize that they have to maintain bilateral communication channels for exchanging master data with business partners in addition to using data pools. If data pools supported private communication channels that companies could use to exchange confidential information on a bilateral basis, e.g., prices or discounts, this would probably be perceived as a helpful functional enhancement.

Implications for Users of Data Pools

Before adopting data pools, firms must implement company-wide master data management and internal data synchronization mechanisms. They should also adapt knowledge and capabilities of staff members, business processes, IT infrastructure, and organization and governance structures. Otherwise, they take the risk of incompatibilities between their internal structures and data pools.

When calculating a business case for data pool adoption, companies should not only criticize the perceived high costs of data pools. They should rather compare it to
all benefits and costs of the error-prone, labor-intensive, and costly processes of traditional master data synchronization.

A detailed business case could also help to ensure senior management commitment for adopting a data pool.

Implications for Scholars

It is noticeable that only few of the barriers to adopting data pools can directly be related to the theories and models discussed in section 2.2. It should be considered to put forward a new theoretical basis for conducting research on data pool adoption.

The coding scheme presented in section 2.3 was derived from theories and conceptual approaches and from literature on data pools for product information sharing. However, we assume that more research could lead to more detailed schemes for structuring and categorizing barriers to the adoption of data pools.

We recommend more empirical research into current issues relating to data pools, e.g., in the form of case studies or surveys.

A more detailed analysis into different categories of data pools, e.g., distinguishing between GS1 data pools and others, could help to obtain a more accurate picture of barriers to adoption of specific data pools.

References

5. de Corbière, F., Rowe, F.: From Ideal Data Synchronization to Hybrid Forms of Interconnections: Architectures, Processes, and Data. JAIS 14, 550–584 (2013)