AMCIS 2023 Proceedings

Strategic & Competitive Uses of Information and Digital Technologies (SCUIDT)

Aug 10th, 12:00 AM

Creating Data Policies for Digital Business Ecosystems

Vitor Hugo Machado Ribeiro
University of Coimbra, vhribeiro@dei.uc.pt

João Barata
University of Coimbra, barata@dei.uc.pt

Paulo Rupino da Cunha
University of Coimbra, rupino@dei.uc.pt

Recommended Citation

Ribeiro, Vitor Hugo Machado; Barata, João; and da Cunha, Paulo Rupino, "Creating Data Policies for Digital Business Ecosystems" (2023). *AMCIS 2023 Proceedings*. 14.
https://aisel.aisnet.org/amcis2023/sig_scuidt/sig_scuidt/14

Creating Data Policies for Digital Business Ecosystems

Completed Research

Vítor Ribeiro
University of Coimbra, CISUC, DEI
Coimbra, Portugal
vhribeiro@dei.uc.pt

João Barata
University of Coimbra, CISUC, DEI
Coimbra, Portugal
barata@dei.uc.pt

Paulo Rupino da Cunha
University of Coimbra, CISUC, DEI
Coimbra, Portugal
rupino@dei.uc.pt

Abstract

Data policies are high-level guidelines that define how the organization handles data. In the age of cross-organizational digital business ecosystems, organizations are facing the need to define data policies that allow them to operate in decentralized scenarios. We conducted a Design Science Research (DSR) project to develop an approach for data policy management for digital business ecosystems. Our artifact was developed and demonstrated in a leading European IT provider. Our results include (1) an approach for data policy creation and (2) a data policy cycle. For theory, our work extends the literature with an approach for data policy management in digital business ecosystems in a highly regulated sector. For practice, our approach can support practitioners in developing the necessary data policies based on their context, considering their location, data-related regulations, available data assets, and organizational environment.

Keywords

Data governance, data policy management, design science research, digital business ecosystem

Introduction

The exponential growth in data volume, format, and typology makes data governance a prime concern (Khatri & Brown, 2010; Lis & Otto, 2021). Traditionally, data governance provides company-wide mechanisms and tools for dealing with multiple data sources, developing new data-based solutions, handling data-related risks, and improving data quality (Abraham et al., 2019). Why every organization should develop a data governance strategy and deploy policies is now well accepted (Alhassan et al., 2019).

However, data governance has recently been moving beyond organizational boundaries (Davidson et al., 2023). Examples of this trend are the digital business ecosystems that define a setup of a collaborative environment that integrates multiple organizations that co-create value enabled by communication and information technologies (Nachira et al., 2007). In this scenario, complexity rises due to decentralized data assets and the need to align the objectives of different parties (Jagals & Karger, 2021; Lee et al., 2019; Scholz et al., 2022). Decentralization requires new ways of defining data ownership and access rights and dealing with multiple regulations (Lee et al., 2019). Furthermore, governance must foster collaboration and facilitate data exchange (Lis & Otto, 2020). However, extant research on decentralized data governance is still immature (Davidson et al., 2023; Lee et al., 2017).

Data policies are key pillars of data governance frameworks (Abraham et al., 2019). They are short statements defining the high-level rules to handle data (Data Management Association, 2017). Therefore, organizations must find appropriate ways to create, deploy, and monitor their data policies (Alhassan et al., 2019). Data policies depend on the organizational context (e.g., geographical location, regulations, and strategic objectives) (Zuiderwijk & Janssen, 2014). For example, the General Data Protection Regulation (GDPR) has forced organizations using European citizen's data to develop data policies for privacy purposes (e.g., data storage, analysis, and retention requirements). Yet, notwithstanding their importance, data protection is only one facet of modern data governance requirements, and there is a shortcoming of data policy creation guidelines for business ecosystems (Davidson et al., 2023). Our literature review revealed a lack of an approach for data policy management in digital business ecosystems since the current solutions focus on a single organization (e.g., data policy management cycle (Loshin, 2010)).

This paper presents the results of a Design Science Research (DSR) project in cooperation with a leading European IT provider (e.g., software-as-a-service solutions to be used by business parties) and its business ecosystem partners. The data governance requirements of the company changed drastically when they decided to extend the traditional products for telecommunications operators (e.g., routers for the Internet) with other digital platforms heavily dependent on data analytics, artificial intelligence (AI), and machine learning (ML) models. Their data policies were obsolete for scenarios where data may be acquired, transformed, or sold across a network of companies. Therefore, the research question was formulated:

RQ: *What are the necessary mechanisms for implementing data policy management in digital business ecosystems?*

The remainder of this paper is structured as follows. The background section provides an overview of data governance, data policies, and data policy management. Then, the research methodology to develop the artifacts is described. Subsequently, we present the proposed approach for data policy management in digital business ecosystems. Next, the demonstration is reported, followed by the discussion. The last section presents the main conclusion and limitations.

Background

Data Governance in Digital Business Ecosystems

Data governance aims to capitalize on data as a key corporate asset and address data-related risks (Abraham et al., 2019). Generally, a data governance framework defines the decision rights and the accountability for the organization's data assets (Khatri & Brown, 2010). Furthermore, it defines the organizational structures and the policies, processes, and standards that guide all the data-related activities of the organization (Data Management Association, 2017).

With organizations shifting from centralized to decentralized scenarios, there is a need to develop new data governance mechanisms (Davidson et al., 2023). The increasing collaboration between the parties requires deploying data-sharing mechanisms and the technical structures that support them (De Prieelle et al., 2020). Moreover, there is a need to enforce the ecosystem's policies, standards, and processes across the partners (Lis & Otto, 2020). Contribution measurement (e.g., metrics to measure the data contribution of each partner) and members' access rights (e.g., restricting the use of shared data for a specific purpose) are critical in ecosystems (Lee et al., 2019). Therefore, data usage must be monitored across the ecosystem's members, including conformance, to avoid data misuse and access (De Prieelle et al., 2020).

Data policies can govern all aspects of the data lifecycle (Loshin & Reifer, 2018). Consequently, modern data governance frameworks should define a set of data policies (e.g., data security, and data quality) and establish compliance monitoring mechanisms (Abraham et al., 2019; Loshin & Reifer, 2018). Inter-organizational data policies are also a requirement (Lee et al., 2019), as presented in the next section.

Data Policies

Policies establish the general regime of organizations regarding a subject (e.g., data, environment), how it is implemented, and how it produces its actual impact (Zuiderwijk & Janssen, 2014). Data policies provide general guidelines and rules related to data creation, collection, storage, security, quality, and acceptable use (Alhassan et al., 2019), covering critical objectives for data, data accountabilities, data roles, and data retention periods (Data Management Association, 2017). Data standards and processes define "how" data is dealt with, while data policies complement them by describing what is possible "to do" and "not to do" by the members of the organization (Data Management Association, 2017). We define and position our work according to the view of Abraham et al. (2019) on data policies and data governance.

Data policies can be of several types: data lifecycle, data security, data privacy, data quality, data architecture, data provenance, data integration, and data storage and retention (Abraham et al., 2019; Data Management Association, 2017; Loshin & Reifer, 2018). Data security policies can ensure that the right people can use and update data correctly and that all inappropriate access and update is restricted (Data Management Association, 2017). Data privacy policies aim primarily to promote compliance with regulations to avoid irregular situations, establishing high-level guidelines for the privacy of data assets (Data Management Association, 2017). Most data-related regulations consider defining policies for data access control, data retention and deletion, auditing, and categorizing sensitive data (TM Forum, 2022). Data storage policies are deployed to manage the organization's data assets' retention, archival, and disposition (Loshin & Reifer, 2018). Data quality policies address periodic quality audits, data standards, and best practices (Data Management Association, 2017). Data architecture policies include the specification of data models and technologies to transmit data (Loshin & Reifer, 2018).

Literature addressing data policy design is scarce. A noteworthy example presented by Joel et al. (2001) proposes a template for data security policies, including the scope, the statement, the responsibilities, the enforcement mechanisms, and the review and monitoring process (Joel et al., 2001). The data policies must be implementable (e.g., using system administration procedures), enforceable, viable in the long term, and independent of technological decisions (Joel et al., 2005). The types of data policies needed vary according to the context of the organization (e.g., market, geographical location, wealthiness, compliance requirements, and type of data used) (Kraemer et al., 2002; Zuiderwijk & Janssen, 2014). Surprisingly, how to deploy data policies adapted to inter-organizational contexts is absent in the literature.

Data Policy Management

Policy development follows the cycle of (1) "*agenda setting*", (2) "*policy formulation*", (3) "*policy implementation*", (4) "*policy evaluation*", and (5) "*policy change or termination*" (Stewart Jr et al., 2007). Data policies must be continuously revised to explore new business opportunities (Alhassan et al., 2019). Moreover, the organization must show evidence of compliance with internal (e.g., data security) and external (e.g., GDPR) data policies (TM Forum, 2022). Data policy management is cross-departmental and includes people (e.g., the appointment of roles for data policy management), processes (e.g., data policy monitoring processes), and technology-related topics like the deployment of tools to support the monitorization of policies (TM Forum, 2022), as shown in Table 1.

Domain	Mechanisms
People	• Define data policy management roles and responsibilities. • Assign data policy managers responsible for leading the development, communication, and implementation of data policies.
Processes	• Define processes to monitor data policy compliance, review and assess the existing data policies, and implement data policies. • Document and manage all the processes related to data policies. • Provide detailed documentation on the data policies.
Technology	• Deploy tools to support data policy compliance, monitoring, and enforcement. • Use tools to automate data policy management tasks. • Develop a data policy inventory and a link to the respective data assets.

Table 1. Data Policy Management Mechanisms, adapted from TM Forum (2022)

Data policy management in business ecosystems includes decentralized data policy enforcement (Kravets & Zimmermann, 2012) and compliance audit (Lee et al., 2019). Moreover, the data policies must be available in a shared repository, and all the relevant stakeholders must be notified whenever policies are added or modified (Stanford University, 2011). Loshin (2010) introduces a lifecycle approach to data policy management. However, this approach is focused on a single organization, since it does not consider the ecosystem's business analysis, partner's data usage rights, and decentralized policy monitoring. Despite the need for data policy management in business ecosystems, there is a lack of an approach in the literature. As a result, policies are often "*imprecise, and thus how, when, and who uses the data*" is unclear (e.g., imprecise data ownership, untraceable data usage) (Lee et al., 2019). This critical challenge must be handled (Espinosa et al., 2019; Lee et al., 2019) to promote a more trustful relationship between the parties (Schreieck et al., 2016).

Methodology

DSR is an iterative process to design artifacts to solve observed problems, make research contributions, evaluate the designs, and communicate the results to appropriate audiences (Hevner et al., 2004). Furthermore, the artifact should be relevant to solving a relevant business problem, and its utility, quality, and efficiency must be rigorously evaluated (Hevner et al., 2004; Peffers et al., 2007). DSR can turn individual experiences into usable data, provide practical application content, and use a fluid operational structure instead of rigid guidelines (Hevner et al., 2004).

We identified as a solution the development of an approach for data policy management for digital business ecosystems that includes policy creation, implementation, monitoring, and review. We selected DSR to develop and evaluate our proposed approach for data policy management since it is a problem-solving paradigm that relies on kernel theories to produce innovative artifacts intended to solve identified organizational problems (Hevner et al., 2004). Peffers et al. (2007) suggest an iterative process that includes "*problem identification and motivation, define objectives of a solution, design and development, demonstration, evaluation, and communication*". Our DSR project had a problem-centered initiation (Peffers et al., 2007) and included a background review of data governance, policies, data policies, and data policy management. Moreover, we have established contacts with industry experts in healthcare software solutions development. The literature review provided 126 results in Google Scholar using the keywords "data policy management" OR "data policies management" OR "data policy creation" OR "data policy definition" OR "data policies creation" OR "data policies definition", excluding citations and patents. For Scopus and WoS, the same keyword returned 5 and 9 results, respectively. The selection of papers addressing data policy creation was presented in the previous section.

After developing the artifact, we demonstrated and evaluated the results in a real-world case. We instantiated the artifact in a business ecosystem, developing its data governance framework and defining a set of data policies. The next section describes the artifacts that were created during the DSR project.

A Data Policy Management Approach for Digital Business Ecosystems

How to create and continuously update data policies of different types (Abraham et al., 2019; Data Management Association, 2017; Loshin & Reifer, 2018) that are coherent, not redundant, and easy to follow by different departments or organizations is the challenge.

We started the design and development of our artifact together with the experts. They were involved in a digital business ecosystem that aims to develop an intelligent healthcare platform to provide advice on training exercises to diabetic people based on their current diabetes blood values. Healthcare is a highly regulated context, and health-related data is extremely sensitive.

The platform team includes (1) a central hospital, (2) a university research and development team, (3) and an IT provider integrated into a major telecommunications company in Europe. The central hospital is responsible for selecting and integrating the participants in the research, which will provide diabetes measurements to build the training data set. They will follow the platform results, validate the clinical outcomes, and disseminate the system to their patients and healthcare professionals. The university research team is responsible for developing the AI and ML models used by the application to suggest the types of exercises people should do. These models use open and patient-specific data, which may require interoperability between different healthcare actors (e.g., hospitals, private clinics, and gyms). The IT provider is responsible for wearable technology infrastructure, mobile app, and integrating the ML and AI models. Moreover, based on the project results, the IT provider is studying the possibility of offering ML as a service to healthcare facilities. Lastly, the IT provider acts as the digital business ecosystem leader. The product is classified as a moderate-high risk (IIB) medical device, according to the European Union's Medical Device Coordination Group (2021).

The platform team is identifying and developing a set of data policies (e.g., data security policies, data quality policies). Several workshops were held to obtain feedback from platform stakeholders about data policy creation. However, we confirmed the literature findings about the lack of awareness of the different data policy types, which is insufficient for effective data governance. Moreover, the IT provider is not interested in generalizing healthcare-related data policies since its portfolio of IT products is vast and data policies vary significantly (e.g., routers need specific data policies not specifically concerned with privacy).

Our first artifact aimed to guide the business ecosystem members in the lifecycle of data policy creation. A Plan – Do – Check -Act (PDCA) -inspired approach (Deming & Gogue, 1988; Loshin, 2010; Stewart Jr et al., 2007) is used to manage the data policies, including (1) *Data Policy Creation*, (2) *Data Policy Implementation*, (3) *Data Policy Monitoring*, and (4) *Data Policy Review*. Then, we gathered inspiration from the IS literature on how to guide each step. Figure 1 introduces the proposed approach for data policy continuous improvement.

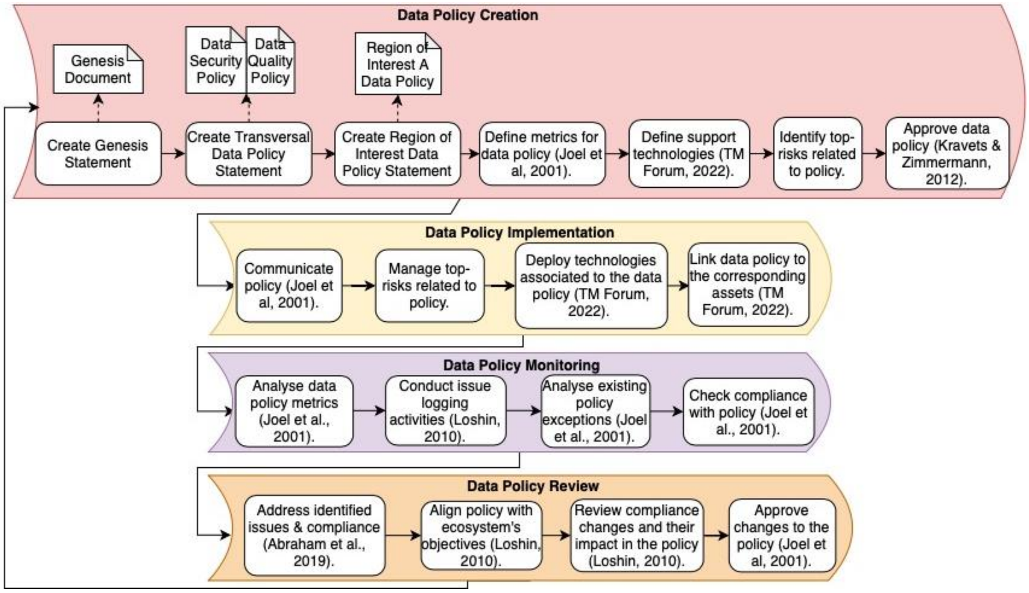
Data Policy Creation
Genesis Document
Data Security Policy
Data Quality Policy
Region of Interest A Data Policy
Create Genesis Statement
Create Transversal Data Policy Statement
Create Region of Interest Data Policy Statement
Define metrics for data policy (Joel et al, 2001).
Define support technologies (TM Forum, 2022).
Identify top-risks related to policy.
Approve data policy (Kravets & Zimmermann, 2012).
Data Policy Implementation
Communicate policy (Joel et al, 2001).
Manage top-risks related to policy.
Deploy technologies associated to the data policy (TM Forum, 2022).
Link data policy to the corresponding assets (TM Forum, 2022).
Data Policy Monitoring
Analyse data policy metrics (Joel et al., 2001).
Conduct issue logging activities (Loshin, 2010).
Analyse existing policy exceptions (Joel et al., 2001).
Check compliance with policy (Joel et al., 2001).
Data Policy Review
Address identified issues & compliance (Abraham et al., 2019).
Align policy with ecosystem's objectives (Loshin, 2010).
Review compliance changes and their impact in the policy (Loshin, 2010).
Approve changes to the policy (Joel et al, 2001).

Figure 1. Data Policy Management in Digital Business Ecosystems

The first phase presented in Figure 1 is the *Data Policy Creation*, following the proposal of Joel et al. (2001), including the assignment of roles and responsibilities, the statement of purpose, the statement, the scope, the monitoring mechanisms, the implementation plan, and the review process. Moreover, metrics (Loshin & Reifer, 2018) and support technologies (TM Forum, 2022) for each data policy are defined. A risk assessment is conducted regarding the elements that can affect the data policy accomplishment (e.g., cyberattacks). Finally, the data policy is approved (Kravets & Zimmermann, 2012).

The *Data Policy Implementation* stage follows in the sequence. The data policies must be adequately communicated to the organization's members (Joel et al., 2001). Next, the identified risks are managed, followed by the deployment of tools to support data policies (e.g., compliance monitoring and enforcement tools) (TM Forum, 2022). Lastly, there is the need to link the data policies with the respective data assets and departments according to the defined scope (TM Forum, 2022).

The *Data Policy Monitoring* stage focuses on monitoring the implementation and compliance with the data policies (Loshin & Reifer, 2018). Monitoring elements (e.g., monitoring processes, supervisors, tools) are put in place to monitor compliance with the internal and external data policies (TM Forum, 2022) based on defined metrics (e.g., number of exceptions) (Joel et al., 2001). Issue logging activities retrieve data across the activities (Loshin, 2010). The exceptions to the data policy (e.g., policies that do not apply in the case of a cyberattack to the organization) and the conditions under which they apply must be monitored to avoid policy exception misapplication (Joel et al., 2001).

The *Data Policy Review* stage focuses on the processes to periodically review and update the existing data policies (Alhassan et al., 2019) to avoid them becoming obsolete (Joel et al., 2005). The alignment with the ecosystem's and partner's objectives should be considered in the review process (Loshin, 2010). Compliance changes and updates should be monitored to understand their impact on the organizations (Abraham et al., 2019). The reviewed policies are discussed and approved (Joel et al., 2001). Finally, the data policy changes must be communicated and disseminated (Stanford University, 2011).

The first DSR results offered a framework for continuously developing data policies, which is essential because these statements are dynamic and vary according to the ecosystem members' regulations or even strategic decisions. Nevertheless, it was insufficient to answer other critical dilemmas of the project participants for policy creation (the first stage presented in Figure 1): *"We will have a single data policy document or several documents?"; "If we have different policy types how do we ensure that each organization/department knows exactly what they have to follow?"; "How to eliminate inconsistencies and ensure that all statements are aligned?"*. Therefore, we proposed three distinct levels for a data policy structure: Genesis, Transversal, and Region of Interest data policies. The Genesis policy is a high-level organizational guideline from which the Transversal data policies will be derived. Its role is like a quality manual in the popular ISO 9001 standard providing high-level policies to the organization, the data policy strategy, and the types of data policies needed. Transversal data policies, on the other hand, define the data policies applicable to all departments and projects of each organization of the business ecosystem (may vary in each organization). However, we found the need to include an additional level of detail: the Region of Interest: may be an element (e.g., a new product to be developed, the integration in a digital business ecosystem, or a new project) that requires the development of specifically tailored data policies. The neuroimaging field often refers to the term "Region of Interest" which describes image areas relevant to understanding a specific phenomenon (Brett et al., 2002), and the concept was already adopted in IS research (Portugal & Barata, 2021). For a specific Region of Interest, it is possible to derive tailored data policies only applicable within that scope. Figure 2 introduces additional details for the Data Policy Creation stage, specifying the steps to create the Genesis, Transversal, and Region of Interest data policies.

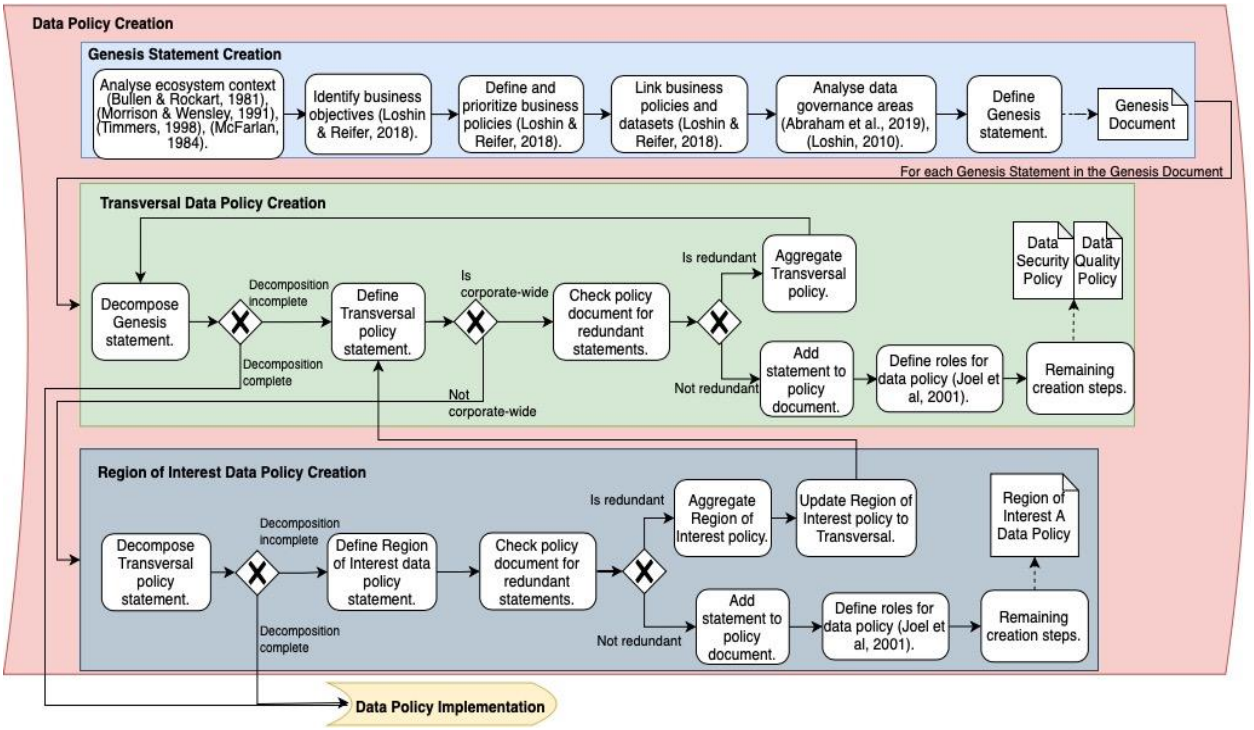
Data Policy Creation
Genesis Statement Creation
Analyse ecosystem context (Bullen & Rockart, 1981), (Morrison & Wensley, 1991), (Timmers, 1998), (McFarlan, 1984).
Identify business objectives (Loshin & Reifer, 2018).
Define and prioritize business policies (Loshin & Reifer, 2018).
Link business policies and datasets (Loshin & Reifer, 2018).
Analyse data governance areas (Abraham et al., 2019), (Loshin, 2010).
Define Genesis statement.
Genesis Document
For each Genesis Statement in the Genesis Document
Transversal Data Policy Creation
Decompose Genesis statement.
Decomposition incomplete
Decomposition complete
Define Transversal policy statement.
Is corporate-wide
Not corporate-wide
Check policy document for redundant statements.
Is redundant
Aggregate Transversal policy.
Not redundant
Add statement to policy document.
Define roles for data policy (Joel et al, 2001).
Remaining creation steps.
Data Security Policy
Data Quality Policy
Region of Interest Data Policy Creation
Decompose Transversal policy statement.
Decomposition incomplete
Decomposition complete
Define Region of Interest data policy statement.
Check policy document for redundant statements.
Is redundant
Aggregate Region of Interest policy.
Update Region of Interest policy to Transversal.
Not redundant
Add statement to policy document.
Define roles for data policy (Joel et al, 2001).
Remaining creation steps.
Region of Interest A Data Policy
Data Policy Implementation

Figure 2. Data Policy Creation Stage in Detail

The process starts with the ecosystem's context and objectives analysis. The Critical Success Factors (Bullen & Rockart, 1981) method can provide support in analyzing the environmental (e.g., data regulations), strategic (e.g., development of a strong business case), industry-specific (e.g., defining clear software requirements and specifications), and temporal factors (e.g., development of data-based solutions for COVID-19) that may influence the organization's data policies. The ecosystems' stakeholder analysis (Bryson, 2004) can provide support in defining the roles and responsibilities of the data policy (e.g., policy approval, policy review, policy scope). The Boston Matrix (Morrison & Wensley, 1991) can provide insights into the specific data policies that must be developed to address the market gaps. McFarlan's Strategic Grid (1984) can provide support in identifying the organization's data policies in the (1) support quadrant (e.g., define an open data policy), (2) turnaround quadrant (e.g., define a data policy for AI and ML), (3) factory quadrant (e.g., define a data security policy), or (4) strategic quadrant (e.g., define a data privacy policy), by considering the data policies for each of the defined quadrants. The ecosystem's business objectives and policy identification follows (Loshin & Reifer, 2018). A linkage between these policies and the available datasets is subsequently established (Loshin & Reifer, 2018). At this stage, the organization will have created the *Genesis* document, which includes a set of Genesis statements to be considered as priorities regarding each data governance dimension (e.g., data quality, data security, and data architecture).

The *Transversal Data Policy Creation* starts with the decomposition of a Genesis statement, in which we verify how the statement is related to each data governance dimension. The data policy documents are checked for redundancies (e.g., if a similar policy already exists) if the Transversal data policy statement is corporate-wide. If no similar statement is detected, it is added to the policy document. If a similar Transversal statement already exists, the corresponding Transversal policies are merged. If the Transversal data policy is not corporate-wide, there is a need to define a Region of Interest data policy. At the end of this stage, the organization will have built a data policy document for each data governance dimension (e.g., data quality policy document, data security policy document).

The *Region of Interest Policy Creation* follows a similar process to the one described in the last paragraphs. Whenever there is a new Region of Interest (e.g., a new product to be developed or a new ecosystem integration), there is the need to reflect on the existing Transversal data policies that may require some tailoring to be applied to that specific region. If a similar policy already exists in other Regions of Interest policies documents, the policy can be considered a possible Transversal data policy, and its hierarchy can be updated. At this stage, the organization will have defined a document for each Region of Interest, that complements the Transversal Data Policies but only applies to specific stakeholders (e.g., healthcare products data policy document).

Demonstration

We deployed the data policy management approach in the digital business ecosystem to demonstrate our artifact. We started with a context analysis, in which we developed a stakeholder map to model the participants, their roles, and interactions (the blue stage in Figure 2). The partners defined as a business objective the goal to develop high-quality products. Therefore, high-quality healthcare data standards must be guaranteed to develop an accurate, unbiased, and fair algorithm that provides correct advice to patients. Figure 3 depicts an extract of the data policy's structure and links to the business ecosystem.

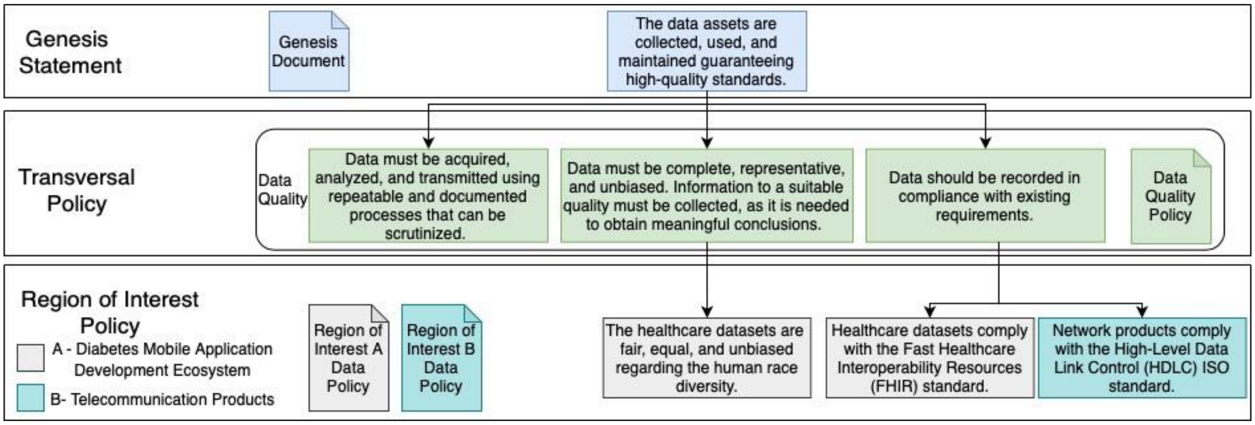
Genesis Statement
Genesis Document
The data assets are collected, used, and maintained guaranteeing high-quality standards.
Transversal Policy
Data Quality
Data must be acquired, analyzed, and transmitted using repeatable and documented processes that can be scrutinized.
Data must be complete, representative, and unbiased. Information to a suitable quality must be collected, as it is needed to obtain meaningful conclusions.
Data should be recorded in compliance with existing requirements.
Data Quality Policy
Region of Interest Policy
A - Diabetes Mobile Application Development Ecosystem
B- Telecommunication Products
Region of Interest A Data Policy
Region of Interest B Data Policy
The healthcare datasets are fair, equal, and unbiased regarding the human race diversity.
Healthcare datasets comply with the Fast Healthcare Interoperability Resources (FHIR) standard.
Network products comply with the High-Level Data Link Control (HDLC) ISO standard.

Figure 3. Data Policy Management Demonstration

We have identified the Genesis statement that is part of the Genesis document as follows: "The data assets are collected, used, and maintained guaranteeing high-quality standards". Based on this Genesis statement, we performed the steps for developing the Transversal and Region of Interest data policies. As an example, we obtained three Transversal data policies deriving the Genesis Statement: (1) "Data must be acquired, analyzed, and transmitted using repeatable and documented processes that can be scrutinized", (2) "Data must be complete, representative, and unbiased. Information to a suitable quality must be collected, as it is needed to obtain meaningful conclusions", and (3) "Data should be recorded in compliance with existing requirements" (the green stage in Figure 2). These policies were incorporated in the Data Quality Policy document.

Considering the specificities of healthcare products, we reflected on these Transversal Statements, that translated to Region Specific data policies. We obtained "The datasets are fair, equal, and unbiased regarding the human race diversity" and "Healthcare datasets comply with the Fast Healthcare Interoperability Resources (FHIR) standard" (the dark-grey stage in Figure 2). A distinct Region of Interest data policy was previously defined for telecommunication products: "Network products comply with the High-Level Data Link Control (HDLC) ISO standard" (the dark-grey stage in Figure 2). For each Region of Interest, we specify a distinct data policy document.

The identified data policy statements can be shared with the ecosystem's members in a shared repository. For example, the hospital responsible for conducting the data collection processes will have to consider the creation of an unbiased, fair, and equal dataset that reflects the human population diversity representation. Based on this statement, the hospital can develop data quality, data provenance, and metadata internal policies to ensure these characteristics are met (required by the University to obtain valid models). All the digital business ecosystem's members must comply with the FHIR healthcare standard when using and exchanging data. Therefore, all the organizations will need to consider the development of internal Transversal and Region of Interest data policies that guarantee the accomplishment of these standards.

Finally, the team defined metrics, risks, and technologies related to each of the data policies (salmon in Figure 1). As an example, we defined the "The healthcare datasets are fair, equal, and unbiased regarding the human race diversity" policy metrics (e.g., the percentage of datasets that are documented, verified for fairness requirements, and compliant with sector's standards), risks (e.g., diabetes values sensor malfunction), and support technologies (e.g., data quality check tool).

Discussion

The artifacts developed in this DSR provide a structured approach to data policy management in digital business ecosystems. The top-down approach can be used to develop data policies for the organizations, considering the Genesis statements, Transversal data policies, and Region of Interest data policies. According to the evidence collected in this case, business ecosystems must define a strategy and a structure for data policies. It is not viable to create ad-hoc policies merged into long documents (difficult to read and evaluate compliance). It is essential to create a logic for coherent statements at different levels (transversal/specific to a region of interest with particular characteristics that may require unique policies).

The approach can support the practitioners in performing data policy review that considers ecosystem environmental dynamics (e.g., new regulations, market shifts, innovative technologies). Based on the reviewed policies, it is possible to derive the specific standards, rules, and processes that define how the ecosystem and its members will achieve and comply with its data policies. Based on the Region of Interest data policies, it is possible to tailor and derive the specific standards, processes, and rules for the activities that are conducted within that working team. Environmental changes (e.g., new data-related regulations, new available technologies, integration in a digital business ecosystem) can be handled by making necessary data policy adjustments that will affect the related standards and processes.

The developed approach can contribute to creating contractual agreements between the organizations (e.g., data sharing agreements, service level agreements) regarding the existence of a Region of Interest that involves participation in a digital business ecosystem. In this case, the identified Region of Interest data policies can apply to the remaining participants, which must obey them through defined contracts that reflect the policy content. Moreover, other organizations can adopt the data policy structure to use a bottom-up to define their data policies that must be considered to comply with the established digital business ecosystem contractual conditions. For this purpose, a shared data policy repository can be created to present this information to the relevant partners. The developed data policies document can provide internal audit mechanisms to verify the alignment between the policies and the defined standards, processes, and rules. Moreover, it can be used to disclose the organizational data policies to audits and demonstrate the data policy cycle evidence.

Conclusion

This paper reports a DSR cycle to (1) create an approach for data policy management in digital business ecosystems, (2) supported by a logical structure of policies, and (3) perform a demonstration of the artifacts in a real case. The importance of data to create innovative products is increasing. Current approaches to data policy management, as usually included in the company websites, are insufficient in decentralized contexts of data use. The GDPR revealed the importance of data privacy, but, as we showed in this case, organizations must implement many other critical types of data policies.

There are also limitations at this stage that need to be stated. First, although we have identified the steps and the tools that support data policy management in digital business ecosystems, we have developed our case in a single case. Other steps and tools may be developed with insights from practitioners in other organizations. Second, in other cases, the data policy management process may need to be tailored to the organization's context, considering market volatility, technical requisites, and compliance requirements. Moreover, the data policies specification can change according to the data domain that is being targeted (e.g., the roles and responsibilities for data privacy policies are different from the ones required for data quality). Third, the organization that participated in our research does not represent the entire industry and the data policies were not yet assessed by organizations external to the business ecosystem, like insurance companies or assessors. Lastly, we have not conducted a formal evaluation of our artefacts. Upcoming stages can focus on defining data quality criteria to assess our approach in a real case.

Future work can extend our research for data policy management in digital business ecosystems. First, the approach can be extended to include a stage that targets the development of the specific processes by deriving the data policies. Second, the development of a tool to support data policy management across the digital business ecosystem. This tool could aggregate the ecosystem's member's data policies in a shared repository and allow the decentralized monitorization of policy compliance across the partners. Moreover, this tool could act as an intermediate platform to communicate the data policy review process results, including details on the elements that may affect the data policies of the remaining ecosystem's members.

Acknowledgements

This work is funded by the project POWER (grant number POCI-01-0247-FEDER-070365), co-financed by the European Regional Development Fund (FEDER), through Portugal 2020 (PT2020), and by the Competitiveness and Internationalization Operational Programme (COMPETE 2020). It is also co-funded by the FCT - Foundation for Science and Technology, I.P./MCTES through national funds (PIDDAC), within the scope of CISUC R&D Unit - UIDB/00326/2020 or project code UIDP/00326/2020

REFERENCES

Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. International Journal of Information Management, 49(January), 424–438.

Alhassan, I., Sammon, D., & Daly, M. (2019). Critical Success Factors for Data Governance: A Theory Building Approach. Information Systems Management, 36(2), 98–110.

Brett, M., Anton, J.-L., Valabregue, R., Poline, J.-B., & others. (2002). Region of interest analysis using an SPM toolbox. 8th International Conference on Functional Mapping of the Human Brain, 16(2), 497.

Bryson, J. M. (2004). What to do when stakeholders matter: Stakeholder Identification and analysis techniques. Public Management Review, 6(1), 21–53.

Bullen, C. V., & Rockart, J. F. (1981). A primer on critical success factors. Working Papers, 69, 1–64.

Data Management Association. (2017). DAMA-DMBOK : data management body of knowledge. Technics Publications, LLC.

Davidson, E., Wessel, L., Winter, J. S., & Winter, S. (2023). Future directions for scholarship on data governance, digital innovation, and grand challenges. Information and Organization, 33(1), 100454.

De Prieelle, F., De Reuver, M., & Rezaei, J. (2020). The Role of Ecosystem Data Governance in Adoption of Data Platforms by Internet-of-Things Data Providers: Case of Dutch Horticulture Industry. IEEE Transactions on Engineering Management, 69(4), 1–11.

Deming, W. E., & Gogue, J.-M. (1988). Qualité: la révolution du management. Economica.

Espinosa, J. A., Kaisler, S., Armour, F., & Money, W. H. (2019). Big data redux: New issues and challenges moving forward. HICSS 2019 Proceedings, 2019-Janua, 1065–1074.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. MIS Quarterly: Management Information Systems, 28(1), 75–105.

Jagals, M., & Karger, E. (2021). Inter-organisational data governance: A literature review. Twenty-Ninth European Conference on Information Systems, June, 1–19. https://aisel.aisnet.org/ecis2021_rp/57

Joel, B., Sunps, W., Global, S. M., Practice, S., Martin, C. R., & Java, S. (2001). Data Security Policy - Structure and Guidelines. Structure, December.

Joel, B., Sunps, W., Global, S. M., Practice, S., Martin, C. R., & Java, S. (2005). Developing a Security Policy. December, 1–7.

Khatri, V., & Brown, C. V. (2010). Designing data governance. Communications of the ACM, 53(1), 148–152.

Kraemer, K. L., Gibbs, J., & Dedrick, J. (2002). Environment and Policy Factors Shaping E-Commerce Diffusion: a Cross-Country Comparison. ICIS 2002 Proceedings, 19(1), 325–335.

Kravets, J., & Zimmermann, K. (2012). Inter-organizational Information Alignment : A Conceptual Model of Structure and Governance for Cooperations Cooperations.

Lee, S. U., Zhu, L., & Jeffery, R. (2017). Data governance for platform ecosystems: Critical factors and the state of practice. PACIS 2017 Proceedings.

Lee, S. U., Zhu, L., & Jeffery, R. (2019). Data governance decisions for platform ecosystems. Proceedings of the Annual Hawaii International Conference on System Sciences, 2019-Janua, 6377–6386.

Lis, D., & Otto, B. (2020). Data governance in data ecosystems - Insights from organizations. 26th Americas Conference on Information Systems, AMCIS 2020, 0–10.

Lis, D., & Otto, B. (2021). Towards a taxonomy of ecosystem data governance. Proceedings of the Annual Hawaii International Conference on System Sciences, 2020-Janua, 6067–6076.

Loshin, D. (2010). Operationalizing Data Governance through Data Policy Management. Kalido, 1(301), 1–11.

Loshin, D., & Reifer, A. (2018). Modern Data Governance : Strategies for Data Policies that Stick. January. https://www.eckerson.com/articles/modern-data-governance

McFarlan, F. W. (1984). Information technology changes the way you compete. In Harvard Business Review (Vol. 62, Issue 3). Harvard Business Review, Reprint Service.

Medical Device Coordination Group. (2021). Guidance on classification of medical devices. Mdcg 2021-24, 1–11. https://health.ec.europa.eu/latest-updates/updated-joint-implementation-plan-actions-considered-necessary-ensure-sound-functioning-new-2022-02-07_en

Morrison, A., & Wensley, R. (1991). Boxing up or boxed in?: A short history of the boston consulting group share/growth matrix. Journal of Marketing Management, 7(2), 105–129.

Nachira, F., Dini, P., & Nicolai, A. (2007). A network of digital business ecosystems for Europe: roots, processes and perspectives. European Commission, Information Society and Media, 106(Com), 5–24.

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.

Portugal, T., & Barata, J. (2021). Enterprise architecture erosion: A definition and research framework. AMCIS 2021 Proceedings.

Scholz, N., Wieland, J., & Schäffer, T. (2022). Towards a Framework for Enterprise & Platform Ecosystem Data Governance. Twenty-Eighth Americas Conference on Information Systems, 0–10.

Schreieck, M., Wiesche, M., & Krcmar, H. (2016). Design and governance of platform ecosystems - Key concepts and issues for future research. ECIS 2016 Proceedings, 16, 12–15.

Stanford University. (2011). Stanford Data Governance Maturity Model. https://www.lightsondata.com/data-governance-maturity-models-stanford/

Stewart Jr, J., Hedge, D., & Lester, J. P. (2007). Public policy: An evolutionary approach. Nelson Education.

TM Forum. (2022). TM Forum Data Governance Guidebook. https://www.tmforum.org/resources/standard/gb1023-data-governance-guide-book-v3-0-0/

Zuiderwijk, A., & Janssen, M. (2014). Open data policies, their implementation and impact: A framework for comparison. Government Information Quarterly, 31(1), 17–29.