Data Network Game: Enabling Collaboration via Data Mesh

Bove Lucaleonardo[1,2][a], Totaro Nicolò G.[1,2][b] and Massimiliano Gervasi[1,2][c]
[1]Department of Engineering for Innovation, University of Salento, Lecce, Italy
[2]Centre for Applied Mathematics and Physics for Industry (CAMPI), University of Salento, Lecce, Italy

Keywords: Data Sharing, Big Data Analytics, Data Value, Data Network, Cooperative Game Theory, Data Mesh.

Abstract: Organizations aim to transform raw data into valuable insights using advanced analytical methods. Since data can be replicated and shared, multiple actors can simultaneously utilize the same information. This study presents the *Data Network*, a theoretical framework representing potential collaborations among organizations sharing data in large-scale big data projects, using *Data Mesh* as a supporting architecture. The Data Network Game (DNG) extends this model by applying game theory to analyze inter-organizational collaborations, incorporating market-imposed constraints that limit compatibility. Various scenarios, defined by distinct benefit and cost functions, are explored to understand their impact on coalition formation and market dynamics. A simplified theoretical example shows how coalitions can achieve greater value through collaboration than by acting independently. This model serves as a practical tool for assessing the trade-offs of cooperation and offers insights into managing emerging data-driven markets.

1 INTRODUCTION

Data are raw informational assets that organizations can transform into value to enhance business process knowledge and support strategic decision-making (Ylijoki and Porras, 2019; Ramchand and Mahmood, 2022; Wu et al., 2022; Gervasi et al., 2023b; Angelelli et al., 2024b; Catalano et al., 2024; Corallo et al., 2023). Data that exceed specific thresholds in characteristics such as velocity, variety, and volume are designated as big data (Laney, 2001). Organizations extract value from big data primarily through big data analytics (Gervasi et al., 2023b; Corallo et al., 2023; Catalano et al., 2024), typically within initiatives described in the literature as *big data initiatives* (Braganza et al., 2017) or *big data projects* (Tiefenbacher and Olbrich, 2015; Huang et al., 2015; Louati and Mekadmi, 2019; Grander et al., 2022). In fact, data can be considered the fundamental resource for big data projects and, consequently, for big data analytics (Ylijoki and Porras, 2019; Gupta and George, 2016; Gervasi et al., 2023b; Catalano et al., 2024). However, data and the information derived from it differ from traditional resources as they are non-exclusive, allowing multiple actors to simultaneously utilize them (Hensler and Huq, 2005). Consequently, organizations might collaborate by sharing their data (Bertsekas and Gallager, 2021) to generate greater value than each could achieve independently (Dong and Yang, 2020). For this reason, data sharing across organizations must be well-regulated, requiring suitable data architectures that facilitate multi-actor data sharing, such as *Data Vault* for data integration (Lindstedt et al., 2009) or *Data Mesh* (Dehghani, 2022). Moreover, data may be associated with a price that organizations would need to pay in order to access and utilize it. In a dynamic context, this price fluctuates based on supply and demand within what is defined as a *Data Market* (Koutroumpis et al., 2020).

To model collaboration among organizations wishing to share data for common objectives, it is necessary to calculate potential incentives arising from such collaborations. This involves defining different types of values associated with data, such as the potential value of data (Angelelli et al., 2024b; Corallo et al., 2023), the *extractable value* of data, and the related *business value* (Gervasi et al., 2023b; Angelelli et al., 2024b), while also understanding how these values might change through sharing and collaboration among multiple organizations. In this study, we define a *Data Network* as the complex structure of potential collaborations between organizations within the data market. To formalize potential coalitions among organizations and analyze the various interactions within the Data Network,

[a] https://orcid.org/0009-0000-1927-5037
[b] https://orcid.org/0009-0001-0845-538X
[c] https://orcid.org/0000-0001-6114-1454

Bove, L., Totaro, N. G. and Gervasi, M.
Data Network Game: Enabling Collaboration via Data Mesh.
DOI: 10.5220/0013283800003944

In *Proceedings of the 10th International Conference on Internet of Things, Big Data and Security (IoTBDS 2025)*, pages 81-92
ISBN: 978-989-758-750-4; ISSN: 2184-4976

we apply formalism and results from game theory (Badewitz et al., 2020). Similar approaches have been used for data markets (Agarwal et al., 2019) to model data pricing (Bi et al., 2024; Liang et al., 2018), trading, and protection (Liang et al., 2018).

Finally, the proposed model introduces constraints that limit coalitions between organizations, addressing privacy, security (Van Panhuis et al., 2014; Angelelli et al., 2024a; Gervasi et al., 2023a), data valuation (Coyle and Manley, 2024; Angelelli et al., 2024b), regulatory (Kathuria and Globocnik, 2020; Graef et al., 2019), and market concerns.

The structure of the paper is as follows: The importance of collaboration in big data projects is covered in Section 2, which also highlights the choice of data mesh as an enabling architecture for inter-organizational data sharing. The Data Network Game concept is theoretically formalized in Section 3. Section 4 offers an in-depth analysis of the model's components, underlying assumptions, and potential implications. Lastly, a real-world examples from the healthcare industries is presented in Section 5, which also examine possible model expansions.

2 THE IMPORTANCE OF COOPERATION IN BIG DATA PROJECTS

According to the various analogies in the literature, we consider data as a raw resource with a potential extractable value, such as metallic ores or oil (Ackoff, 1989; Ylijoki and Porras, 2019; Saltz, 2015). Thus, it is possible to associate data with the same characteristics as other resources, following the VRIO (*Valuable, Rare, costly to Imitate, Organizationally embedded*) model (Barney, 1991). Like traditional resources, even data are associated with *facilitating factors* that enable a company to establish a competitive advantage, such as *exclusive access to data*, and *exploitative access to data* (Fast et al., 2021). Moreover, while data have all the characteristics of the VRIO model, the ease of duplicating and sharing it with other actors is an atypical characteristic compared with other resources, which often enjoy exclusivity properties (Gervasi et al., 2023a). Beyond data, the knowledge they generate can be consumed by multiple actors, not only by its creator (Hensler and Huq, 2005). This feature opens up positive considerations regarding mutual data sharing for greater value creation. Some challenges, such as the unavailability of specific data or the lack of high-quality data, could be overcome by architectures that can accommodate, manage, and make available a wide variety of data from diverse sources and organizations (Chen et al., 2017; Chen et al., 2014; Gervasi et al., 2023a). Sharing expertise not only within firms but also across firms, clearly under assumptions of complementarity and connectivity, could lead to an evolution of the classic big data value chain towards a value network triggered by big data (Wu et al., 2022). System Theory, and in particular the concept of *synergy*, explains how well-orchestrated and shared resource utilization can generate greater value than the sum of the individuals; namely, the value is a *super-additive set function*:

$$Value(d_1 \cup d_2) \geq Value(d_1) + Value(d_2), \quad (1)$$

with d_1 and d_2 being sets of shareable data/resources (Dong and Yang, 2020). These aspects have significant and tangible effects on firm performance (Dong and Yang, 2020; Tanriverdi and Venkatraman, 2005). The proliferation of interactions and the growing importance of stakeholders become fundamental in the distributed co-creation of value, accompanied by the formation of ad-hoc ecosystems (Del Vecchio et al., 2018; Malthouse et al., 2019; Roos, 2018). This leads to a network of collaborations that produces value for the involved entities. The definition of value network we provide aims at formalizing the super-additivity property (1) into a broader context of interdependence and symbiosis among different actors. This collaboration among the network's members fosters additional value that could not be derived from the mere effort of organizations if they acted alone (Malthouse et al., 2019; Roos, 2018). In practical terms, organizations require the adoption of new data architectures that enable such collaboration to establish a data-driven strategy that can be implemented by multiple actors.

2.1 Data Mesh: A New Paradigm for Data Sharing

The development of new data architectures for the sharing and management of data across multiple organizations is essential to preserve their potential value, facilitate sharing among diverse actors, ensure privacy and security, and support analytical processes (Priebe et al., 2021; Hechler et al., 2023). In a data-driven market, one of the key challenges is therefore to preserve data quality without reducing its potential (Reggio and Astesiano, 2020), which requires selecting the most suitable data architecture for specific goals. We have identified and adopted the data mesh as an enabling architecture for multi-actor data sharing within the Data Network and the framework we define as the Data Network Game, as it decentralizes data ownership, enhances scalability, and ensures domain-driven accountability. This architecture is designed for sharing data between domains that own it, unlike other architectures, such as, data vault, which focuses on historical data integration (Lindstedt et al., 2009), or data fabric, which emphasizes centralized orchestration of distributed data (Sharma et al., 2023).

The data mesh approach is particularly effective in contexts of data sharing and reuse (Dehghani, 2022; Azeroual and Nacheva, 2023), and it is based on four foundational principles:

- **Domain Ownership:** each dataset is labeled with the relevant information, including who is responsible for its content.
- **Data as a Product:** data are treated as products, which require investment and whose value is directly linked to their quality.
- **Self-Serve Data Platform:** data are stored on a user-friendly platform, allowing each actor to locate accessible data that is ready for use without requiring pre-analytics processing.
- **Federated Computational Governance:** decentralized approach combining domain autonomy with global standards, ensuring data security, compliance, quality, and interoperability.

The principle of domain ownership aims to decentralize the ownership of analytical data by assigning it to business domains that are closest to the data sources or primary consumers, thereby segmenting the data and managing its life cycle within each domain (Dehghani, 2022).

Assuming the adoption of a cross-domain data mesh architecture for inter-organizational data sharing, the big data value chain demands a fundamental redefinition. The linear value chain can evolve into a graph structure by reusing the same data product across multiple analytics processes (Gervasi et al., 2023a).

2.2 Cooperative Game Theory for Data Sharing

In the literature, among related models, we find examples such as the *Data Provision Game* (Badewitz et al., 2020) and the *Data Marketplace* (Agarwal et al., 2019), which focus on data pricing and revenue sharing. Notably, the Data Network Game moves beyond the concept of directly assigning a specific price to a data domain. Instead, the model proposes that individual organizations select others with which to share their data, thereby recognizing value even potential value in such data. The selection of domains, and consequently of organizations, becomes an integral part of the decision-making process, serving as a guarantee for security, reliability, and, above all, data quality. In this context, adopting a data mesh as an enabling architecture for multi-actor data sharing unlike the traditional big data value chain (Badewitz et al., 2020; Gervasi et al., 2023b) represents an innovative approach.

The total benefit generated by a coalition depends on the combined resources and efforts of its members. In the context of data sharing, these benefits often display increasing returns to scale, meaning that the value derived from the shared data grows more than proportionally as the amount or diversity of data shared increases (Konsynski and McFarlan, 1990). A key question in this context is how to fairly distribute the collective benefits among the members of a coalition in a way that encourages active participation. Various methods for distributing payoffs are found in cooperative game theory, such as in profit-sharing games (Kleinberg and Oren, 2022; Bilò et al., 2023b; Bilò et al., 2023a), where each player (e.g., a domain) selects a resource (e.g., a coalition) to maximize their individual payoff. Our model adopts a mechanism where each agent's payoff is proportional to their data contribution. Specifically, each agent's reward is calculated based on the relative value of their data compared to the total data value within the coalition. This proportional distribution ensures that agents are compensated according to the value they add to the coalition, motivating them to share valuable data and engage in collaborative efforts. However, agents may face certain incompatibility constraints that prevent them from collaborating with others. These constraints could stem from legal restrictions, privacy concerns, or competitive interests (Myerson, 1980). Such limitations reduce the pool of feasible coalitions and must be taken into account when modeling coalition formation. For a coalition structure to be sustainable, it must be stable, meaning no agent or group of agents has any incentive to leave and form a new coalition.

3 DATA NETWORK GAME

The *Data Network Game* is a conceptual framework designed to represent collaborations among organizations that share their data. As previously discussed, we posit that the data mesh serves as an enabling data architecture for multi-actor data sharing and management. Within the context of the Data Network Game, the domains of the data mesh are thus considered as players, or “actors”, engaged in a cooperative or competitive game system aimed at maximizing the value derived from data sharing.

In the foundational version of the Data Network Game presented in this study, we assume that each organization owns only one data domain. Consequently, each domain is viewed as representing an individual organization that operates autonomously but is incentivized to collaborate with other domains (organizations) to increase the total value generated. In this model, domain coalitions enable each participant to achieve a potential gain greater than what could be realized independently, reflecting the principles of cooperative game theory. The primary assumptions underlying the model are as follows:

- **Single-Domain Ownership:** each organization is associated with a single domain.
- **Incompatibility Constraints:** the model accounts for market-imposed incompatibility constraints that limit coalition formation between domains. These constraints represent legal, ethical, or competitive limitations, ensuring that the Data Network Game reflects realistic conditions for cooperation.
- **Coalition Formation:** domains are permitted to form coalitions, provided they adhere to incompatibility constraints.
- **Incremental Value Through Coalition:** it is assumed that participation in a coalition generates incremental value for the involved domains, exceeding the sum of values they would achieve independently.

The cost function associated with coalition formation considers configuration costs, data integration expenses, and compliance requirements. The benefit function reflects the added value from data sharing and access to larger and more diverse datasets. Finally, the distribution of gains among coalition members is proportional to each domain’s contribution, fostering balanced and sustainable collaboration. Under these assumptions, the Data Network Game seeks to model a collaborative network of domains that maximizes the value of shared data while respecting market restrictions and promoting fair and dynamic competition among actors.

3.1 Domains and Market Structure

We consider a *market* M as a set of n *organizations*, $M = \{O_1, O_2, \ldots, O_n\}$ where each organization O_i manages a set of *domains* D_i. In the model presented we assume that each organization has only one domain. Consequently, the total set of domains in the market is $D = \{d_1, d_2, \ldots, d_n\}$. Specifically, a *domain* $d_i \in D$ is a business entity within an organization responsible for managing a certain quantity of data, and $v(d_i) \geq 0$ denotes the *value* of the data owned by d_i. Therefore, we identify each d_i with the set of data it manages and $v(d_i)$ with its overall value. Table 1 provides the nomenclature used in the modeling of the Data Network Game for clarity.

The market imposes *incompatibility constraints* (I), where $I \subseteq \{\{d_i, d_j\} : d_i, d_j \in D,\ i \neq j\}$. Whenever $\{d_i, d_j\} \in I$, it means that domains d_i and d_j cannot belong to the same coalition due to legal, ethical, or other competitive considerations. These constraints prevent certain domains from collaborating and must be respected when forming coalitions. In realistic scenarios, constraints that limit the free sharing of data among organizations are imposed to protect competition, consumers, and their personal data (Graef et al., 2019). Although the replicability of data is an incentive for building a confederate, it is clear that multiple incompatibility factors related, for example, to privacy and security must be taken into account (Cavanillas et al., 2016; Gervasi et al., 2023a). Indeed, incompatibility constraints include legal and political issues, such as restrictive policies, data ownership, and privacy protection. Incompatibility constraints derive from rules that are external to organizational choices, as they are the responsibility of public authorities (Kathuria and Globocnik, 2020; Graef et al., 2019), which ensure compliance with laws, both administrative and financial. In the context of collaborative settings, partial information availability about sensitive data is also a source of uncertainty that may influence decision-making regarding security constraints and prioritization of interventions, which requires appropriate methodologies for their evaluation (Angelelli et al., 2024a). Lastly, it is important to highlight that technical, economic, or technological issues due to inefficient data management or lack of resources are not constraints of incompatibility.

3.2 Game Structure

Each domain $d_i \in D$ is treated as a player in the cooperative game. The objective of each domain is to maximize its payoff by joining a coalition that adheres to the market-imposed incompatibility constraints. For this reason, we consider a *feasible coalition structure* $\mathcal{C} = \{C_1, C_2, \ldots, C_k\}$ as a partition of D, where each coalition C_h satisfies the incompatibility constraints (namely, each coalition C_h is *feasible*). Here, k is the total number of coalitions in the structure, while h (with $1 \leq h \leq k$) is simply an *index* denoting the h-th coalition within the partition. So, the *feasible strategy space* S_i of each domain d_i consists of all coalitions $C \subseteq D$ that include d_i and respect the incompatibility constraints:

$$S_i = \{C \subseteq D \mid d_i \in C, \forall d_j \in C\ \{d_i, d_j\} \notin I\}. \quad (2)$$

This set represents all possible coalitions d_i that can feasibly join, given the incompatibility constraints.

Table 1: DNG: Terminological Foundations.

	Description
M	Market: set of organizations
D	Set of domains in M
O_i	i-th organization
d_i	Domain of i-th organization
$v(d_i)$	Value by domain d_i
$\mathcal{C}$	Coalition structure
C_i	i-th coalition
S_i	i-th strategy
$b(C_i)$	Benefit function for coalition C_i
$c(C_i)$	Cost function for coalition C_i
$u_i(C_j)$	Payoff of i-th domain in coalition C_j

3.3 Benefit and Cost Functions

Once some domains are part of the same coalition, we need to define a function capable of expressing the individual return to each individual domain as a function of its own value and that generated by the coalition. This return will depend on the domains of the same coalition. In the first step, we define a *benefit function* $b(C_j)$ for a coalition C_j as:

$$b(C_j) = f\left(\sum_{d_i \in C_j} v(d_i)\right), \quad (3)$$

where $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a non-decreasing convex function with $f(0) = 0$ that depends on the big data characteristics, such as volume, variety, velocity, and other factors (Geerts and O'Leary, 2022). This formalizes the super-additivity property (1) and ensures increasing returns as the quantity of data increases, capturing the power of data value generation. Therefore, the *total benefit* of the market for a coalition structure $\mathcal{C} = \{C_1, C_2, \ldots, C_k\}$ is:

$$\text{SUM}_b(\mathcal{C}) = \sum_{j=1}^{k} b(C_j). \quad (4)$$

This represents the aggregated benefit generated by all coalitions in the market.

Regarding the costs for domains to participate in the coalition, including configuration, data integration, and compliance costs, we define a *cost function* $c(C_j)$ for a coalition C_j as:

$$c(C_j) = c_0 + c_k \cdot |C_j|^\gamma, \quad (5)$$

where c_0 is a fixed cost, c_k is the marginal cost for domains in the coalition C_j, $|C_j|$ denotes the number of domains in coalition C_j, and $\gamma > 0$ is a parameter that adjusts cost growth as a function of coalition size. This captures both the coalition formation costs and the incremental cost of adding new domains to the coalition. The *total cost* of the market for a coalition structure $\mathcal{C} = \{C_1, C_2, \ldots, C_k\}$ is:

$$\text{SUM}_c(\mathcal{C}) = \sum_{j=1}^{k} c(C_j). \quad (6)$$

3.4 Payoff Function and Stability

Considering that each coalition produces a benefit for all the participating domains, it is interesting to determine a feasible way to distribute the benefit among all of them. To take into account the contribution given by each domain d_i in coalition C_j to the total benefit, we propose the following payoff function:

$$u_i(C_j) = w_i(C_j) \cdot [b(C_j) - c(C_j)], \quad (7)$$

where $w_i(C_j) := \frac{v(d_i)}{\sum_{d_h \in C_j} v(d_h)} \in [0, 1]$ is the weight that models the fraction of total value produced by domain d_i in coalition C_j.

The inclusion of such payoff functions defines an interaction among agents that could potentially lead to stable coalition structures, where no player (or group of players) has any incentive to deviate. The resulting strategic interaction can be modeled by resorting to some tools from cooperative game theory. We point out that our model fits in the general structure of hedonic games (Bogomolnaia and Jackson, 2002). Then, we propose to apply to our model two notions of stability widely adopted for the general class of hedonic games: the *core stability* and the *Nash stability*.

Definition 1 (Core-stability)**.** *A coalition structure $\mathcal{C} = \{C_1, \ldots, C_k\}$ is* **core-stable** *if, for any subset $C' \subseteq D$ that respects the incompatibility constraints,* *there exists at least one domain $d_i \in C'$ such that $u_i(C(d_i)) \geq u_i(C')$, where $C(d_i)$ denotes the coalition of C containing d_i.*

A coalition structure is core-stable if, for any possible subset of domains C' that decides to form a separate feasible coalition, there exists at least one domain in C' whose individual payoff would not increase by deviating. This means that the coalition structure is resilient to group deviations because no subset of domains can collectively break away to form a new coalition that would make all its members strictly better off. Therefore, in a core-stable structure, there is no incentive for any group of domains to leave and form a better arrangement.

If the strategic environment among agents is not highly cooperative, it might happen that agents independently choose which coalitions to join. In such a scenario, we can adopt a (often weaker) stability condition by assuming that a coalition structure is stable if no domain can improve its utility through unilateral deviations only (i.e., by moving from its current coalition to another). This form of stability is typically referred to as Nash-stability (or individual stability), and it suggests that, while domains might not be able to prevent group deviations, they have no incentive to leave their coalition individually to join another, as doing so would not improve their utility.

Definition 2 (Nash stability)**.** *A coalition structure $C = \{C_1, \ldots, C_k\}$ is* **Nash-stable** *if, for any domain $d_i \in D$, (i) $u_i(C(d_i)) \geq u_i(C_j \cup \{d_i\})$ holds for any $C_j \in C$ and (ii) $u_i(C(d_i)) \geq u_i(\{d_i\})$, where $C(d_i)$ denotes the coalition of C containing d_i.*

We observe that the first condition of the above definition ensures that no domain has an incentive to leave its current coalition to improve its utility by joining another group (i.e., jumping into another coalition). The second condition states that no domain can improve its utility by forming a coalition where it is the sole participant. This implies that being alone, or acting independently, would not provide a better outcome compared to remaining in the current coalition. Together, these conditions ensure that the coalition structure is stable with respect to both unilateral shifts to other groups and complete isolation.

In the literature related to hedonic games, it has been shown that core-stable or Nash-stable coalition structures do not always exist, and their computation can be intractable (see, for instance, (Woeginger, 2013b; Woeginger, 2013a; Peters and Elkind, 2015; Sung and Dimitrov, 2010)). Nevertheless, investigating the existence and computation of stable coalition structures within our specific model remains worthwhile, as this research could lead to a deeper understanding of cooperative behavior in our strategic environment and its implications for overall system efficiency (e.g., measured by total benefit or total cost).

3.5 Theoretical Case Study

In Table 2 we present a theoretical case in which there are 10 domains $D = \{d_0, d_1, \ldots, d_9\}$ and their incompatibility constraints. For illustrative purposes, the value of each domain, denoted $v(d_i)$, is assigned randomly from a predefined range (e.g., via a uniform distribution) to capture variability in data quality, volume, and synergy potential.

Table 2: The value and the incompatibility constraints of each domain in the market.

Domain	Value	Constraints
d_0	42.33	d_1, d_6, d_7
d_1	32.57	d_0, d_3, d_8
d_2	43.49	d_3, d_9
d_3	51.41	d_1, d_2, d_4, d_5
d_4	70.23	d_3, d_6, d_9
d_5	30.13	d_3, d_7, d_9
d_6	40.22	d_0, d_4, d_7, d_8
d_7	49.40	d_0, d_5, d_6, d_9
d_8	69.85	d_1, d_6
d_9	44.30	d_2, d_4, d_5, d_7

The fixed costs c_0 used in this example are equal to 5, $\gamma = 0.3$, and the marginal costs c_k related to the coalitions that will form were considered as uniformly random amounts in the range $[1, 30]$. The cost function used is $c(C_j) = 5 + c_k \mid C_j \mid^{0.3}$. The selection of $\gamma = 0.3$ ensures scalability across domains, emphasizing shared architectures' efficiency. Its general applicability requires empirical parameter estimation based on defined data and enabling technologies. For the benefit function relative to coalition C_j, we used a non-decreasing and convex function $b(C_j) = \left(\sum_{d_i \in C_j} v(d_i)\right)^{1.2}$. The benefit function $f(x) = x^{1.2}$ satisfies the super-additivity property of data value and grows with the number of organizations in the coalition under ideal conditions. In real-world scenarios, however, a threshold may emerge beyond which the benefits significantly decrease.

Finally, the incentive for each domain $d_i \in D$ to participate in coalitions is determined by the payoff function u_i defined in (7). In Table 3, we report for each organization the respective benefit, cost, and payoff, assuming that each organization is a coalition in its own right.

Table 3: The benefits, costs, and payoffs referred to coalitions formed by a single domain.

Coalition	Benefit	Cost	Payoff
$\{d_0\}$	89.52	21.99	67.53
$\{d_1\}$	65.36	34,70	30.66
$\{d_2\}$	92.50	29.44	63.06
$\{d_3\}$	113.05	21.04	92.01
$\{d_4\}$	164.37	13.45	150.92
$\{d_5\}$	59.54	12.17	47.37
$\{d_6\}$	84.20	31.63	52.57
$\{d_7\}$	107.75	18.59	89.16
$\{d_8\}$	163.30	12.52	150.78
$\{d_9\}$	94.56	17.36	77.20

Given the above parameters, all possible coalition structures were evaluated. Among all admissible coalition structures, a core-stable coalition structure $\mathcal{C}_E = (\{d_2,d_4,d_7,d_8\},\{d_1\},\{d_3,d_6,d_9\},\{d_0,d_5\})$ was identified. In Figure 1, for each domain d_i, the payoff without participating in any coalition is compared with the payoff generated by the same domain within the coalition $\mathcal{C}_E$ that includes it, showing that the payoff is greater when the domain is part of a coalition.

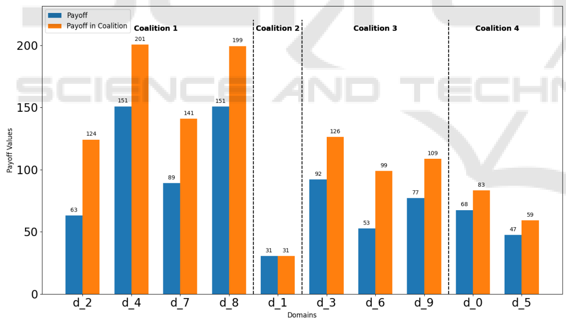

Figure 1: Comparison of Payoffs Inside and Outside Coalitions.

Clearly, the value of data is extractable, quantifiable, and realizable only in the presence of a big data initiative. It is not possible to think of data value as agnostic and not contextualized. In this sense, the intangible nature of data value relates to value perception by those involved in the Data Network; in specific contexts, such perception can be quantified, e.g., through information-theoretic notions (Corallo et al., 2020). Thus, in the example, we assume that each coalition is engaged in a set of big data initiatives where the data they share are recognized as having potential and extractable value.

Figure 2 represents the core-stable coalition structure identified in the example.

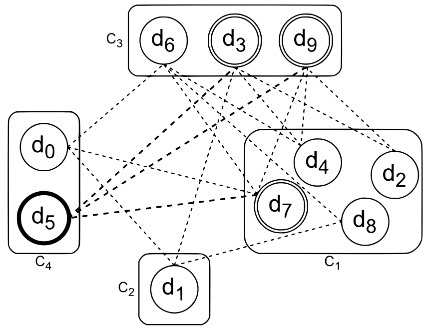

Figure 2: Core-stable coalition structure of the example.

The configuration is $\mathcal{C}_E = \{C_1, C_2, C_3, C_4\}$, where $C_1 = \{d_2,d_4,d_7,d_8\}$, $C_2 = \{d_1\}$, $C_3 = \{d_3,d_6,d_9\}$, and $C_4 = \{d_0,d_5\}$. The dotted lines represent the incompatibility constraints between the various domains. The constraints of domain d_5 have been highlighted as an example.

4 DISCUSSION

The Data Network Game models the dynamics of potential collaborations between organizations interested in data sharing. At the core of this model, the concept of data mesh emerges as a key element for data sharing, promoting domain-oriented management, the use of a self-serve data platform, the adoption of shared governance among all members of the coalition, and the view of data as products (Dehghani, 2022). A potential extension of the model could include the sharing of additional resources. Beyond data, the sharing of technologies (Technology Mesh) could further increase the value extracted from the data, fostering greater interoperability between domains (Gervasi et al., 2023a). In addition to the exchange of digital resources, the model could be extended to include the sharing of human resource skills available to each domain, with a mutual knowledge transfer (Angelelli et al., 2024b) and exchange of abilities.

Within the Data Network Game, there exists a wide range of possible configurations due to the numerous parameters that characterize it, such as the *potential value* attributed to the data, the *extractable or generable value*, the corresponding *business value* (Angelelli et al., 2024b), *benefit* and *cost functions*, and *incompatibility constraints*. Organizations are therefore called upon to analyze various scenarios and assess the utility of collaborations, considering not only specific big data initiatives but also the possible temporal evolution of the coalitions themselves.

Indeed, a coalition should not be understood as a temporary collaboration between organizations with a common and defined goal. Rather, a coalition in the market should be viewed as a stable strategic alliance that exists within the market, aimed at consolidating its position or emerging as a new key player.

In the Data Network Game, each domain aims to maximize its own yield, namely the value generated through the strategic sharing and utilization of data via targeted collaborations with other domains, in line with the principles of hedonic games. For simplicity, the proposed model assumes a one-to-one correspondence between organizations and domains, treating them almost as synonyms. However, it is plausible to extend this representation by assuming that an organization may manage multiple domains (or business units) and consequently participate simultaneously in multiple coalitions. A crucial constraint of the model establishes that each domain can share its data exclusively with one coalition at a time, ensuring the exclusivity of the data within each coalition. However, it is possible for two domains belonging to the same organization to be part of the same coalition. This scenario opens new perspectives for the model, where organizations do not merely optimize the performance of individual domains but aim at maximizing the overall returns derived from all the domains under their control.

Within the Data Network Game, data are conceived as products that organizations seek to enhance and combine strategically. While in the data mesh paradigm the principle of *data as a product* emphasizes data quality and usability, in the Data Network Game the focus shifts to the value generated by the data itself. The value of a domain depends on the intrinsic potential of the data it contains and on the opportunities to integrate them with data from other domains, thereby increasing their overall information value. Therefore, it is essential to formally define a value function associated with the data, considering both its raw form and the transformation processes it may undergo. Additionally, it is necessary to understand how the value of data may evolve over time and in relation to market dynamics (Angelelli et al., 2024b). For example, in the theoretical case study, the value function was not explicitly defined, nor was a temporal dependency considered. Such a dependency is crucial in big data initiatives to model the obsolescence of extracted information in relation to the market and its impact. In this context, domains, based on the value recognized by the market, can be assimilated to *stocks* in a data market (Agarwal et al., 2019). The value of these stocks could fluctuate based on the number of coalitions interested in integrating that domain, and the value attributed to it by those coalitions.

It thus becomes crucial to identify the factors that influence the formation and stability of coalitions so that the potential value extracted from the data can be maintained over time and across multiple big data initiatives. The theoretical example presented was constructed to show how domains benefit more by participating in coalitions $\mathcal{C}_E$ rather than operating independently. An aspect that could be useful to analyze and possibly model concerns the dynamics that could generate market polarization, favoring some players over others. For instance, in a Data Network Game, the market rules might favor coalitions between smaller organizations, but it is also possible that they could favor larger organizations (e.g., *superstar firms* (Fast et al., 2021)). In the latter case, collaboration among market-leading organizations could result in the creation of an almost monopolistic market. However, antitrust authorities are placing an increasing amount of emphasis on preventing the concentration of knowledge and market power in the hands of a few strong corporations, as this could impede innovation and limit consumer choice (Spulber, 2023). Therefore, the market rules, and hence those of the Data Network Game, must be refined to ensure a market that does not polarize and does not favor any particular category of organization in advance.

In the context of the Data Network Game, adhering to strict data security and protection standards plays a fundamental role. Although privacy and security are critical considerations for data sharing, it is essential to integrate concrete methodologies to strengthen the framework of the Data Network Game, such as *federated learning* for model training without sharing raw data and *differential privacy* for confidentiality ensures robust practices and data governance compliance (Faroukhi et al., 2020). It is plausible that such standards could become a key factor in coalition formation, potentially evolving into a new type of incompatibility constraint. This constraint would no longer be imposed solely by market regulators but would arise directly from the organizations themselves, which could set stricter security requirements to access certain collaborations. This behavior could encourage organizations to adopt higher and more restrictive security measures to avoid exclusion from strategic coalitions. The federated computational governance of the data mesh is pivotal in regulating coalitions and their associated constraints, ensuring adherence to shared standards both within coalitions and at the market level. Furthermore, it enables the effective management of control issues, a key challenge in coalition conflicts (Inkpen, 2005; Gervasi et al., 2023a).

5 FUTURE WORK

The Data Network Game presents significant prospects for development, aimed at refining the theoretic model and enhancing its ability to represent complex and evolving contexts. Among the proposed extensions, a notable one is the possibility of allowing each organization to participate in the Data Network through a **plurality of domains**, thus providing a more accurate modeling of the operational reality of large multi-sector organizations. This generalization would enable capturing the diversity and specificity of the benefits arising from interactions between distinct domains. Simultaneously, it becomes necessary to introduce formal constraints regulating the possibility for an organization to affiliate with multiple coalitions simultaneously, specifying the assumptions and operational conditions under which such multiple affiliations are permissible in the market while ensuring consistency in competitive and collaborative dynamics within the model. **Hypergraphs of incentives and constraints**, which each represent the possible advantages of coalition relationships and the restrictions imposed by the market, might be used to expand the model in this way.

Additional potential extensions of our model might result from investigating potentially more **realistic payoff functions**, in which payoffs are established using a network that simulates the efficacy of interactions within the same coalitions. Therefore, investigating new generalizations of the Data Network Game that incorporate graphs or hypergraphs into their payoff structure (e.g., in (Aloisio et al., 2020; Aloisio et al., 2021; Aloisio et al., 2024; Apt et al., 2017; Bilò et al., 2022; Bogomolnaia and Jackson, 2002)) would be a valuable direction for further research. The **formalization of constraints** is crucial to prevent excessive concentrations of informational power within coalitions, reducing the risk of non-competitive or asymmetric configurations. Constraints define the restrictions that the Data Network Game must adhere to. These can be external, such as regulatory requirements related to privacy and security, or internal, aimed at ensuring market balance. External constraints may result in incompatible configurations within the model, while internal constraints can take the form of penalties or incentives designed to promote specific coalitions over others. The implementation of external constraints or the definition of mechanisms such as penalties must be overseen by a designated authority acting as the regulator and governor of the data market. Consequently, the model must incorporate a **dynamic representation of constraints**, capable of adapting to new regulations without destabilizing the market game, and integrating them into the cost and benefit functions of participating organizations. Additionally, the model must identify network structures that violate fair play principles and detect coalitions that may create power imbalances.

A final central aspect in the generalization of the Data Network Game concerns the possibility of structuring a **dynamic data market** in which the data and information generated are treated as tradable assets, and their value fluctuates over time based on key parameters such as quality, rarity, and utility value. In such a configuration, the value of data can be quantified through information metrics well suited to a dynamic setting (Angelelli et al., 2020), enabling real-time evaluation of the effectiveness of coalitions and monitoring the evolution over time of the benefits generated by the big data projects implemented by organizations (Angelelli et al., 2024b). This approach provides a flexible and adaptive framework capable of optimizing the management of informational resources and supporting strategic decisions based on variations in value and demand in the market. Regarding competition and cooperation strategies between coalitions, the model could be enriched to account for the balance between incentives for collaboration and competitive pressures. This would allow for the exploration of more complex scenarios in which coalitions do not only operate as cooperative entities but as actors in a strategic competition.

5.1 Healthcare Case Study: Insights and Implications

The theoretical framework of the Data Network Game requires validation through real-world applications, particularly in sectors such as healthcare, finance, and public services. These domains face significant challenges in balancing privacy, security risks, and the benefits of data sharing among various entities and actors (Cavanillas et al., 2016; Pellegrino et al., 2024). For instance, Systems-of-Systems approaches have been applied to hospital facility management across districts or regions to handle events like pandemics (Pellegrino et al., 2024; Cheng et al., 2022). In this context, extending the model to multi-domain scenarios becomes essential. Despite the advantages of data sharing in healthcare (Shen et al., 2019), numerous barriers persist, especially regarding the risks associated with sharing sensitive data (Van Panhuis et al., 2014). These obstacles are not always rooted in incompatibility constraints but must be addressed to foster a high-quality, relevant data ecosystem that generates community-wide value.

Defining a realistic value function for data domains is a key challenge. Building on collaborations with stakeholders and analyses of real-world use cases, future work will propose and evaluate: (1) A set of value functions centered on data quality, characterized across multiple dimensions (Xiang et al., 2013); (2) constraints based on a multilayer approach for security issues (Faroukhi et al., 2020); and (3) modeling and quantifying the information extracted from data along with its temporal evolution, including obsolescence. In particular, these aspects were selected because data quality is crucial in analytics, and the cost associated with achieving specific quality standards is offset by the benefits such data provide to the overall system (Badewitz et al., 2020). Meanwhile, information extracted from data often loses value over time. Temporal effects on information can be effectively addressed when data are collected over time, using information-theoretic methods (e.g., cross-entropy-based approaches) designed to balance previously acquired information with newly collected data (Angelelli et al., 2020; Angelelli and Konopelchenko, 2021). This, in turn, could provide deeper insights into how coalitions might evolve over time and, consequently, how the data market itself may transform.

6 CONCLUSIONS

In conclusion, the Data Network Game provides an advanced theoretical framework for modeling and representing potential collaborations between organizations focused on data sharing within big data projects. The model emphasizes the importance of structuring stable strategic coalitions, in compliance with regulatory, security, and market constraints, integrating both competitive and collaborative dynamics. The potential value of data, treated as tradable assets, and the informational value generated from these, through the big data projects conducted by the coalitions, are key elements of the Data Network Game. The model is applicable both in community-oriented sectors, such as public administrations and healthcare institutions, and in profit-driven contexts, such as through the creation of data markets. Future perspectives include the expansion of the model to include multi-sector organizations, the creation of a dynamic data marketplace, the integration of adaptive mechanisms to align constraints based on evolving regulatory requirements, and the introduction of multivariate dynamic incentives and constraints within the model. The ultimate goal is to identify market configurations within real-world scenarios that ensure an optimal balance between cooperation and competition, promoting the stability and effectiveness of coalitions in the long term and, therefore, the market itself.

REFERENCES

Ackoff, R. L. (1989). From data to wisdom. *Journal of applied systems analysis*, 16(1):3–9.

Agarwal, A., Dahleh, M., and Sarkar, T. (2019). A marketplace for data: An algorithmic solution. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 701–726.

Aloisio, A., Flammini, M., Kodric, B., and Vinci, C. (2021). Distance polymatrix coordination games. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI 2021*, pages 3–9.

Aloisio, A., Flammini, M., and Vinci, C. (2020). The impact of selfishness in hypergraph hedonic games. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, Proceedings*, pages 1766–1773.

Aloisio, A., Flammini, M., and Vinci, C. (2024). Generalized distance polymatrix games. In *Theory and Practice of Computer Science - 49th International Conference on Current Trends in Theory and Practice of Computer Science, SOFSEM 2024, Proceedings*, pages 25–39.

Angelelli, M., Arima, S., Catalano, C., and Ciavolino, E. (2024a). A robust statistical framework for cyber-vulnerability prioritisation under partial information in threat intelligence. *Expert Systems with Applications*, 255:124572.

Angelelli, M., Ciavolino, E., and Pasca, P. (2020). Streaming generalized cross entropy. *Soft Computing*, 24(18):13837–13851.

Angelelli, M., Gervasi, M., and Ciavolino, E. (2024b). Representations of epistemic uncertainty and awareness in data-driven strategies. *Soft Computing, https://doi.org/10.1007/s00500-024-09661-8*.

Angelelli, M. and Konopelchenko, B. (2021). Entropy driven transformations of statistical hypersurfaces. *Reviews in Mathematical Physics*, 33(02):2150001.

Apt, K. R., de Keijzer, B., Rahn, M., Schäfer, G., and Simon, S. (2017). Coordination games on graphs. *Int. J. Game Theory*, 46(3):851–877.

Azeroual, O. and Nacheva, R. (2023). Data mesh for managing complex big data landscapes and enhancing decision making in organizations. In *KMIS*, pages 202–212.

Badewitz, W., Kloker, S., and Weinhardt, C. (2020). The data provision game: researching revenue sharing in collaborative data networks. In *2020 IEEE 22nd Conference on Business Informatics (CBI)*, volume 1, pages 191–200. IEEE.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of management*, 17(1):99–120.

Bertsekas, D. and Gallager, R. (2021). *Data networks*. Athena Scientific.

Bi, Y., Wu, Y., Liu, J., Ren, K., and Xiong, L. (2024). When data pricing meets non-cooperative game theory. In *2024 IEEE 40th International Conference on Data Engineering (ICDE)*, pages 5548–5559. IEEE.

Bilò, V., Bove, L., and Vinci, C. (2023a). Utility-sharing games: How to improve the efficiency with limited subsidies. In *Proceedings of the 24th Italian Conference on Theoretical Computer Science, 2023*, volume 3587, pages 314–327.

Bilò, V., Ferraioli, D., and Vinci, C. (2022). General opinion formation games with social group membership. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022*, pages 88–94.

Bilò, V., Gourvès, L., and Monnot, J. (2023b). Project games. *Theor. Comput. Sci.*, 940(Part):97–111.

Bogomolnaia, A. and Jackson, M. O. (2002). The stability of hedonic coalition structures. *Games and Economic Behavior*, 38(2):201–230.

Braganza, A., Brooks, L., Nepelski, D., Ali, M., and Moro, R. (2017). Resource management in big data initiatives: Processes and dynamic capabilities. *Journal of Business Research*, 70:328–337.

Catalano, A. A., Catalano, C., Del Vecchio, V., and Lazoi, M. (2024). *Big Data Analytics and Visualization Solutions in Industry: a Literature Review*. Proceedings IFKAD 2024.

Cavanillas, J. M., Curry, E., and Wahlster, W. (2016). The big data value opportunity. *New horizons for a data-driven economy: A roadmap for usage and exploitation of big data in Europe*, pages 3–11.

Chen, H.-M., Kazman, R., Garbajosa, J., and Gonzalez, E. (2017). Big data value engineering for business model innovation. In *50th Hawaii International Conference on System Sciences, HICSS 2017, Hawaii, USA*. ScholarSpace / AIS Electronic Library (AISeL).

Chen, M., Mao, S., and Liu, Y. (2014). Big data: A survey. *Mobile networks and applications*, 19:171–209.

Cheng, W., Lian, W., and Tian, J. (2022). Building the hospital intelligent twins for all-scenario intelligence health care. *Digital Health*, 8:20552076221107894.

Corallo, A., Crespino, A. M., Del Vecchio, V., Gervasi, M., Lazoi, M., and Marra, M. (2023). Evaluating maturity level of big data management and analytics in industrial companies. *Technological Forecasting and Social Change*, 196:122826.

Corallo, A., Fortunato, L., Massafra, A., Pasca, P., Angelelli, M., Hobbs, M., Al-Nasser, A. D., Al-Omari, A. I., and Ciavolino, E. (2020). Sentiment analysis of expectation and perception of MILANO EXPO2015 in twitter data: a generalized cross entropy approach. *Soft Computing*, 24:13597–13607.

Coyle, D. and Manley, A. (2024). What is the value of data? a review of empirical methods. *Journal of Economic Surveys*, 38(4):1317–1337.

Dehghani, Z. (2022). Data mesh: Delivering data-driven value at scale. *O'Reilly Media, Inc.*, 1.ed - preview version.

Del Vecchio, P., Di Minin, A., Petruzzelli, A. M., Panniello, U., and Pirri, S. (2018). Big data for open innovation in smes and large corporations: Trends, opportunities, and challenges. *Creativity and Innovation Management*, 27(1):6–22.

Dong, J. Q. and Yang, C.-H. (2020). Business value of big data analytics: A systems-theoretic approach and empirical test. *Information & Management*, 57(1):103124.

Faroukhi, A., El Alaoui, I., Gahi, Y., and Amine, A. (2020). A multi-layer big data value chain approach for security issues. *Procedia Computer Science*, 175:737–744. The 17th International Conference on Mobile Systems and Pervasive Computing (MobiSPC),The 15th International Conference on Future Networks and Communications (FNC),The 10th International Conference on Sustainable Energy Information Technology.

Fast, V., Schnurr, D., and Wohlfarth, M. (2021). Data-driven competitive advantages in digital markets: An overview of data value and facilitating factors. In Ahlemann, F., Schütte, R., and Stieglitz, S., editors, *Innovation Through Information Systems*, pages 454–454, Cham. Springer International Publishing.

Geerts, G. and O'Leary, D. (2022). V-matrix: A wave theory of value creation for big data. *International Journal of Accounting Information Systems*, 47:100575.

Gervasi, M., Totaro, N. G., Fornaio, A., and Caivano, D. (2023a). Big data value graph: Enhancing security and generating new value from big data. In *Proceedings of the Italian conference on cyber security (ITASEC 2023)*, volume 3488. CEUR-WS, Bari.

Gervasi, M., Totaro, N. G., Specchia, G., and Latino, M. E. (2023b). Unveiling the roots of big data project failure: a critical analysis of the distinguishing features and uncertainties in evaluating big data potential value. In *Proceedings of the 2nd Italian conference on big data and data science (ITADATA 2023)*, volume 3606. CEUR-WS, Naples.

Graef, I., Tombal, T., and De Streel, A. (2019). Limits and enablers of data sharing. an analytical framework for eu competition, data protection and consumer law. *SSRN Electronic Journal*.

Grander, G., Da Silva, L. F., Gonzalez, E. D. R. S., and Penha, R. (2022). Framework for structuring big data projects. *Electronics*, 11(21):3540.

Gupta, M. and George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, 53(8):1049–1064.

Hechler, E., Weihrauch, M., and Wu, Y. (2023). Evolution of data architecture. In *Data Fabric and Data Mesh Approaches with AI: A Guide to AI-based Data Cataloging, Governance, Integration, Orchestration, and Consumption*, pages 3–15. Springer.

Hensler, D. A. and Huq, F. (2005). Value creation: knowledge flow, direction, and adaptation. *International Journal of Learning and Intellectual Capital*, 2(3):278–287.

Huang, T., Lan, L., Fang, X., An, P., Min, J., and Wang, F. (2015). Promises and challenges of big data computing in health sciences. *Big Data Research*, 2(1):2–11.

Inkpen, A. C. (2005). Strategic alliances. *The Blackwell handbook of strategic management*, pages 403–427.

Kathuria, V. and Globocnik, J. (2020). Exclusionary conduct in data-driven markets: limitations of data sharing remedy. *Journal of Antitrust Enforcement*, 8(3):511–534.

Kleinberg, J. M. and Oren, S. (2022). Mechanisms for (mis)allocating scientific credit. *Algorithmica*, 84(2):344–378.

Konsynski, B. R. and McFarlan, F. W. (1990). Information partnerships–shared data, shared scale. *Harvard Business Review*, 68(5):114–120.

Koutroumpis, P., Leiponen, A., and Thomas, L. D. (2020). Markets for data. *Industrial and Corporate Change*, 29(3):645–660.

Laney, D. (2001). 3-D Data Management: Controlling Data Volume, Velocity, and Variety. Technical report, META Group Res. Note.

Liang, F., Yu, W., An, D., Yang, Q., Fu, X., and Zhao, W. (2018). A survey on big data market: Pricing, trading and protection. *Ieee Access*, 6:15132–15154.

Lindstedt, D., Graziano, K., and Hultgren, H. (2009). *The business of data vault modeling*. Lulu. com.

Louati, R. and Mekadmi, S. (2019). Toward a conceptualization of big data value chain: From business problems to value creation. In *Business Transformations in the Era of Digitalization*, pages 76–92. IGI Global.

Malthouse, E. C., Buoye, A., Line, N., El-Manstrly, D., Dogru, T., and Kandampully, J. (2019). Beyond reciprocal: the role of platforms in diffusing data value across multiple stakeholders. *Journal of Service Management*, 30(4):507–518.

Myerson, R. B. (1980). Conference structures and fair allocation rules. *International Journal of Game Theory*, 9:169–182.

Pellegrino, G., Gervasi, M., Angelelli, M., and Corallo, A. (2024). A conceptual framework for digital twin in healthcare: Evidence from a systematic meta-review. *Information Systems Frontiers, https://doi.org/10.1007/s10796-024-10536-4*.

Peters, D. and Elkind, E. (2015). Simple causes of complexity in hedonic games. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015*, pages 617–623.

Priebe, T., Neumaier, S., and Markus, S. (2021). Finding your way through the jungle of big data architectures. In *2021 IEEE International Conference on Big Data (Big Data)*, pages 5994–5996. IEEE.

Ramchand, S. and Mahmood, T. (2022). Big data architectures for data lakes: A systematic literature review. In *2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC)*, pages 1141–1146. IEEE.

Reggio, G. and Astesiano, E. (2020). Big-data/analytics projects failure: a literature review. In *2020 46th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, pages 246–255. IEEE.

Roos, G. (2018). Resource deployment system implications of migrating the firm into a digital value creation paradigm. *Knowledge Management Research & Practice*, 16(3):281–291.

Saltz, J. S. (2015). The need for new processes, methodologies and tools to support big data teams and improve big data project effectiveness. In *2015 IEEE International Conference on Big Data (Big Data)*, pages 2066–2071. IEEE.

Sharma, V., Balusamy, B., Thomas, J. J., and Atlas, L. G. (2023). *Data Fabric Architectures: Web-Driven Applications*. Walter de Gruyter GmbH & Co KG.

Shen, B., Guo, J., and Yang, Y. (2019). Medchain: Efficient healthcare data sharing via blockchain. *Applied sciences*, 9(6):1207.

Spulber, D. F. (2023). Antitrust and innovation competition. *Journal of Antitrust Enforcement*, 11(1):5–50.

Sung, S.-C. and Dimitrov, D. (2010). Computational complexity in additive hedonic games. *European Journal of Operational Research*, 203(3):635–639.

Tanriverdi, H. and Venkatraman, N. (2005). Knowledge relatedness and the performance of multibusiness firms. *Strategic management journal*, 26(2):97–119.

Tiefenbacher, K. and Olbrich, S. (2015). Increasing the value of big data projects–investigation of industrial success stories. In *2015 48th Hawaii International Conference on System Sciences*, pages 294–303. IEEE.

Van Panhuis, W. G., Paul, P., Emerson, C., Grefenstette, J., Wilder, R., Herbst, A. J., Heymann, D., and Burke, D. S. (2014). A systematic review of barriers to data sharing in public health. *BMC public health*, 14:1–9.

Woeginger, G. J. (2013a). Core stability in hedonic coalition formation. In *SOFSEM 2013: Theory and Practice of Computer Science, 39th International Conference on Current Trends in Theory and Practice of Computer Science, Proceedings*, pages 33–50. Springer.

Woeginger, G. J. (2013b). A hardness result for core stability in additive hedonic games. *Math. Soc. Sci.*, 65(2):101–104.

Wu, X., Liang, L., and Chen, S. (2022). How big data alters value creation: Through the lens of big data competency. *Management decision*, 60(3):707–734.

Xiang, J., Lee, S., and Kim, J. K. (2013). Data quality and firm performance: Empirical evidence from the korean financial industry. *Information Technology and Management*, 14.

Ylijoki, O. and Porras, J. (2019). A recipe for big data value creation. *Business Process Management Journal*, 25(5):1085–1100.