Qualitative Structural Model for Capabilities in Open Data Organizations

Fatemeh Ahmadi Zeleti
Insight Centre for Data Analytics @ NUI Galway
Ireland
fatemeh.ahmadizeleti@insight-centre.org

Adegboyega Ojo
Insight Centre for Data Analytics @ NUI Galway
Ireland
adegboyega.ojo@insight-centre.org

Abstract

Open data is increasingly becoming an essential asset for many organizations. However, large numbers of organizations fall short when it comes to utilizing open data effectively to fully leverage the potential of it. There are ample evidences that this shortcoming is attributable to the poor understanding of what types of capabilities are required to successfully conduct data related activities. At the same time, research on open data capabilities and how they relate to one another remains sparse. Based on the theoretical foundation constructed from the integration of Capability-based Theory and Dynamic Capability Theory and, extant literature and interviews of leadership of open data organizations, we attempt to address this knowledge gap by investigating open data capabilities and relationships between them. Findings help validate the two theories in the open data organizations and reveal unknown knowledge about open data capability areas and how they affect one another.

1. Introduction

Open data is an essential organizational asset for many organizations and large numbers of new start-ups are beginning to benefit from the potential of this asset for a wide range of new products and services [1]. In spite of high investment in developing open data technical and infrastructure capabilities [2], large numbers of these organizations fail to effectively use open data and fully leverage its potential [3][4]. A major reason for this is that these organizations do not clearly know what specific capabilities are required to effectively harness open data for their business needs and organizational goals [5]. For example, In 2015, the result of PwC survey of 1,800 organizations [3] show that 75% of them lack the capabilities to utilize open data. In 2016, an exploratory research of 33 open data organizations in UK [6] concludes that related capabilities and activities remain vague in these organizations. Another data capability project in UK in 2016 [4] reveals that lack of understanding and capabilities put organizations at risk. Studies show that in order to compete and survive in the fast changing and competitive open data industry, open data

organizations are required to plan and develop capabilities for generating value from open data, increasing agility and competitive advantage of the organization [7][4].

To our knowledge, no previous scholarly work has attempted to comprehensively identify open data capability areas and their relationships to each other and, articulate a structural model for capabilities in open data organizations. Many studies such as the Open Data for Business report of The World Bank [8], UK's Data Strategy study [9], and a study of Dynamic Capabilities [10] strongly suggest further research into open data capabilities. Therefore, in this work, we attempt to robustly address this research gap driven by two research questions (RQ): RQ1) what are the main elements of the structural model and RQ2) how do these elements relate to each other?.

In this research, we define *open data organizations as both non-profit and for-profit organizations that use, produce, or otherwise invest in open data as a key aspect of their operation for generating customer value and achieving organization's mission goals*.

2. Theoretical Background

The theoretical background of this paper is created by conducting an extensive review of extant literature on organizational capabilities and related theories. We elaborate on the two theories: Capability-based View and Dynamic Capability Theory. The two theories help us explain the competitive advantage of the organizations from the capability-based perspective. Organizations generate value, respond to the changing environment, and compete when they develop or acquire a set of organizational capabilities [11]. Therefore, we also elaborate on the three organizational capability types and areas associated with each type. Last in this section, based on the background and related theories, we present the resulting theoretical model and research hypothesis (H) for this study (Figure 1).

2.1. Related theories
2.1.1.   Capability-based view

In the area of firm competitive advantage, the Capability-based View (CBV) has been an influential theoretical model for understanding how competitive advantage within an organization is achieved and how that advantage might be sustained over time [11]. This view suggests that capabilities are the source of competitive advantage while resources are the source of capabilities and that organizations can gain competitive advantage from its ability to apply its capabilities to perform important activities within the organization [12]. According to [13], organizational capabilities are the source of competitive advantage, but not the organizational resources. In the CBV of the organization, an organization's performance is affected by organization-specific capabilities such as specific physical (e.g., specialized equipment and infrastructure), human (e.g., expertise), and organizational (e.g., superior sales force) capabilities [14], that can be used to implement value-creating activities [15] and, are fundamental to the competitive advantage of an organization [16][17].

2.1.2. Dynamic capability theory

CBV is not able to provide explanations as to how organizations respond in a timely manner to the market changes and, product and process innovation, along with the management capability to effectively coordinate and redeploy internal and external competences [18]. Therefore, CBV has been criticized for conceptual vagueness and for its inadequacy in a context characterised by unpredictable change [16][19], termed high-velocity or dynamic markets [14]. As a result, the organizational capability literature provides another major capability-based perspective named Dynamic Capability Theory (DCT) to address how organizational capabilities can be created and refreshed in changing environments [18][20]. If an organization possesses processes, resources, and competencies but, lacks dynamic capabilities, it has a chance to make a competitive return for a short period, but superior returns cannot be sustained. The possession and deployment of dynamic capabilities provide the business enterprise with a chance to generate superior profitability over the longer run. When organizations are dynamic, management will be active at sensing and seizing opportunities [21]. According to [22], dynamic capabilities allow organizations to renew and leverage their internal and external capabilities thereby enabling it to coordinate inter-organizational activities and respond rapidly, in a flexible manner, to global competitors' strategies [17].

2.2. Organizational Capability Types and Areas

In the literature, three types of organizational capability areas are identified and described based on the well-known edicts of CBV and DCT. The three capability areas include *1) Value capabilities*, *2) Dynamic capabilities*, and *3) Competitive capabilities* [23]. Below, we briefly describe each capability areas and its core elements as being presented in the related literature. Table 1 is a summary table of the discovered capability types and areas from the literature.

Value capabilities

The creation of 'value' is the key in every organization. 'Value' in the products and services is what makes customers and end users satisfied and loyal with the organization's offering [24]. Capabilities are required for every organization to develop this 'value'. This includes capabilities that are characterized by value, heterogeneity, and imperfect mobility. Value capabilities include all capabilities which assist an organization to deliver the organization value to the customers. While value capabilities are not the source of competitive advantage, they are necessary to produce customer value. Value capabilities includes: *Individual Competences, Business Processes, Organizational capabilities, IT and Technological Infrastructure, and Management and Governance capabilities* [23][25][26].

Dynamic capabilities

The majority of the studies on dynamic capability assert that dynamic capabilities are the ability of the organization to renew its capabilities to deal with rapidly changing environments [27]. [28] defines dynamic capabilities as "a firm's capacity to deploy resources, usually in combination, using organizational processes, to effect a desired end". Dynamic capabilities allow the organization to search and explore, acquire, and assimilate new resources and capabilities that can help the organization to develop new opportunities [29]. Dynamic capabilities include: *Process Innovation, Knowledge Management and Organizational Learning, Value Chain Performance, Relationship Infrastructure, and Management Functions* [23][26][30].

Competitive capabilities

This strategic level capability includes all he capabilities that foster the organization's competitive advantage and allow organizations to stay competitive and outperform competitors. Competitive capabilities are the key to the success and profitability of the organization [23][19]. Because, as the level of dynamics in business environments increases, the development of strategies that will differentiate the

organization from its competitors becomes the key success factor [5]. Competitive capabilities include: *Enterprise Infrastructure Strategy, Product and Service Strategy, Business Development Strategy, and Relational Rent Strategy* [31][23][26].

Table 1: Three organizational capability types and areas

Value Capability Areas
Individual Competences
Business Processes
Organizational
IT and Technological
Management and Governance
Dynamic Capability Areas
Process Innovation
Knowledge Management and Organizational Learning;
Value Chain Performance
Relationship Infrastructure
Management Functions
Competitive Capability Areas
Enterprise Infrastructure
Product and Service
Business Development
Relational Rent

2.3. Relationships between Capability Areas

Evidence from the last decades shows that organizations can successfully engage in both product and service development and performance and process innovation and create relationships between the two [32]. A better product and service innovation capability can create products with higher value than those of competitors [33]. Also, organizational knowledge management and learning capabilities is critical to product and service performance and innovation. Organization with innovative knowledge is able to introduce innovative products or services, potentially helping it become a market leader [11][34]. Similarly, according to [35], knowledge management is one of the main resources responsible for results in terms of flexibility, expressed as innovation and responsiveness to clients in regard to product and service performance and improvement upon client's request. Moreover, intensive use of knowledge management capabilities

enables information to be identified, captured and capitalized as input to business process development and innovation [35][36]. [35] highlights the importance of incorporating intellectual capital as a nodal capability in the pursuit of process efficiency and flexibility. Therefore, [35] proved that there is a relationship between knowledge management and business process improvement and that process innovation or improvement is dependent on the availability of both internal and external knowledge and learning to the organization. Organization that begins with superior knowledge; it is more likely to gain further knowledge because of its prior knowledge [23].

At the strategy level, business strategy is so important to an IT and technological strategy. A focused, driven business strategy will lead to the most efficient application of IT expenses as the result of the appropriate IT strategy implementation. With no IT strategy, an enterprise inherits an IT and architectural maze that becomes so expensive to maintain and support for business constituents, they will eventually rebel at the high costs and suboptimal service that IT provides [37][38].

2.4. Theoretical Model and Hypothesis

Relying on our research presented in sections 2.1, 2.2, and 2.3, this section contributes to development of a theoretical model and research hypothesis for this study. As shown in Figure 1, the theoretical model consists of three major capability areas for generating value from open data (black blocks), agility (dark gray blocks), and competitive advantage (light gray blocks) based on the literature presented in section 2.2 and, it also includes hypothetical relationships between every two main capability areas based on the literature presented in section 2.3.

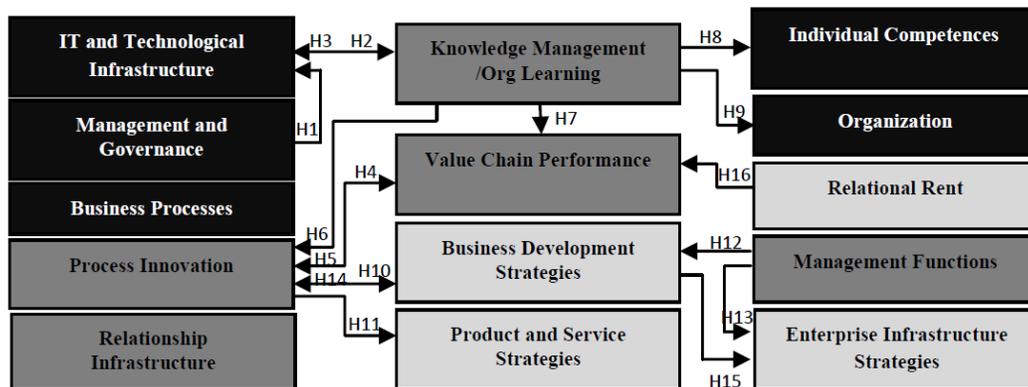## 3. Methodology

3.1. Research Objectives

Figure 1. Theoretical Model and Research Hypothesis

The objectives of this research are 1) to provide a conceptual approach to develop a structural model for capabilities for value creation, agility, and competitive advantage in open data organizations and 2) to refine the theoretical model for open data, based on empirical evidence collected through the in-depth interviews of 11 successful and revenue generating open data organizations located in different geographical areas. The structural model developed in this work helps research and practice community to understand capability areas that are important for open data organizations and the relationship between them.

3.2. Research Method and Approach

Following qualitative research method, we mainly rely on existing literature and theories to develop our theoretical model where we later refine based on the empirical study of 11 open data organizations. Moreover, we explore deductive research approach [39] to explore known theories and to test whether the theories are valid in a given circumstances or not.

This approach starts with analysis of existing work and related theories, and then, it leads us to develop theoretical model and hypothesis to be tested which either lead to confirmation or rejection. This approach allows us to deliver numbers of research stages including: 1) Analysis of existing work and related theories (sections 2.1, 2.2, and 2.3), 2) Developing a theoretical model (section 2.4), 3) Formulating research hypothesis (section 2.4), 4) Testing the hypothesis based on the analysis of the data collected through in-depth interviews of 11 open data organizations (section 4), and 5) Refining the theoretical model for open data organizations and developing the structural model (section 5).

3.3. Research Process

3.3.1. Review of literature and related theories

Through the review of the literature and related theories, we establish understanding of the domain knowledge. We specifically aim to understand different organizational capability types and capability areas for value creation, agility, and competitive advantage of the organization. The model developed includes three main capability types and capability areas associated with each type (Table 1). We further review existing literature in 'general business' domain to discover possible relationships (H) or links between different capability areas.

3.3.2. Research hypothesis and theoretical model

Guided by our research questions (See Introduction) and building upon the knowledge generated through the previous step (3.3.1), we proposed research hypothesis (H) and developed the theoretical model for this study (Figure 1).

3.3.3. Interview protocol and instrument

In this step, we aim to conduct in-depth interviews of 11 CEOs or decision makers of successful and revenue generating open data organizations to better understand the domain (open data capabilities) and, empirically testing the hypothesis. The choice of semi-structured interviews has been determined by the lack of existing work on open data capabilities which arises the need to explore and dig into the interviewee's perspectives on the topic under investigation [40]. Therefore, the collected data supports analysis of capabilities of open data organizations and guides us in development of the capability structural model for open data organizations.

Theoretical model was used as the basis to design and develop our semi-structured interview protocol. Through the interview sessions, we mainly wanted to explore the relationships between capability areas (hypothesis) and specific open data capability associated with each relationship.

For the selection of potential organizations, we request to access the third party's dataset of 685 organizations around the world that use open data in some forms. We mainly looked for those organizations that: 1) rely on open data as one of their key resources to achieve mission goals, 2) the application of open data is primarily in developing new products and services, and 3) are generating revenue.

Taking into consideration the above criteria, 68 organizations were shortlisted and contacted. We managed to conduct interviews of 11 open data organizations (seven profit organizations and four non-profit organizations) that showed interest and agreed to participate in our study.

Participants from different geographical locations are considered because they each bring different perspective which enriches our understanding of the domain. The interview was designed to take no more than 1 hour and 15 minutes. Interviews are voice recorded following the consent of the interviewees and each of the recorded files has been carefully transcribed into a separate text document.

The interview instrument comprises three main parts: Organizational Background, Use of Open Data in the Organization and, Open Data Capabilities in the Organization.

3.3.4. Synthesis - coding and analysis

In this step, we code and analyze the interviews data to synthesize and construct our open data capability model. To code the data, we develop concept hierarchy [41] based on the elements presented in the theoretical model (Figure 2). We specifically did this by eliciting key aspects of organizational capabilities (tackled by the RQ1) outlined in our theoretical model (Figure 1) and then we code specific word, label, or phrase from the interviews data to the appropriate concept (coding example is shown in Figure 3).

Regarding the relationships (hypothesis in Figure 1) between the capability areas (RQ2), we develop a logic that allows coding and analysis of the data that represent each relationship. For example, we code all the data that shows that a relationship exist between Capability X and Capability Y under 'X impacts Y'. A partial snapshot of the relationship coding interface is shown in Figure 4.

To code each transcript based on the concept hierarchy, we used NVivo. NVivo is a strong and comprehensive qualitative data analysis software platform which can be used to organize and analyze any types of qualitative data [42][43] and to "obtain rigor in dealing with such data" [43]. Through the coding process, we 1) select a particular phrase, sentence, paragraph or whole section of the text and 2) assign this fragment to a specific concept following the developed concept hierarchy. Any text that could not be placed to any existing concept would be given a new concept to be code to [44]. This increases trustworthiness as we make sure that we capture all possible variables or concepts from the transcripts and enables the theoretical model to be refined and extended as the coding progresses [45].

To analyze the data based on the coding and concept hierarchy, we follow standard steps to qualitative data analysis [42][43][46][45]. Furthermore, we adopt *Data Analysis Triangulation* through adopting the three data analysis techniques: *Classical Content Analysis, Taxonomic Analysis* and *Frequency Check* [47]. Through Classical Content Analysis, we allow new concepts or codes to emerge following our concept hierarchy. Through Taxonomic Analysis, we allow new categories to emerge from the concepts or codes which may not be covered in our hierarchy. Through Frequency Check, we take note of the number of times each specific category and concept is being coded – this will assist us in identifying the perceived importance of different capabilities in open data organizations.

4. Analysis

Analysis of the data collected from open data organizations revealed numbers of main open data capability areas, relationships between these capability

Figure 2. Concept hierarchy addressing RQ1

Figure 3. An example of coded interview data to a concept (Business Development Strategic Capability) – organization's names are removed for anonymity

Figure 4. Analysis addressing RQ2 – An example of coded interview data for a relationship (Process Innovation Capability Impacts Business Development Strategic Capability) - organization's names are removed for anonymity

areas, and specific capabilities associated with each relationship.

In Table 2, we present all the main open data capability areas that have emerged from our data analysis (RQ1). We give unique code to each discovered open data capability area to simplify presentation of the analysis of the relationships between the areas.

Table 2. Main open data capability areas discovered during the data analysis

Code	Main Open Data Capability Areas
Capability areas for value generation from open data	
C1	Individual Competences and Expertise
C2	Open Data Processes
C3	Organization
C4	IT and Technological Infrastructure
C5	Management and Data Governance
Capability areas for enabling agility in open data organizations	
C6	Process Innovation
C7	Knowledge Management and Organizational Learning
C8	Value Chain Performance
C9	Relationship Infrastructure
C10	Management Functions
Capability areas for improving competitive advantage of open data organizations	
C11	Enterprise Infrastructure
C12	Product and Service
C13	Business Development
C14	Relational Rent

Our analysis of the open data organizations also revealed that, capability areas presented in Table 2 are impacting one another through other specific open data capabilities. In this analysis, capability area 'C_i' is impacting capability area 'C_j' if there exist specific capabilities associated with 'C_i' that are impacting 'C_j'. In Table 3, we show the relationships and specific open data capabilities that are associated with each relationship (RQ2). As an example from Table 3, R26 which is discovered during the analysis of the data shows that, there exist a relationship between C1 and C4 indicating that C1 is impacting C4 through other specific capabilities. In this example, individual's capabilities related to their 'Prior Open Data Learning and Experiences' can directly impact implementation of C4 (IT and Technological Infrastructure capability).

Table 3. Relationships between capability areas and their associated specific open data capabilities

Relation As shown in structural model	Relation As shown in the theoretical model	Capabilities associated with each relationship
R26 C1-C4	Discovered relationship	Prior Open Data Learning and Experiences
R9 C1-C7	Discovered relationship	Regular Reports, Prior Open Data Learning and Experiences
R8 C7-C1	**H8** supported	Assess and Identify Open Data Training Areas, Assess, Capture, and Analyse Internal Knowledge, Collaborative Open Data Projects and Learning, Involve Individuals in Knowledge Creation Process, Knowledge Acquisition and Mergers, Open Data Market Learning
R27	Discovered	Adopting 3rd Party Vendor's Data Processes, Alliance-Based Data Processes, Assess and Adopt Data Process Best Practices, Assess and
C6-C2	relationship	Adopt New Data Processes and Tools, Assess and Increase Data Process Efficiency, Open Data Related Resource Availability and Allocation
R10 C7-C3	**H9** supported	Assess and Analyze Acquired Knowledge, Assess, Capture, and Analyse Internal Knowledge, Knowledge Management Tools (e.g. Atlassian, GitHub), Knowledge Pattern Development and Use
R18 C4-C5	Discovered relationship	Analytics Tools
R1 C5-C4	**H1** supported	Data Standards, Data Privacy, Governance Model, Technical Knowledge, Data Policy, Data Value Governance, Data Modelling, Data Concept, Open Data Guidelines, EU Directives, Legal Frameworks, Data Best Practices, Data Cultural Shift
R2 C4-C7	**H2** supported	Data and Big Data Technologies, Data Collection Infrastructure, Data Infrastructure Performance Monitoring System, Sensors, Analytics Tools, Data Management Systems, Data Store and Computing Power, APIs and Channels
R3 C7-C4	**H3** supported	Assess and Adopt Data Best Practices, Access and Analyze Acquired Knowledge, Assess, Capture, and Analyse Internal Knowledge, Collaborative Open Data Projects and Learning, Engage and Participate in Data Ecosystem, Knowledge Acquisition and Mergers, Knowledge Clustering and Classification, Knowledge Management Tools (e.g. Atlassian, GitHub), Open Data Market Learning, Open Data Offering's Problem Scope Identification
R25 C4-C11	Discovered relationship	Data Infrastructure Performance Monitoring System
R6 C7-C6	**H6** supported	Assess and Adopt Data Best Practices, Assess and Analyze Acquired Knowledge, Assess, Capture, and Analyse Internal Knowledge, Collaborative Open Data Projects and Learning, Engage and Participate in Data Ecosystem, Knowledge Acquisition and Mergers, Knowledge Clustering and Classification, Knowledge Management Tools (e.g. Atlassian, GitHub), Open Data Market Learning, Knowledge Pattern Development and Use, Open Data Offering's Problem Scope Identification, Research and Development
R4 C6-C8	**H4** supported	Adopting 3rd Party Vendor's Data Processes, Alliance-Based Data Processes, Assess and Adopt Data Process Best Practices, Assess and Adopt New Data Processes and

		Tools, Assess and Increase Data Process Efficiency, Open Data Related Resource Availability and Allocation
R5 C8-C6	**H5** supported	Assess and Increase Data Domain, Assess Open Data Offering Performance and Requirements, Co-Creating Open Data Offering, Feedback Loop and Data Request
R12 C6-C12	**H11** supported	Assess and Adopt New Data Processes and Tools, Assess and Increase Data Process Efficiency, Open Data Related Resource Availability and Allocation
R20 C12-C6	Discovered relationship	Data Quality Continuous Monitoring, Constantly Increase Data Quality, Data Quality Attribute and Measurement, Support Data with Metadata, Data Quality Evaluation, Using Data Quality Assessment Tools, Fast and Efficient Delivery through APIs, Flexible Design for Data Service Expansion, Criteria Search-based Quick Response, Update-based Quick Response
R11 C6-C13	**H10** supported	Alliance-Based Data Processes, Assess and Adopt New Data Processes and Tools, Assess and Increase Data Process Efficiency, Open Data Related Resource Availability and Allocation
R15 C13-C6	**H14** supported	Advance Integration, Analysis of the State of the Art, Collecting Large amount of Data, Data Relationship Discovery, Disruptive Data Product Innovation, Identify Good Data Sources, Engaging in Various Projects, Focus on Principles of Big Data, Move from Close Data to Shared Data, Targeting New Sectors, Understand and Develop Open Data and Big Data Capabilities, Cost Containing, Resource and Budget Allocation, Understanding Legal Framework within each Sector, Find and Provide to a Related Open Data Market, New Contracts and Agreements with Data Users, Operate in a New Open Data Sector, Oversee Data Market, Targeted Market, Product-save-time Pricing, Resource-based Pricing, Value-based Pricing
R22 C6-C14	Discovered relationship	Adopting 3rd Party Vendor's Data Processes, Alliance-Based Data Processes, Assess and Adopt New Data Processes and Tools
R19 C14-C6	Discovered relationship	Open and Provide Data to Others, Develop Open Data Site
R7 C7-C8	**H7** Not supported *Relationship changed from direct to indirect*	*This indirect relation returns no interpretation from the data. However, we decided to keep the relationship because it is found to be different in the open data context*
		Collaborative Open Data Projects

R24 C7-C9	Discovered relationship	and Learning, Engage and Participate in Data Ecosystem, Open Data Offering's Problem Scope Identification
R21 C14-C7	Discovered relationship	Open your Data to Other Companies, Open your Data through Open Data Portal
R17 C14-C8	**H16** supported	Team Collaboration and Sharing , Internal Data Warehouse, External Data Warehouse, Build Example from your Data and Linked Data, Develop Testable Prototype for Customers, Open your Data and Tools to Customers to Investigate Value and Needs, Collect Data User's Stories, Provide as Open Source, Open your Data to Other Companies, Merging with Other Companies, External Data Warehouse, Share of Resources and Capabilities
R14 C10-C11	**H13** supported	Financial and Investment Assessment, Market and Clients Assessment, Technical Assessment
R13 C10-C13	**H12** supported	Agile Mind-Set, Financial and Investment Assessment, Identify New Areas of Operation, Market and Clients Assessment, Research
R16 C13-C11	**H15** supported	Intelligent Data Models, Operational Efficiency, Products Usable by All, More Automated Data Services, Collecting Large amount of Data, Open to try new Data Tools and Infrastructure, Publish all the Data, Moving into Business Intelligence and IoT, Targeting New Sectors, Tailor and Fit-to-use Existing Tools, More Open Source, Operate in a New Open Data Sector, Open all for all, Collect more Data and Expand the Market
R23 C13-C12	Discovered relationship	Operate in a New Open Data Sector, Targeted Market

5. Findings

The structural model for open data capabilities is developed based on the analysis of the interview data. This model is a refinement of the theoretical model presented in section 2.4. In Figure 5, we show the structural model for open data capabilities which includes 14 capability areas and the relationships between the capability areas. Analysis shows that, the 14 capability areas presented in the theoretical model are all true in open data context. Moreover, out of the 16 hypothesis, analysis supports 15 hypotheses (H1 to H6 and H8 to H16). Analysis did not support H7. In addition, analysis revealed 11 new relationships (shown in Table 3 and Figure 5).

We establish the perceived degree of importance of different capability areas based on the Frequency Check analysis. In the category of 'Value', our finding

suggests that, the two most important open data value capability areas are Individual Competences and Expertise, and Management and Data Governance. In both profit and non-profit organizations, capabilities related to data competences and expertise of working individuals plays a significant role in value generation from open data. However, Management and Data Governance capabilities are found to be more important to profit organizations.

In the category of 'Agility/Dynamic', our finding shows that, the two most important open data dynamic capability areas are Knowledge Management and Organizational Learning and, Process Innovation capabilities. Knowledge Management and Organizational Learning capabilities in profit organizations are significantly higher than non-profits. According to our findings, this capability area has revenue potentials for profit organizations and can contribute to increasing competitive advantage and profitability in these organizations because, knowledge can contribute to better and faster problem solving than rivals. However, this capability area has other potentials for non-profits such as: 1) increasing transparency and trust, 2) improving data management, 3) providing training, 4) locating affordable resources, 5) unleash the potentials of team members, and 6) Identifying the right data infrastructure. Moreover, our finding suggests that Process Innovation capabilities are underdeveloped in non-profits compared to for-profits.

In the category of 'Competitiveness', our finding shows that, the two most important capability areas are Business Development and Product and Service Strategic Capabilities. Our finding highlights that,

Business Development capability is equally important to both profit and non-profit organizations however, this capability in non-profits are mostly defined at a higher level, whereas, in profit organizations, this capability area captures a wider spectrum of activities, stakeholders, and strategic planning. Product and Service competitive capabilities are developed in both profit and non-profit organizations to a certain extent.

In addition to the above findings, our analysis suggests that, least attention is given to the development of dynamic capabilities which indicates that open data organizations do not yet find themselves in a favorable situation to response to open data market dynamics and therefore they cannot position themselves based on their agility strength.

As can be seen from our analysis, data does not support H7. In this regard, finding suggests that Knowledge Management and Organizational Learning capabilities have an impact on Value Chain Performance through Process Innovation capabilities. Knowledge and learning that is created through developing Knowledge Management and Organizational Learning capabilities are required to improve process innovation capabilities and increase the efficiency of data processes therefore, this can impact the value chain performance of the organization.

6. Validation

In this part, we put forward the arguments for the validity of our structural model for open data capabilities which aims is to ensure that *the developed structural model accurately preserves the relationships*
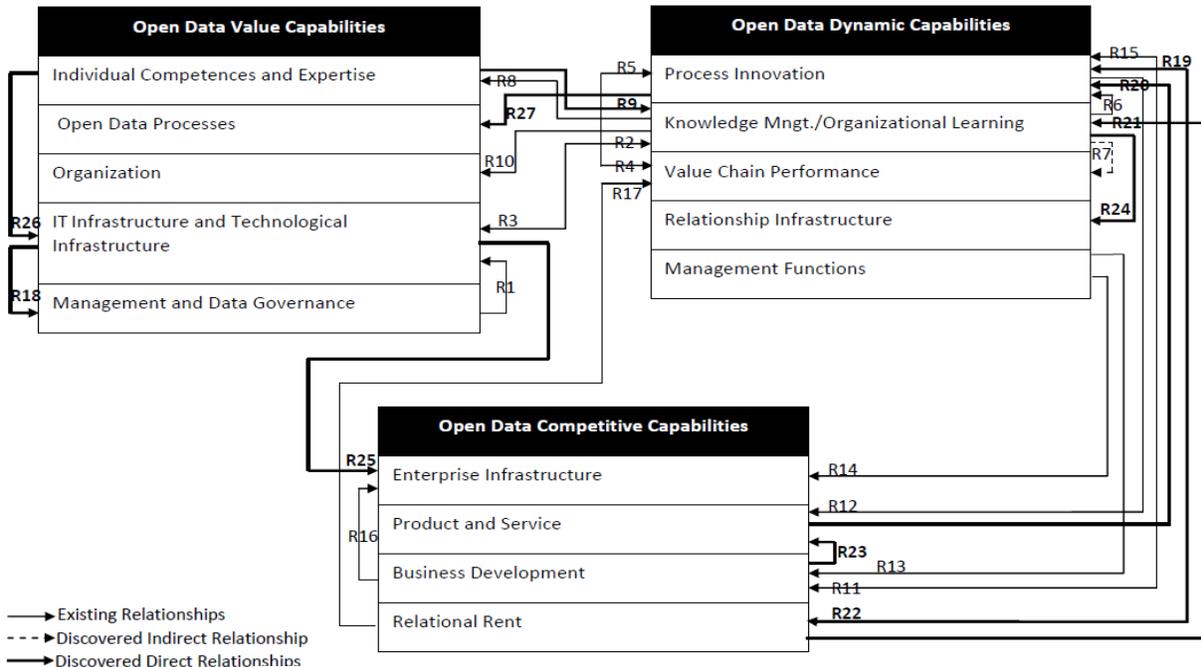

Figure 5. Structural model for open data capabilities

and elements from the interviews data [48]. Our claim mainly rests on the 1) developed instrument which was based on the two related well-known theories and literature (theoretical model) and 2) adopted qualitative research validation approach [48][49][50][51] which assists us in looking to see whether the developed model including all its elements and relationships preserves the interviews data collected and analyzed or not. During the validation process, we detected two errors and we made an attempt to fixing the problems. As a result of validation, the data implies the structural model for open data capabilities presented in Figure 4.

In addition, the instrument and data was peer reviewed by an experienced researcher and domain expert. Moreover, the results and findings are consistent with the CBV and DCT.

7. Discussion

The results presented in Section 5 and validated in Section 6 strengthen, through empirical evidence, the model and research hypothesis in Figure 1. The developed structural model for open data capabilities in Figure 5 differs from those that have been proposed for the open and big data domains [1][52][53][54][55][9][56]. For example, insights from a scenario, interviews, and a survey study [1] reveals three capability areas: IT, Information and Data, and Human. The developed model in Figure 5 confirms these three capabilities and adds to it by outlining other open data capability areas that are equally important to open data organizations. Similarly, in other cited studies, some aspects or areas appear to be relevant but, no structural model has been developed that shows the main open data capability areas, relationships between the areas, and specific capabilities associated with each relationship. Therefore, the structural model developed in this work advances the current body of knowledge with new findings.

In addition, consistent with the CBT and DCT, through our theoretical model, we showed that open data capabilities contribute to improving agility and competitive advantage in open data organizations. Therefore, the theories used in this work are applicable to the open data context. In addition, the theories could also be useful to other individuals or researchers who want to apply them in similar situations.

In our future work, we aim to quantify the strength of each relationship in our structural model to better understand the influence of each capability area.

8. Conclusion

Building upon the holistic enterprise perspectives of the CBT and DCT and, the results of expert interviews, we have developed a structural model for open data capabilities for open data organizations. The model gives an initial, yet unique and, empirically and theoretically grounded view of the capabilities that organizations require to generate value from open data, improve agility and obtain competitive advantage. The model includes main open data capability areas, relationships between the capability areas and, a set of capabilities associated with each relationship. By developing the model, we aim to help organizations or start-ups whose aim is to use open data to meet their business objectives to better understand open data capabilities and how capability areas are related to one another. By adopting the structural model, open data organizations can create a solid foundation for effectively harnessing open data. In general, supported by the adopted theories, we claim that, the developed structural model not only helps reinforce the competitive advantage of open data organizations but also other organizations in general business domain.

In addition, the developed structural model provides governments with the bases to develop better and more informed strategic decisions to support opening up more data to both public and private organizations and to measure the performance of these organizations and perform benchmarking.

9. References

[1]   A. Zuiderwijk, M. Janssen, K. Poulis, and G. van de Kaa, "Open Data for Competitive Advantage: Insights from Open Data Use by Companies," in *Proceedings of the 16th Annual International Conference on Digital Government Research (dg.o 2015)*, 2015, pp. 79–88.

[2]   D. Sammon and T. Nagle, "The Data Value Map: A framework for developing shared understanding on data initiatives," in *25th European Conference on Information Systems*, 2017, pp. 1439–1452.

[3]   C. Reid, R. Petley, J. McClean, K. Jones, and P. Ruck, "Seizing the Information Advantage: How organizations can unlock value and insight from the information they hold," 2015.

[4]   HEDIIP, "Data Capability: A call to action," UK, 2016.

[5]   J. M. Gathungu and J. K. Mwangi, "Dynamic capabilities, talent development and firm performance," *DBA Africa Manag. Rev.*, vol. 2, no. 3, pp. 83–100, 2012.

[6]   J. Fawcett, "Examining open data as a source of competitive advantage for big businesses," in *Open Data Research Symposium*, 2016.

[7]   Booz Allen Hamilton, "Tips for Building a Data Science Capability," Hamilton, 2015.

[8]   The World Bank, "Open Data for Business (OD4B) Tool," Washington DC, 2015.

[9]   HM Government, "Seizing the data opportunity: A strategy for UK data capability," UK, 2013.

[10]   B. Klievink and M. Janssen, "Realizing joined-up government - Dynamic capabilities and stage models for transformation," *Gov. Inf. Q.*, vol. 26, no. 2, pp. 275–284, 2009.

[11]   H. Wang, "Theories for competitive advantage," in *Being Practical with Theory: A Window into Business Research*, Wollongong, Australia, 2014, pp. 33–43.

[12]   R. M. Grant, "The Resource-based theory of competitive advantage: implications for strategy formulation," *Calif.*

Manag. Rev., vol. 33, no. 3, pp. 114–135, 1991.

[13] R. Amit and P. J. H. Schoemaker, "Strategic Assets and Organizational Rent," *Strateg. Manag. J.*, vol. 14, no. 1, pp. 33–46, 1993.

[14] E. M. Daniel and H. N. Wilson, "The role of dynamic capabilities in e-business transformation," *Eur. J. Inf. Syst.*, vol. 12, no. 4, pp. 282–296, 2003.

[15] B. Wernerfelt, "A Resource based view of the firm," *Strateg. Manag. J.*, vol. 5, no. 2, pp. 171–180, 1984.

[16] K. M. Eisenhardt and J. a. Martin, "Dynamic capabilities: what are they?," *Strateg. Manag. J.*, vol. 21, no. 10–11, pp. 1105–1121, Oct. 2000.

[17] G. A. Alsos, O. J. Borch, E. Ljunggren, and E. L. Madsen, "Dynamic Capabilities – Conceptualization and Operationalization," in *The Academy of Management Conference*, 2008, pp. 1–28.

[18] V. Ambrosini and C. Bowman, "What are dynamic capabilities and are they a useful construct in strategic management?," *Int. J. Manag. Rev.*, vol. 11, no. 1, pp. 29–49, Mar. 2009.

[19] P. Oliveira, A. V Roth, and W. Gilland, "Achieving competitive capabilities in e-services," *Technol. Forecast. Soc. Chang.*, vol. 69, pp. 721–739, 2002.

[20] M. J. Leiblein, "What do resource- and capability-based theories propose?," *J. Manage.*, vol. 37, no. 4, pp. 909–932, 2011.

[21] M. Augier and D. J. Teece, "Dynamic Capabilities and the Role of Managers in Business Strategy and Economic Performance," *Organ. Sci.*, vol. 20, no. 2, pp. 410–421, 2009.

[22] D. a Griffith and M. G. Harvey, "A Resource Perspective of Global Dynamic Capabilities," *J. Int. Bus. Stud.*, vol. 32, no. 3, pp. 597–606, 2013.

[23] G. D. Bhatt and V. Grover, "Types of Information Technology Capabilities and Their Role in Competitive Advantage: An Empirical Study," *J. Manag. Inf. Syst.*, vol. 22, no. 2, pp. 253–277, 2005.

[24] A. Lindgreen and F. Wynstra, "Value in business markets: What do we know? Where are we going?," *Ind. Mark. Manag.*, vol. 34, no. 7 SPEC. ISS., pp. 732–748, 2005.

[25] F. Ahmadi Zeleti and A. Ojo, "Open data value capability architecture," *Inf. Syst. Front.*, 2016.

[26] F. Ahmadi Zeleti and A. Ojo, "Capability Model for Open Data: An Empirical Analysis," in *11th International Conference on Theory and Practice of Electronic Governance*, 2018.

[27] C. E. Helfat and M. a. Peteraf, "The dynamic resource-based view: capability lifecycles," *Strateg. Manag. J.*, vol. 24, no. 10, pp. 997–1010, Oct. 2003.

[28] S. Fainshmidt, A. Pezeshkan, M. L. Frazier, A. Nair, and E. Markowski, "Dynamic Capabilities and Organizational Performance: A Meta-Analytic Evaluation and Extension," *J. Manag. Stud.*, vol. 53, no. 8, 2016.

[29] G. D. Bhatt and A. F. Emdad, "An analysis of the virtual value chain in electronic commerce," *Logist. Inf. Manag.*, vol. 14, no. 1, pp. 78–85, 2001.

[30] F. Ahmadi Zeleti and A. Ojo, "Critical Factors for Dynamic Capabilities in Open Government Data Enabled Organizations," in *Proceedings of the 17th International Digital Government Research Conference on Digital Government Research*, 2016, pp. 86–96.

[31] F. Ahmadi Zeleti and A. Ojo, "Competitive Capability Framework for Open Government Data Organizations," in *Proceedings of the 18th Annual International Conference on Digital Government Research*, 2017, pp. 250–259.

[32] T. Fonseca, "Combining Product and Process Innovation: Is Organizational Innovation the crucial complement?," in *DRUID Academy conference in Rebild*, 2014.

[33] C. Lin and H.-L. Tsai, "Achieving a firm's competitive advantage through dynamic capability," *Balt. J. Manag.*, vol. 11, no. 3, pp. 260–285, 2016.

[34] L. Liu and Z. Jiang, "Influence of technological innovation capabilities on product competitiveness," *Ind. Manag. Data Syst.*, vol. 116, no. 5, pp. 883–902, 2016.

[35] A. Cristiano, S. Melo, M. Aparecida, C. Netto, V. J. Martins, F. Filho, and E. Fernandes, "Knowledge Management for Improving Business Processes: an Analysis of the Transport Management Process for Indivisible Exceptional Cargo," *Pesqui. Operacional*, vol. 30, no. 2, pp. 305–330, 2010.

[36] K. Ashogbon, "Achieving Business Process Improvement Through Knowledge Management," HAMK University of Applied Sciences, 2012.

[37] Gartner Research, "Six Building Blocks for Creating Real IT Strategies," 2002.

[38] F. Wu, S. Yeniyurt, D. Kim, and S. T. Cavusgil, "The impact of information technology on supply chain capabilities and firm performance: A resource-based view," *Ind. Mark. Manag.*, vol. 35, no. 4, pp. 493–504, 2006.

[39] S. Greener, *Business Research Methods*. Ventus Publishing ApS, 2008.

[40] C. Boyce and P. Neale, "Conducting in-depth interviews: A Guide for designing and conducting in-depth interviews," 2006.

[41] Y. Lu, "Concept Hierarchy in Data Mining: Specification, Generation and Implementation," SIMON FRASER UNIVERSITY, 1997.

[42] Social Science Data and Software, "Using NVivo for Qualitative Data Analysis," vol. 1. Stanford University, Stanford, pp. 7–7, 2011.

[43] A. H. Hilal and S. S. Alabri, "Using NVivo for Data Analysis in Qualitative," *Int. Interdiscip. J. Educ.*, vol. 2, no. 2, pp. 181–186, 2013.

[44] N. Helgevold and V. Moen, "Three Approaches to Qualitative Content Analysis," *Qual. Health Res.*, vol. 15, no. 9, pp. 29–42, 2005.

[45] P. Mayring, "Qualitative Content Analysis," Klagenfurt, 2014.

[46] B. C. Ozkan, "Using NVivo to Analyze Qualitative Classroom Data on Constructivist Learning Environments," 2004.

[47] N. Leech and A. Onwuegbuzie, "An array of qualitative data analysis tools: A call for data analysis triangulation," *Sch. Psychol. Q.*, vol. 22, no. 4, pp. 557–584, 2007.

[48] H. Noble and J. Smith, "Issues of validity and reliability in qualitative research," *Evid. Based. Nurs.*, vol. 18, no. 2, pp. 34–35, 2015.

[49] D. Horsburgh, "Evaluation of qualitative research," *J. Clin. Nurs.*, vol. 12, no. 2, pp. 307–312, 2003.

[50] C.-E. Lin, K. M. Kavi, F. T. Sheldon, K. M. Daley, and R. K. Abercrombie, "A Methodology to Evaluate Agent Oriented Software Engineering Techniques," *Syst. Sci. 2007. HICSS 2007. 40th Annu. Hawaii Int. Conf.*, p. 20, 2007.

[51] N. Prat, I. Comyn-Wattiau, and J. Akoka, "Artifact Evaluation in Information Systems Design Science Research - A Holistic View," *PACIS 2014 Proc.*, vol. Paper 23, pp. 1–16, 2014.

[52] T. H. Davenport, J. G. Harris, D. W. De Long, and A. L. Jacobson, "Data to Knowledge to Results: Building an Analytics Capability," 2001.

[53] R. Agarwal, E. Baltassis, J. Brock, and J. Platt, "Enabling Big Data: Building Capabilities that Really Matter," Boston, 2014.

[54] C. Dremel, S. Overhage, S. Schlauderer, and J. Wulf, "Towards a Capability Model for Big Data Analytics," *Proc. der 13. Int. Tagung Wirtschaftsinformatik*, pp. 1141–1155, 2017.

[55] J. Attard, F. Orlandi, and S. Auer, "Value Creation on Open Government Data," *2016 49th Hawaii Int. Conf. Syst. Sci.*, pp. 2605–2614, 2016.

[56] J. S. Marfo and R. Boateng, "Developing big data capabilities in developing countries: Evidence from a cross industry study in Ghana," in *Regional Conference of the International Telecommunications Society*, 2015.