Agile Mechanisms for Open Data Process Innovation in Public Sector Organizations: Towards Theory Building

Fatemeh Ahmadi Zeleti
Insight Centre for Data Analytics
National University of Ireland
Galway, Ireland
Fatemeh.AhmadiZeleti@Insight-Centre.org

Adegboyega Ojo
Insight Centre for Data Analytics
National University of Ireland
Galway, Ireland
adegboyega.ojo@deri.org

ABSTRACT

Process innovation in public organizations is widely documented and has increasingly been the subject of empirical scrutiny. However, no study has attempted to investigate process innovation in open data organizations in public sector. Guided by the Dynamic Capability Theory and based on the detailed study of four open data organizations, we synthesize a theoretical model and a process model for open data process innovation in public sector organizations. Specifically, the study sought to understand how open data process agility is achieved in these organizations. The results highlight the specific agile mechanisms that enable and improve open data process innovation in public sector organizations. The results also provide perspectives on how open data organizations in public sector can change data processes to transform the way they respond to changing demands and external environment.

CCS CONCEPTS

Information systems, **Model development and analysis, Reference models**, *Business rules*, *Business-IT alignment*, *Performance*, *Implementation management*

KEYWORDS

Process Innovation, Open Data, Open Data Processes, Data Processes, Public Sector Organization, Agile Mechanisms

ACM Reference format:

F. Ahmadi Zeleti, A. Ojo. 2019. Agile Mechanisms for Open Data Process Innovation in Public Sector Organizations: Towards Theory Building. In *Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance (ICEGOV2019), Melbourne, VIC, Australia, April 3-5, 2019*, 11 pages. https://doi.org/10.1145/3326365.3326387

ICEGOV2019, April 3–5, 2019, Melbourne, VIC, Australia

ACM ISBN 978-1-4503-6644-1/19/04…$15.00
https://doi.org/10.1145/3326365.3326387

1 INTRODUCTION

Public organizations are seen as active providers of information and input to their beneficiaries such as general public and private entities [1]. Therefore, public organizations must regularly adopt innovative mechanisms to improve their products and services in response to changes in the external environment, resource scarcity, and constant increasing expectations of end users [2][3]. Scholars have identified process innovation as one essential area for enabling organizational agility [3].

With the vast availability and improvement of open data and open government data, the demand for using and processing this data has increased in public sector to improve public value [4], which in turn leads to providing wide range of innovative open data products and services. To realize the real value of the data held in public sector and to address the fast changing expectations of end users, public organizations need to adopt mechanisms which improve data process agile.

Despite earlier studies that have sought to address process innovation in public sector and the existence of longstanding evidence base on the factors enabling process innovation [5], agile mechanisms for process innovation in public open data organizations are yet to be investigated. We currently have no theoretical and empirical explanations on how to enable open data process innovation or agility in public organizations [6].

Thus, this study seeks to make two contributions. First is to provide theoretical framework for our empirical investigation of four diverse public open data organizations. Diverse sampling enhances generalization in case study research [7]. For the first contribution, we commenced by discussing the Dynamic Capability Theory (DCT) as our theoretical lens to this case study research and present process innovation and its constructs. Based on the theoretical framework and, adopting within-case and cross-case analysis approaches to building theory from case studies [7][8], we analyzed our four cases. Second, we induced nine propositions which helped us identify constructs which are necessary and can be used for strong theory building [7].

Adopting definition of agility presented in [9] and [10], we define process agility as organization's ability to sense emerging needs and quickly match them with the available data processes to enable delivery of innovative data-driven products and services.

2 THEORETICAL UNDERPINNING

DCT has been an influential theoretical model for understanding how organizational capabilities can be created and refreshed in changing environments [11][12]. DCT suggests that, the possession and deployment of dynamic capabilities or mechanisms provide the organization with a chance to generate superior performance over the longer run. According to [13], dynamic capabilities allow organizations to renew and leverage their internal and external capabilities which in turn enhance its power in its global relationships, thereby enabling it to coordinate inter-organizational activities and respond rapidly, in a flexible manner, to global competitors' strategies. According to Teece et al. [14], DCT includes three fundamental unit of analysis namely *Process*, *Position*, and *Path/Strategies* available to the organiztaions. In this research, we focus on the Process.

Process is refered to as the way things are done in the organization [15]. The term process innovation encompases the implementation of change [16] in and significant restructuring/renewing [17] of existing processes or routines. Davenport [16] further reports that there exist number of drivers for process innovation such as improving efficiency of the developed products and services, lowering the cost of development and and production, reducing unnecessary spending, improve profitability, achieving competitive advantage [16], and improving organizational performance [18]. Process innovation allows organizations to improve or change the existing processes – specifically the processes to integrate, reconfigure, gain and release new resources [19][20]. Based on the well-known edicts of DCT, three constructs of process innovation include: *reconfiguration* which is the ability to sense the need to reconfigure, transform, and recombine existing processes; *integration* which is the ability to integrate and coordinate different processes; and *learning and leveraging* which is the ability to transform knowledge and learning acquired from the ecosystem [6] in order to perform activities better and quicker and, to extend existing processes by deploying it into a new domain [14][21].

In Figure 1, we present the theoretical framework for this study.

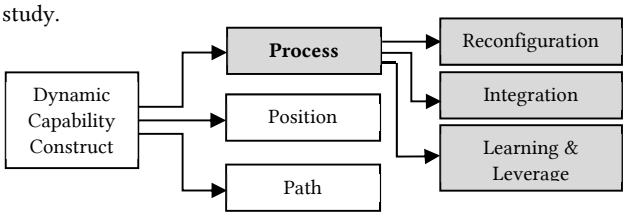

Figure 1. Theoretical framework based on DCT [14]

3 METHODOLOGY

Objectives - The objectives of this research is to investigate open data process innovation mechanisms in public sector based on empirical evidence collected through the in-depth interviews of four public open data organizations in the data ecosystem. In particular, the goals of this study are 1) to make sense of each individual case and identify categories as representing process innovation mechanisms in each investigated organization *(Within-Case analysis)* and 2) to elicit categories that are common among all organizations *(Across-Case analysis)* [22][23].

Research Method - One of the strength of qualitative research is its ability to illustrate the particulars of human experiences in the context of common phenomenon [22]. Following qualitative research method, we use DCT as our theoretical lens to study open data process innovation mechanisms in public sector.

Cross-Case Approaches to Data Analysis - For cross-case analysis, we adopt two approaches: *case-oriented* and *tactical 2x2* approach to cross-case analysis [7][24]. The first, looks for commonalities across multiple cases to build a general explanation that fits each of the individual cases, even though the cases will vary in their details [7]. This can help delineate the combination of agile mechanisms for open data process innovation and can contribute to generalization [24]. Moreover, within case-oriented approach, we adopt *multicase method technique* developed by Stake in 2006 [25]. According to Stake [25], multicase method is applied to cases which are similar in some ways (e.g. person, organization) and help the researchers find common issues across cases and contribute to generalization and theory building. The latter, looks for more rigorous analysis and sophisticated understanding of the cases [7] by conducting 2x2 analysis of every pair of cases to find out similar mechanisms between each pair. This tactic allows us to extend our results from the case-oriented approach and facilitate more rigorous analysis for theory building and generalization.

Data Collection and Case Selection - The choice of semi-structured interviews has been determined by the lack of existing work on open data process innovation which raises the need for in-depth exploration of the interviewee's perspectives on the topic under investigation [26]. Through the interview sessions, our goal is to explore agile mechanisms for open data process innovation which are implemented in open data organizations in public sector. The selection criteria include public organizations that: 1) rely on open data as one of their key operational resources to achieve organization's mission goals, 2) the application of open data is primarily in developing new products and services, and 3) has long history of practice and consider themselves mature and successful. In order to develop a study that is applicable across public organizations, we purposefully selected diverse organizations [7] filling the following categories: geographical location, sector (upstream and midstream), size, and data domain. Diverse sampling enhances the generalizability of the study [7]. We request to access the third party's dataset of 685 organizations. The reason for this selection is that, this dataset is the largest and fast growing dataset of organizations using open data around the word. Taking into consideration the above criteria, 43 organizations were shortlisted and contacted. We managed to conduct interviews of four organizations (Table 1) who have agreed to participate in this study. 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 two main parts: organizational background or context and open data process innovation activities of the organization.

Data Coding Process and Analysis - For our data analysis, we applied interpretive analysis that is also referred to as '*de-contextualization*' and *'re-contextualization'*. In qualitative case

study data analysis, data are de-contextualized when they are separated from individual cases in which they are originated into units of meaning through coding process and re-contextualized as they are reintegrated into categories or themes that represent units of like meaning. Re-contextualization creates a reduced data set drawn from across all cases [22]. We follow standard steps to qualitative data analysis [27][28] and adopt *Classical Content Analysis* as our technique to analyze qualitative data [29]. Through Classical Content Analysis, we allow new concepts or codes to emerge. For coding each data transcript, we used NVivo which is a strong and comprehensive qualitative data analysis software platform which can be used to organize and analyze any types of qualitative data [27][28] and to "obtain rigor in dealing with such data" [28]. Through the coding process, we 1) select a particular phrase, sentence, or paragraph of the text and 2) assign this fragment to a specific concept that can best describe the information being coded. This allows us to generate concept as the coding progresses. Any text that could not be placed to any developed concept would be given a new concept to be code to [23]. This increases trustworthiness as we make sure that we capture all possible variables or concepts from the transcripts which can allow us to better understand these capabilities [30].

4 ANALYSIS

Merriam [31] and Eisenhardt [7] called for multiple case studies to be performed in two steps; 1) within-case analysis followed by 2) cross-case analysis. In this section, we present the analysis of the data following the within-case and cross-case qualitative data analysis strategies. These choices were influenced foremost by the purpose of the study but also by specific data management and analysis practices prescribed by the method.

Table 1. Profile of the cases

Cases	Location	Size	Sector	Data Domain	Data Type	Data Application	Business Model	Revenue Stream	Clients
Case 1	Ireland	51-200	Up-stream	Environment	Open Data, Linked Data	New Product, Service, and Application, Optimization, Advocacy, Strategic	Supporting Primary Business	Government and Public Fund	Public Sector or B2G
Case 2	USA	11-50	Up-stream	All	Open Data	New Product, Service, and Application, Optimization, Advocacy, Strategic	Increasing Quality through Participation	Government and Public Fund	Private and Public Sector, Researcher, Developers, Citizens
Case 3	UK	1-10	Mid-stream	Business, Research, and Consulting	Open Data	Advocacy	Open Source	Government and Public Fund	Developers, Researchers, Citizens
Case 4	Spain	51-200	Mid-stream	Data Research and Innovation	Open Data	New Product, Service, and Application	Supporting Primary Business	Government and Public Fund	Public Sector or B2G

In section 5.1, we provide case by case analysis where we first illustrate the context and the dynamic nature of each case and, later, we present the categories which represent agile mechanisms for open data process innovation. Analysis in this section will provide the window to compare and give insight for the cross-case analysis in Section 5.2 where we aim to explore commonalities across cases by illustrating the mechanisms which are common among the investigated public organizations.

4.1 Within-Case Analysis

According to Eisenhardt [7], there is no standard format for within-case analysis. Different authors present their analysis differently such as using: graphs, tabular display, descriptions, and tables of information. In this study, we first describe each case and latter we present the analysis in a table. In this analysis, the overall idea is to become familiar with each case as a stand-alone entity.

Case 1 analysis

Context - Case 1 with over 10 years of experience in data cataloguing is a State agency responsible for marine research, technology development and innovation in Ireland. Case 1 carries out environmental, fisheries, and aquaculture surveys and monitoring programmes to meet Ireland's national and international legal requirements. Case 1 provides scientific and technical advice to Government to help inform policy and to support the sustainable development of Ireland's marine resource. Case 1 aims to safeguard Ireland's unique marine heritage through research and environmental monitoring. Case 1's research, strategic funding programmes, and national marine research platforms support the development of Ireland's maritime economy.

Dynamic Nature of the Organization - The agility in this organization is inspired by the view that organizational dynamism is not separate from its processes and process innovation capabilities. Moreover, analysis revealed that, this organization follows reactive approach to trends around open data and rather proactive in the area of dynamic capability adoption. This allows the organization to be dynamic in terms of their capabilities. By contrast, dynamic capabilities in particular process innovation capabilities in this organization can lead to reconfiguration in the industry value chain by being a key player in the upstream sector.

Open Data Process Innovation Capabilities - *Reconfiguration* – Case 1 data analysis revealed numbers of process innovation capabilities that facilitate process reconfiguration. According to the analysis, s*earching and acquiring for relevant knowledge from the ecosystem* is essential to identify potential process reconfiguration needs. To be capable of reconfiguring existing data processes, analysis supports the need to *adopt open data standards and follow European directives*; *acquiring knowledge around data management tools*; *acquiring a set of requirements from the open data industry* that can facilitate the development of prototypes; *acquiring knowledge about related best practices* around adding value to the data value chain and routines; and *acquiring knowledge about key stakeholders* (agencies, companies, academia and etc.) who can provide resource exchange required for process reconfiguration needs. In addition, data revealed that sometimes capabilities and other potential resources for the reconfiguration activities can be acquired from other work-in-progress R&D projects in the organization. As a matter of fact, *utilizing internally generated knowledge specifically through R&D activities* should not be neglected. In spite of the importance of internally generated knowledge, analysis uncovers the importance of *improving collaboration and communication with the product and service users* as another valuable source of knowledge to inform process reconfiguration.

Integration – Case 1 data analysis revealed numbers of process innovation capabilities that facilitate process integration and combination. Based on the analysis, *assessing and identifying series of existing best practices* is essential for this organization to integrate various data processes to attain efficiency and effectiveness in data processes and, to avoid possible risks associated with the integration efforts. Analysis further revealed that effective and contextually sound process integration can lead to effective process redesign, automation of data processes, and integration of services, improved data value chain, and the secure sharing of data across numerous applications. Moreover, to integrate or combine data processes, organization should be capable of *assessing and evaluating the many processes, platforms, and applications* that are considered to be coupled. Data suggests that, evaluation helps planning the integration and helps ensure that the integration is sensible and implementation and outcomes are as intended. Data also suggests that, *technical requirements of such integration effort must be taken into consideration* for developing a prototype of the combined processes. *'From pilot to production'* is the strategy to support piloting the resulted process to ensure effectiveness. Nevertheless, analysis disclosed that following more general *data standards* such as ISO19139 (Data Standardization) and ISO 19156 (Observation and Measurements) and, more specific standards such as open data standard for Inspire Directive, standards for Marine Strategy Framework Directive, and data standards for Water Framework Directive enable coherent synthesis of systems and processes.

Learning and Leveraging – Case 1 data analysis revealed numbers of process innovation capabilities that facilitate leveraging data processes. According to the analysis, this organization can leverage the existing processes and extend the application of their current processes by deploying them into a new domain if the organization has the ability to *improve data management processes*; *improve data cataloging processes* (using generic marine related Data Models); *improving linked data*; and link up and *connect to advanced data processes used by network agencies and businesses*. In addition, analysis suggests that, *new prototype of a product and service and, new initiatives* contribute to extending the data processes into a new domain. Initiatives such as Ireland's Marine Atlas, Irish Spatial Data Exchange, Ireland's Digital Ocean, Ireland's Marine Renewable Energy Portal and connecting to Ireland's open data Portal. *Adopting evaluated tools* such as ERDDAP to add value to marine data and enhance data cataloging is seen to be an important capability to determine whether or not a particular data process will meet its intended objectives and functions. Leveraging existing data processes and extending them into new domain may require data/open data to be largely available which should entail certain range of agreements that implicate *data protection and treatment*. To facilitate and support the increasing use of data by other data processes and applications, analysis suggests that *appropriate Data License and Standards to be established and in place*. Expanding the application domain in which data processes could be capable to be used for necessitates standards. Analysis shows that *developing new application that handles data processes standards to be evaluated* against different application domains is of great value to the organization to determine whether or not a particular process can successfully be extended to a particular domain.

Case 2 analysis

Context - Case 2 is a Department of a City in the United States. The government of the City exists to promote and sustain a superior quality of life in the City. In partnership with other communities, the goal of the City is to deliver cost-effective services in a personal, responsive and innovative manner. The City launched Open Data Portal in 2012, includes over 100 datasets and they are adding more all the time. The City encourages all stakeholders to explore these datasets to provide better services at the City level. For the City, the use of open data is not to meet the financial objectives. Open data value is not to meet monetary missions but rather to achieve and deliver greater transparency and trust.

Dynamic Nature of the Open Data Program and Initiatives - In general, the City is committed to a multi-year, bold vision to become a leading digital city. In particular, the open data movement and activities in the City aim to achieve transparency and trust. To achieve these goals and to improve agility in the open data programs and activities, the investigated City department sets strategic goals around 'Data Governance' and 'Infrastructure'. Following these goals, activities will continue to operate and City departments will always respond to rapidly emerging needs or changing priorities of the City.

Open Data Process Innovation Capabilities - *Reconfiguration* – Case 2 data analysis revealed numbers of process innovation capabilities that facilitate process reconfiguration. Data revealed that, an *ample use of knowledge management capabilities* of the organization is vital to properly generate and use insightful organizational knowledge to identify the needs for reconfiguration of processes. The data further shows that *identifying potential third party vendors* are important to reconfigure and improve process of data collection. The City believes that, engaging and partnering with third party organizations can significantly advance the various processes involve in the data value chain and reduce the possible risks associated with technical processes. Data suggests that, third party engagement can contribute to agility, improving value chain performance, and competitive advantage. In addition to the above capabilities, data revealed the importance of *availability of*

sufficient end user's communication channel and process in the City's open data portal to collect user's feedback. This assists the City to gain insights into the effectiveness of their data processes and to revisit existing processes for reconfiguration purposes.

Integration – In the case of the investigated City department in this study, analysis revealed that at the moment (current responsibility and obligation of the City department), the department is optimizing data processes separately.

Learning and Leveraging – Case 2 data analysis revealed numbers of process innovation capabilities that can help the City to perform activities better and quicker and, to extend existing processes by deploying it into a new domain. Data particularly suggests *developing and implementing citywide data governance model* based on a recognized, industry standard framework such as DMBOK. Data management is very critical to the City as the City deals with large volume of data from different sector and municipalities. Therefore, following this collection of best practices and standardization guides (DMBOK), the City can perform data management processes faster, in a more effective and consistent manner, and is able to extend a particular process into different datasets. Moreover, ability of the City to *standardize data storage approach and solutions* has been reported to be essential for data process optimization and improvement. Analysis further suggests that *utilizing a hybrid cloud approach* can significantly contribute to improving flexibility and speed in performing data related processes. *Implementing unified communications* and, *define and develop metrics* can help support process optimization. Adopting, unified communication within the organization can ease and improve data sharing and productivity which will result in quicker response to the needs.

Case 3 analysis

Context - Case 3 is a leading text and data mining company, with headquarters in Cambridge, England. Case 3 is specializing in building open source tools to enable clients to find facts hidden in information. Case 3 is now established as a leading supplier of text and data mining tools to both Higher Education and Knowledge based organizations, with a broad portfolio of clients. Case 3's values include: the liberation of scientific knowledge from public documents to make it useful for everybody, to support an open community that uses and promotes content mining, to create an Open toolbox of software, protocols and resources for mining, to work with partners to create better tools and support their knowledge extraction.

Dynamic Nature of the Organization - Case 3 organization is quite dynamic in terms of following the latest trends in the open data and data mining. Through consulting services and training workshops, Case 3 seeks for and identifies demands of their clients in order to address them faster than rivals. In addition, Case 3 organization is closely working and collaborating with open communities and other business partners which can improve organizational agility through acquiring knowledge and expertise.

Open Data Process Innovation Capabilities - *Reconfiguration* – Case 3 data analysis revealed numbers of process innovation capabilities that facilitate process reconfiguration. Analysis suggests that, the process of aggregating data from different data sources is vital for the day to day data mining operation of the organization. Therefore, it is essential that the organization *constantly assess the existing processes* to understand whether or not a particular process requires reconfiguration. Assessment informs further reconfiguration of the data processes. However, this suggests the organization to be alert for alternatives from the partners and the industry through a *knowledge acquisition process.* Data also revealed that *organizational ability to establish sufficient connectivity with various data sources and collect large amount of data* will inform data process reconfiguration. *Understanding the desire of end users* can also inform data process reconfiguration. An example form this organization is that, Data Publishing Process may need to be altered to address the end users needs and desire.

Integration – Case 3 data analysis revealed two process integration and combination capabilities. Data revealed that, *logical integration and synthesis of data model, data graph, data classification, and data visualization into a process* that can lead to production of more valuable scientific data and a better search mechanisms that can return as close content as possible. Data further suggests that, integration necessitates *process assessment effort* where the expert team assesses various processes activities to identify if they are close enough to be merged together. In this regard, analysis uncovers that not only processes joined in an integrated output, but sometimes integration is performed to address activities within a particular process to support the future activities of the organization.

Learning and Leveraging – Case 3 data analysis revealed numbers of process innovation capabilities that can help perform activities better and quicker and, to extend existing processes by deploying it into a new domain. Data suggests that, *following the established data standards and guidelines*, the organization is able to perform text mining, data extraction process, and source encoding process more smoothly and faster. Based on the data analysis, *sharing capabilities and resource allocation* can also contribute to improved process dynamism. Data suggests that, process innovation and application of data processes in a new domain require organizational resources (technical and nontechnical) and expertise. In this regard, *forming alliance* is vital in shaping the right processes. Alliance-based process development can provide the organization with a collaborative space to showcase their existing data processes and explore the potential of them. This can help inform possible and alternative applications of their processes in different domain.

Case 4 analysis

Context - Case 4 is a non-profit research and innovation centre based in Spain which promotes mission-oriented activities on advanced Internet architectures, applications and services. The centre stands up for a new open innovation framework, fostering the collaboration between companies, public administration, the academic environment and end-users. Case 4's goal is to develop advanced Internet technologies to the benefit of government, companies and citizens of Catalonia and the rest of the world. Case 4's research and innovation units want to produce technologies and solutions with the aim of converting Catalonia into a leading global smart region in a Smart Europe, with a flourishing added-value economy and an innovative society.

Dynamic Nature of the Organization - International relevance through organizational R&D activities incorporation with public and private sector plays a role in making the organization agile in terms of understanding their market and client's needs. Additionally, the organization relies on its key technological assets to follow action plans which set out a clear roadmap for the

use of research and technical developments to address the societal and market challenges and, strengthen the digital society. In this regard, close collaboration with institutions, companies, and organizations from other sectors and, taking advantage of the knowledge created by the R&D projects, this organization widens its horizon with the goal to share knowledge, expertise, and experiences in the field and, continues to increase its know-how regarding how to maintain organizational agility and uphold the impacts of its Innovation Unit.

Open Data Process Innovation Capabilities - *Reconfiguration* – Case 4 data analysis revealed two process innovation capabilities that facilitate process reconfiguration. Analysis revealed that, *exploiting knowledge management capabilities to access and analyze internally and externally generated knowledge* is essential to develop understanding about how the organization can reconfigure and optimize data processes. New configuration supported by processed knowledge can facilitate and inform design of new data product and service. Analysis further suggests that *continues screening of the potential market and partners* can lead to identification of more sufficient processes and tools. However, the organization should have the ability and management expertise to determine whether or not the new configuration serves the need of the organization and can be aligned with the projects needs. *Establishing and sustaining a good level of communication with the end users* can help validate whether the reconfiguration meets its goal or not.

Integration – Case 4 data analysis revealed numbers of process integration and combination capabilities. Data revealed that, to combine data processes, *prior knowledge of the activities* is required to inform classification and merge of similar activities. This is also referred to as *data process matching* which is a process where similar activities are put into a cluster. This organization adopts numbers of algorithms for clustering. Analysis also revealed that, complexity is associated with the implementation of clustering process therefore, *knowledge of the domain and the end user's needs and, clarity in the context* play a key role in merging the elements from two similar antecedent processes to create a single process that can be used to replace the original processes. In addition, analysis suggests that, the organization needs to *adopt proximity metrics* that take into account similarity of activities within processes. Data explains that possibility of *adopting a bottom-up approach to data process integration* where the expected value proposition (the offering) informs and promotes homogeneous process integration should not be overlooked. *Adopting the right practical methods and process matching technique* for integrating processes reported to lessen the complexity associated with the integration effort and help reveal contextual similarity. Data uncovers that, *identifying contextual similarity* is a key in process integration as it helps to see similarities between two processes.

Learning and Leveraging – Case 4 data analysis revealed numbers of process innovation capabilities that can help perform activities better and quicker and, to extend existing processes by deploying it into a new domain. This organization data analysis revealed that, having in place *the right data quality standards and standard evaluation process* can help support fast data processing and publishing. By ensuring the appropriateness of data processes, the quality of the outcome is checked against the functional and nonfunctional requirements. This promotes data process application in other projects, datasets, and data domain. In addition, data suggested that, *leveraging skillsets and automated data processes through partner-to-partner relationship or network (alliance)* can speed up the way the organization is responding to its target market and clients. Forming partnership and establishing network of partners endorse *sharing capabilities and resource allocation* which are required to addressing the organizational needs and client's demands.

In Table 2, we summarize the analysis and present the agile mechanisms for open data process innovation.

4.2 Across-Case Analysis

Within-case analysis facilitates cross-case comparison. To perform cross-case analysis, we first adopt Case-Oriented approach [24]. The central question of interest that we would like to answer following this approach is *in what ways the cases are alike* [24]. Guided by this question and following the main objectives of this research (agile mechanisms for open data process innovation), the special emphasis is only given to analyzing and presenting agile mechanisms (from Table 2) which are in common or alike across the four cases for *reconfiguration*, *integration*, and *leveraging*. The result of this analysis is presented in Table 3 (case-oriented). Next, to complement the case-oriented analysis, we perform 2x2 tactic to cross-case analysis [7] where we analyze every pair of cases and list down the similar mechanisms between each pair. This tactic allows us to complement our results from the case-oriented approach. The result of this analysis is presented in Table 3 (2x2 tactic).

Case-oriented approach to cross-case analysis

Reconfiguration – It is clear from the analysis that *Knowledge Acquisition* has been considered as one of the mechanisms for improving agility of open data processes in the organization. Analysis across the four cases suggests that, looking for, eliciting, and acquiring relevant knowledge from different sources such as network partners, other businesses, and the data market itself play a role in improving open data process innovation. Across-case analysis also suggests that, internally generated knowledge is also considered another essential source of knowledge that can contribute to open data process innovation. Despite the role that knowledge acquisition plays, cross-case analysis suggests that maintaining a good level of *Stakeholders Connectivity* does equally contribute to the data process reconfiguration effort.

Integration – As revealed by the cross-case analysis of the four cases, data shows that cases repeatedly highlight the significance of *Process Assessment* and *Contextual Clarity and Similarity* of the processes to be merged in supporting and improving integration effort of data processes.

Table 2. Agile mechanisms for open data process innovation - summary of within-case analysis

Case	Reconfiguration	Integration	Learning & Leveraging
1	Search and acquire relevant knowledge from the ecosystem; Adopt open data standards; Follow European directives; Acquire knowledge around data management tools; Acquire a set of requirements from the open data industry;	Assess and identify series of existing best practices; Assess and evaluate processes, platforms, and applications; Identify technical requirements for/and contextually sound integration; From pilot	Improve data management processes; Improve data cataloging processes; Improve linked data; Connect to advanced data processes used by network agencies and businesses; Implement

	Acquire knowledge about related best practices; Acquire knowledge about key stakeholders; Utilize internally generated knowledge specifically through R&D activities; Improve collaboration and communication with the product and service users	to production; Follow data standards to avoid risks	prototypes in a new initiative; Adopt evaluation tools; Ensure data protection and treatment; Appropriate Data License and Data Standards; Develop new application that evaluate data processes standards
2	Ample use of knowledge management capabilities; Identify potential third party vendors; Availability of sufficient end user's communication channel and process	*No mechanisms reported.* *Case 2 is assessing and optimizing data processes separately*	Develop and implement citywide data governance model; Standardize data storage approach and solutions; Utilize a hybrid cloud approach; Implement unified communications; Develop metrics
3	Constantly assess the existing processes; Knowledge acquisition from the partners and the industry; Establish sufficient connectivity with various data sources and collect large amount of data; Understand the desire of the end users;	Logical integration and synthesis of data model, data classification, and data visualization into a process; Process assessment effort prior to integration	Follow the established data standards and guidelines; Share capabilities and resource allocation; Alliance-based process development
4	Exploit knowledge management capabilities; Continues screening of the potential market and partners; Establishing and sustaining a good level of communication with the end users	Prior knowledge of the activities; Data process matching; Knowledge of the domain and the end user's needs; Clarity in the context; Adopt proximity metrics; Adopt a bottom-up approach to data process integration; Adopt the right practical methods and process matching technique; Identify contextual similarity	Adopt the right data quality standards and standard evaluation process; Leverage skillsets and automated data processes through partner-to-partner relationship or network (alliance); Share capabilities and resource allocation

Through process assessment mechanisms, an organization is able to focus on evaluating how the existing process achieves its goal and delivers the expected value as they were promised. The assessment output can inform and help prioritize the integration data process effort. Equally important is the contextual clarity of different processes and their similarities. This can foster a logical view to process integration to maximize efficiency.

Learning and Leveraging – The cross-case analysis of the four cases indicate that following *Standards* and *Standardizing* data processes can help maximizing compatibility of data processes and their application in other domains. Following standards helps organizations to run them smoothly and much quicker.

2x2 Tactic to cross-case analysis

Data analysis suggests that Case 1 is similar to Case 2 in implementing agile mechanisms like: acquiring relevant knowledge, improving efficient collaboration and communication with end users, data process assessment and evaluation, and adopting and following appropriate standards around data processes; similar to Case 3 in acquiring knowledge from other partners and ecosystem, evaluating and assessing processes prior to data process integration effort, and following standards in order to avoid risk associated with data integration and processing; similar to Case 4 in exploiting knowledge management capabilities to acquire internally and externally generated knowledge, establishing and improving a good level of communication with end users, understanding technical requirements (e.g. proximity metrics, matching techniques), developing contextual similarities (between processes) for sound and coherent process integration, and adopting standards.

Case 2 is similar to Case 3 in implementing the following agile mechanisms: ample use of knowledge management capabilities to acquire knowledge from partners and the data industry, data process assessment prior to integration effort, and adopting standards; similar to Case 4 in acquiring relevant knowledge from the ecosystem and exploiting knowledge management capabilities, improving and sustaining connection and communication with end users, and searching for and identifying potential business and industry partners.

Case 3 is similar to Case 4 in understanding and developing contextual similarities (between processes) for sound and coherent process integration, forming alliance-based process development which can contribute to learning and extension of processes to other application domains, sharing capabilities and resource allocation, and following standards.

Table 3. Results: Case-oriented and 2x2 Cross-Case Analysis

	Reconfiguration	**Integration**	**Learning & Leveraging**
Case Oriented	Knowledge Acquisition	Process Assessment	Standards and Standardization
	Stakeholders Connectivity	Contextual Process Analysis	×
2x2 Tactic	Knowledge Acquisition	Process Assessment	Standards and Standardization
	Stakeholders Connectivity	Contextual Process Analysis	Capability Sharing and Resource Allocation
	Partner Identification	Technical Requirements Identification	Alliance-Based Processes

5 DISCUSSION

Following the within and cross-case analysis, tentative constructs and their relationships to enabling process agility or process innovation in open data organizations emerged (Figure 2). The emergent relationships fit with the evidence in the cases. Each relationship is confirmed at least by two cases. At this point, the qualitative data analysis presented is useful for providing a good understanding why emergent relationships hold. According to [7] and [8], this is a crucial step towards internal validity.

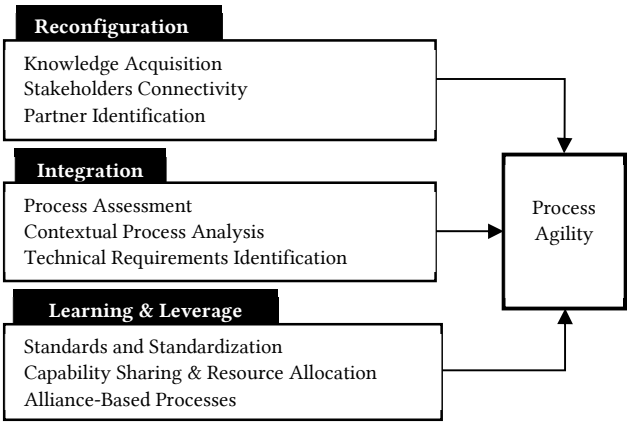

Figure 2. Theoretical model for process agility in public open data organizations

Discussion around data process agility in public organizations has been driven largely by the agility mechanisms presented in Table 3. All four cases represent interesting examples of how process agility is achieved and improved in public sector. Even though each case is unique by its own, commonalities found in the four cases help us synthesize and understand how process agility can be achieved in this particular sector.

Data Process Reconfiguration - One common feature found in all the four cases is the ability of the organization to acquire new knowledge for process reconfiguration effort. Knowledge is either created internally or collected/acquired from external sources (partners, projects, and etc.). Internal and external knowledge needs to be distributed and transformed into valuable asset. [32] appears to view knowledge acquisition and management as a mechanism to improve efficiency in public sector because it is linked to saving resources and most importantly providing product and service innovation opportunities. While all the four cases show that knowledge acquisition in public sector is mainly used to inform the need for process reconfiguration, each of these organizations aim to use new knowledge to generate different organizational assets. Case 1 adopts this mechanism to reconfigure processes to achieve faster data process related *decision making*. According to [32] and [33], knowledge facilitates faster decision making. This is because the experience of the organization is recorded and this helps both to avoid mistakes and discover and use the best solutions to similar issues saving time. Case 2 adopts this mechanism to reconfigure processes which can eventually lead to *improving transparency and trust*. In public sector, transparency is still seen as the main outcome of process innovation [34][35]. Case 3 adopts this mechanism to reconfigure processes to unleash and *employ alternative solutions* to address demands; and Case 4 adopts this mechanism to reconfigure processes to *attract and obtain funding opportunities*. We argue that, each of these different but unique features lies in the characteristics (e.g. aim and mission goal, size, products and services and etc.) of each organization. However, based on the four cases, new knowledge acquisition can inform and facilitate process reconfiguration and enable process agility. This leads us to the following proposition:

PROPOSITION 1. Public organizations acquire knowledge for process reconfiguration to enable process agility.

Another common feature found is the ability of the organization to maintain ongoing end users connectivity and stakeholder's engagement. According to [33], providing more data to users will increase their engagement level and expectations. In addition, [36] has also claim that engaging end users in the process development and assessment as early as possible can promote *'open design'* and encourage participation. However, the key to stakeholders' engagement is the availability of appropriate communication channel and engagement process model. All our cases have confirmed that end users are valuable source of knowledge and information which should be taken into consideration for process reconfiguration efforts. [33] has further claim that, desire and expectations that users raise can be used for assessing data and related processes. Feedback system, data request system (Case 2), and users' stories (Case 1) are the few potential engagement models for maintaining users' engagement for optimal outcome. While one-sided initiatives which are very supply-side driven are very discouraging, more collaborative initiatives where users are allowed to propose a new dataset and solution can encourage participation and knowledge discovery and delivery [33][34]. Therefore, we claim that it is very important to put forward initiatives and programs that advance the end users engagement and participation. This leads us to the following proposition:

PROPOSITION 2. Public organizations employ ongoing stakeholder engagement to enable process agility.

The third common feature found is the ability of the organization to identify both research and industry partners who can provide support for process reconfiguration. Cases claim that knowing the key stakeholder/player can allow the organization to seek help in many data related process activities such as data process reconfiguration. Potential and supportive partners bring the resources and expertise that are required to perform the activity. From the cases, we can claim that there exist two main reasons why public sector organizations seek for partnership (public and private): 1) the organization has already reached maturity and is looking for potential partners to advance productivity, improve agility, achieve collaborative advantage, quickly addressing multi-faceted societal problems by implementing innovative data processes [4], increase in-house expertise by sharing resources and capabilities, benefit funding, and in general to up-scale (Case 1 and 4); 2) to attain goals beyond the capability and capacity of the organization in order to successfully govern and achieve public service delivery in a timely manner (Case 2 and 3). This leads us to the following proposition:

PROPOSITION 3. Public organizations identify partners to support the process agility effort.

Data Process Integration - The fourth common feature found is the ability of the organization to perform process assessment to evaluate the existing processes against best-in-the-class processes and, efficiency and ability to respond to the changing demands/needs of the end users. Stakeholders specifically end users' feedback and input is essential for the team to continue to assess and iterate the data process evaluation and re-design. [35] suggests that feedback mechanisms and systematic end users engagement can better facilitate organizations to assess data and its related processes. This rapid launch-and-improve/improve-and-relaunch process should become the organization's core approach to data process agility [36]. In

addition, from the cases, we can see that process assessment is particularly important to minimize possible risks (costs, effort, time, and resources) that may be associated with the process integration. This is specifically true in the big organizations (Case 1 and 4) where distribution of resources and funds are carefully monitored. One approach to manage risks and implement coherent integration is to develop pilot before the final launch of the new process. Moreover, in bigger organizations, process assessment can advise the technical and non-technical requirements for process integration which is more *aligned with the projects needs of the organization.* Nevertheless, in smaller organizations, we can see that, process assessment is seen sufficient if done to optimize the existing processes to allow *addressing end users demand in a timelier manner* rather than performing any unnecessary integration activities. Regardless of the specific outcome, process assessment helps process agility. This leads us to the following proposition:

PROPOSITION 4. Public organizations perform process assessment for process agility.

The fifth common feature found is the ability of the organization to perform contextual process analysis. Contextual process analysis allows identifying similarities between different processes and activities within each process. This mechanism is essential for making a sound and rational integration. In this regard, [37] claim that, contextual similarity is important for process integration and it can also lead to design of new and removing redundant process. Case 1 and Case 4, as the two big public sector organizations see the importance of contextual clarity. For this, they implement relevant matching techniques and methods which might be costly. We argue that, this is due to the fact that any possible risks must be diminished before the final implementation. Even though we have not seen this in the cases but, we argue that, before analysis, contextual factors and issues should be well understood and identified for a successful integration. Contextual factors play a role in logically mix and match different processes into a more meaningful, sound, reliable, and fast process. We further argue that, the need to identify technical requirements is very well aligned with the need to identify and develop these factors for successful contextual process analysis. This leads us to the following proposition:

PROPOSITION 5. Public organizations perform contextual process analysis for process agility.

The sixth common feature found is the ability of the organization to identify and develop technical requirements for process integration effort. Prior discussions (assessment and contextual analysis) lead to more understanding around this mechanism and note the importance of technical requirements identification for improving process agility. Similar to other areas discussed, we argue that big organizations seem to identify technical requirements by adopting technical techniques and methods (mapping functions, proximity metric, cluster-based similarity, scoring). The resources (human, technical, intellectual) generated in big organizations are superior and difficult to replicate [38] compared to smaller ones. So they have a bigger chance to cope with the changing demands quickly. In addition, taking a closer look at how big public sector organizations (Case 1 and 4) operate, we can see that, most of the times they have the support of the government and funding agencies. These factors to our view can positively influence their ability to adopt more advanced tools for configuration and testing the technical requirements for speedier process integration. However, this gap can certainly limit the technical activities of smaller organizations (Case 3) but, gives them the opportunity to focus on and master few application areas [6]. This leads us to the following proposition:

PROPOSITION 6. Public organizations identify technical requirements for process agility.

Data Process Learning and Leveraging - The seventh common feature found is the ability of the organization to follow available standards such as W3C standards. In all the four cases, adopting standards help the organizations to extend current data processes to different data domain and application areas. [35] suggest that, developing standards assist bigger organizations to recognize and measure value creation which are very complicated practices. This view of Bertot et al. [35] corresponds with this study which shows that big and upstream organizations develop standards and engage in standardization efforts more than smaller organizations who are mostly adaptors. According to [4], standards help connect components of the data processes and allow organizations to figure out how they interrelate. This leads us to the following proposition:

PROPOSITION 7. Public organizations adopt standards for process agility.

The eighth common feature found is sharing process capability and resources to facilitate learning, doing things faster, and extending processes to different application areas. On the one hand, this mechanism received less recognition in upstream organizations (Case 1 and 2). We argue that this may be linked to the institutional attributes of upstream organizations, strategic decisions, data protection, and rigid regulations employed. In upstream organizations, actors or stakeholders who want to access or share resources (Data) must be well framed [39]. We further argue that, relationship, capability, and resource sharing require sufficient level of openness which require upstream organizations to introduce legal procedures [4]. On the other hand, in midstream organizations (Case 3 and 4) where intermediaries fit, sharing process capabilities and resources are essential to maximize organizational ability to achieve process agility. Identifying the right partner can improve the chances of appropriate resource allocation [4]. Resource allocation allows this category of organization to find the primary sources of valuable data. According to [4], sharing and accessing adequate data resources and capabilities enter different partners into an exchange relationship for developing data processes which appears to be more stable and sustainable. However, there is a need for adequate network mechanisms and resource (data) sharing approaches to facilitate sharing capabilities and resource allocation for process agility. This leads us to the following proposition:

PROPOSITION 8. Public organizations share process capabilities and allocate resources for process agility.

The ninth common feature found is the ability of the organization to form alliance with the potential and supportive research and industry partners to develop a more sustainable and agile data processes. According to [35], forming a network of

partners or alliance help produce stronger agenda for developing data processes which are more stable and agile as well as improve transparency and trust between partners as the result of collaborative effort. In the field of Strategic Alliance, NASA Johnson Space Center [40] highlights the importance of creating the win-win strategy and forming alliance to benchmark data processes. As our cases suggest, Midstream organizations adopt this mechanism as they need to maximize ability to improve quality of the process innovation and enhance their portfolio by supplementing internal core capabilities and resources with external capabilities. This leads us to the following proposition:

PROPOSITION 9. Public organizations adopt alliance-based processes for process agility.

Figure 3 builds on the empirical evidence collected from the cases and, authors observations and related literature presented in the discussion. Square blocks represent the nine mechanisms for open data process agility and, the alignments and emergent relationships represent the flow between each two mechanisms.

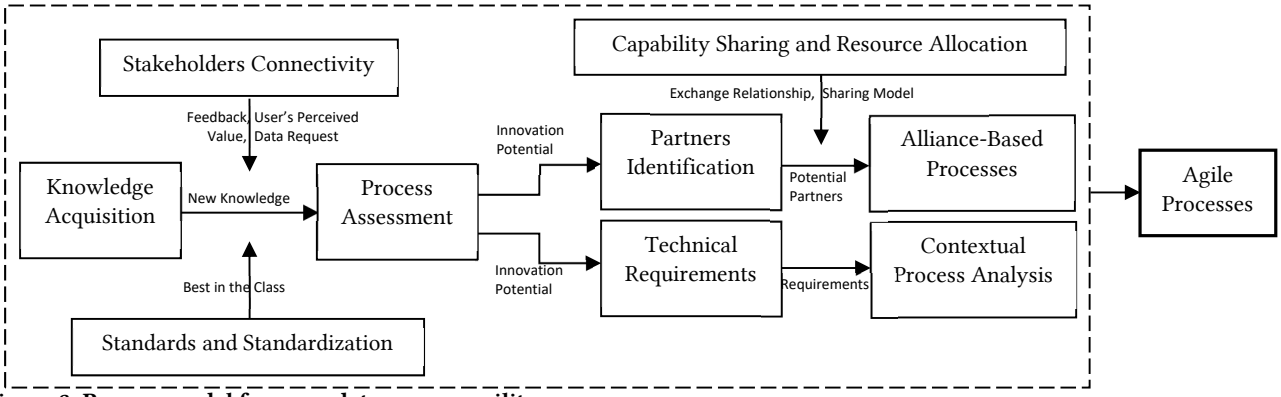

Figure 3. Process model for open data process agility

6 CONCLUSION

Many organizations in the public sector utilize open data to develop wide range of data-driven products and services which can ultimately improve public value. While open data is maturing, attention is shifting to its value chain and associated data processes for example, data collection and publication processes; specifically how fast such processes can change to address the dynamic landscape of the external environment and fast-changing users' expectations.

Guided by the DCT and based on the empirical evidence, this study cumulates agile mechanisms for open data process innovation in public sector. The study further synthesizes a theoretical model and a process model for open data process innovation. Our result showed that, even though organizations vary in their details, agile mechanisms adopted for open data process innovation remain alike in successful open data organization in public sector.

Our contribution in this study is two-fold. First, adopting the DCT as our theoretical lens and our studies of diverse cases, we contributed to the existing literature by providing a refined theoretical model for process innovation that serves as the base for generalization and fresh theory that bridges well from rich qualitative evidence to mainstream deductive research. Second, by inducing nine measurable propositions, this study contributes to establishing empirically-verifiable theory about the open data process innovation in public sector and in general. The results reported in this paper provide empirical evidence to support the propositions.

Moreover, this research contributes to the existing knowledge of the open data organizations in public sector by providing specific agile mechanisms and process model for enabling and improving open data process innovation. It also provides perspectives on how open data organizations in public sector can change data processes in a timely manner in order to transform the way they respond to changing demands and external environment.

We anticipate that the future research will 1) quantify our model to assess the strength of each relationship and 2) apply the model to a larger population.

REFERENCES

[1] P. Koch and J. Hauknes, "Innovation in the Public Sector," Oslo, 2005.
[2] R. Lenart, "Relational Capital As An Instrument Of Increasing Competitiveness," in *The 8th International Management Conference "Management Challenges For Sustainable Development,"* 2014, pp. 14–26.
[3] R. M. Walker, "Internal and External Antecedents of Process Innovation: A review and extension," *Public Manag. Rev.*, vol. 16, no. 1, pp. 21–44, 2014.
[4] F. van Schalkwyk, M. Willmers, and M. L. McNaughton, "The roles of intermediaries in an open data ecosystem," *J. J. Inf. Technol. Dev.*, 2015.
[5] L. B. Mohr, "Determinants of Innovation in Organizations," *Am. Polit. Sci. Rev.*, vol. 63, no. 2, pp. 11–26, 1969.
[6] 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.
[7] K. M. Eisenhardt, "Building Theories from Case Study Research," *Acad. Manag. Rev.*, vol. 14, no. 4, p. 532, 1989.
[8] K. M. Eisenhardt and M. E. Graebner, "Theory building from cases: Opportunities and challenges," *Acad. Manag. J.*, vol. 50, no. 1, pp. 25–32, 2007.
[9] C. Ambrose and D. Morello, "Designing the Agile Organization: Design Principles and Practices," 2004.
[10] E. Overby, A. Bharadwaj, and V. Sambamurthy, "Enterprise agility and the enabling role of information technology," *Eur. J. Inf. Syst.*, vol. 15, no. January, pp. 120–131, 2006.
[11] M. J. Leiblein, "What do resource- and capability-based theories propose?," *J. Manage.*, vol. 37, no. 4, pp. 909–932, 2011.
[12] 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.
[13] 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.
[14] D. Teece, G. Pisano, and A. Shuen, "Dynamic capabilities and strategic

management," *Strateg. Manag. J.*, vol. 18, no. 7, pp. 509–533, 1997.
[15] D. Teece and G. Pisano, "The dynamic capabilities of firms: An introduction," *Ind. Corp. Chang.*, vol. 3, no. 3, pp. 537–556, 1994.
[16] T. H. Davenport, *Process Innovation: Reengineering Work Through Information Technology*. Boston, MA: Ernst & Young, 1993.
[17] J. E. Ettlie and E. M. Reza, "Organizational Integration and Process Innovation," *Acad. Manag. J.*, vol. 35, no. 4, 2017.
[18] G. S. Kenfac, S. Nekoumanesh, and M. Yang, "Process Innovation: Impacts on Organization's Performance - A Qualitative Study of Four Swedish Municipalities," Linnaeus University, 2013.
[19] F. Ahmadi Zeleti and A. Ojo, "Capability Matrix for Open Data," in *15th IFIP working conference on virtual enterprises*, 2014.
[20] B. Verworn and C. Herstatt, "The innovation process : an introduction to process models," 12, 2002.
[21] D. J. Teece, "A dynamic capabilities-based entrepreneurial theory of the multinational enterprise," *J. Int. Bus. Stud.*, vol. 45, no. 1, pp. 8–37, 2014.
[22] L. Ayres, K. Kavanaugh, and K. A. Knafl, "Within-Case and Across-Case Approaches to Qualitative Data Analysis," *Qual. Health Res.*, vol. 13, no. 6, p. 831, 2003.
[23] N. Helgevold and V. Moen, "Three Approaches to Qualitative Content Analysis," *Qual. Health Res.*, vol. 15, no. 9, pp. 29–42, 2005.
[24] S. Khan and R. VanWynsberghe, "Cultivating the Under-Mined: Cross-Case Analysis as Knowledge Mobilization," *Qual. Soc. Res.*, vol. 9, no. 1, 2008.
[25] R. E. Stake, *Multiple Case Study Analysis*. 2006.
[26] C. Boyce and P. Neale, "Conducting in-depth interviews: A Guide for designing and conducting in-depth interviews," 2006.
[27] Social Science Data and Software, "Using NVivo for Qualitative Data Analysis," vol. 1. Stanford University, Stanford, pp. 7–7, 2011.
[28] 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.
[29] 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.
[30] P. Mayring, "Qualitative Content Analysis," Klagenfurt, 2014.
[31] S. B. Merriam, *Qualitative research and case study applications in education*, 2nd ed. San Francisco: Jossey-Bass, 1998.
[32] S. A. Theocharis and G. A. Tsihrintzis, "Knowledge Management Systems in the Public Sector: Critical Issues," *Lect. Notes Softw. Eng.*, vol. 4, no. 1, 2016.
[33] M. SalehAltayar, "Motivations for open data adoption: An institutional theory perspective," *Gov. Inf. Q.*, vol. 35, no. 4, 2018.
[34] I. Mergel, A. Kleibrink, and J. Sörvik, "Open data outcomes: U.S. cities between product and process innovation," *Gov. Inf. Q.*, 2018.
[35] J. C. Bertot, P. McDermott, and T. Smith, "Measurement of Open Government: Metrics and Process," in *45th Hawaii International Conference on System Sciences*, 2012.
[36] K. H. Tan and Y. Zhan, "Improving new product development using big data: a case study of an electronics company," *R&D Manag.*, vol. 474, pp. 570–582, 2017.
[37] E. D. Morrison, A. Menzies, G. Koliadis, and A. K. Ghose, "Business Process Integration: Method and Analysis," in *Sixth Asia-Pacific Conference on Conceptual Modelling*, 2009.
[38] B. Wernerfelt, "A Resource based view of the firm," *Strateg. Manag. J.*, vol. 5, no. 2, pp. 171–180, 1984.
[39] M. Mcleod and M. Mcnaughton, "Mapping an emerging Open Data ecosystem," vol. 12, pp. 26–46, 2016.
[40] NASA Johnson Space Center, "Strategic Alliances Strategies and Processes Benchmarking Study," 2009.