A GENERIC PROVENANCE FRAMEWORK TO DOCUMENT PUBLIC POLICY MAKING PROCESSES

Barkha Javed

This thesis is submitted in partial fulfilment of the requirements of the University of the West of England, Bristol for the degree of Doctor of Philosophy

Faculty of Environment and Technology,
University of the West of England,
Bristol

(2020)

DECLARATION

I declare that this thesis presented by me entitled ‘A process-agnostic and a model-driven based public policy making provenance framework’ is my own investigation and has not previously been submitted for a degree or similar award at the University of the West of England or any other institution. To the best of my knowledge and belief, no material in this thesis has been previously published or written by another person, except where due reference is made.

Barkha Javed

Abstract

Public policies impact the day-to-day activities of individuals. Effective public policy outcomes result in general acceptance in the community. The transparency in policy making process and participation during policy creation holds significant positions for developing trust among the community. Established domains such as e-health employs provenance for creating transparency and trust among the researchers. Public policy making can also use provenance to develop trust and transparency in their processes. At present, however, public policy makers employ various means to manage public policy making. Having no unified platform for the policy making process presents challenges in respect of searching and locating the evidence that was used during policy creation and for ensuring trust and transparency among actors. The absence of such a support also presents challenges for participation in public policy making. To address the given challenges, this research presents the provenance framework that manages the public policy making provenance data and enable participation of diverse actors.

Due to dynamicity attached to public policy making, a provenance framework needs to be adaptable. Therefore, a model-driven approach has been used to frame the public policy making provenance framework. In addition to a model-driven approach, a mechanism is required that can enable the capture of public policy processes. However, the knowledge-intensive dynamics of public policy making presents challenges for using process-based solutions. Therefore, this research work describes a process-agnostic approach inspired from a network-based packet switching approach for tracking policy making processes. Managing public policy provenance data is not the only facet that develops trust. What is required is the facilitation of citizens and non-government bodies in the policy creation process. Therefore, a provenance framework has been designed by considering the principles of smart governance which results in a smart cities solution. In order to evaluate the framework, a proof-of-concept has been designed and implemented. An evaluation has been carried out to determine the suitability of a model-driven and a process-agnostic approach for policy making provenance framework in smart cities. For the evaluation purposes, three public policy making case studies *Shops Opening Hours' Extension Policy, Air Quality Improvement Policy, and Local Planning Policy in Decision-making* were employed. The three case studies were used to derive various experiments to test the provenance framework. The experiments captured the dynamic and knowledge-intensive aspects of the provenance framework. The results collected from the execution of the experiments demonstrated the aptness of a process-agnostic approach and model-driven approach for the policy making provenance framework. Lastly, an end user evaluation was carried out to assess the effectiveness of the provenance framework. The positive responses of end users showed the usefulness of the provenance framework.

Acknowledgement

I am very thankful to my loving parents for supporting me all the way. Their love, support, and encouragement have always been a source of inspiration. I am very grateful for their motivation. I would like to pay my deep gratitude to my lovely sisters for their love and support.

Many special thanks to my supervisors Dr. Richard Mcclatchey and Dr. Zaheer Khan who have always inspired me with their knowledge and dedication to work. I would like to appreciate their encouragement, support, motivation, patience, and guidance that remained with me throughout my research work. I am very grateful for their valuable feedback and time. I would like to thank David Ludlow for guiding me throughout my PhD journey. I thank and appreciate the valuable input of interviewees in my research. I would like to give thanks to Dr. Lindsey Gillies and Dr. Tim Brailsford for helping me, especially, during the last days of my PhD. I am grateful to colleagues at UWE graduate school for their assistance, Marisa, Helen Mulligan, Patricia, Yasmin, Helen Frisby, Claire, Samantha, and Helen Jackson.

Finally, I pay my profound gratitude to my friends Daniel, Punsisi, Millicent, Pastor Jacob, Kayo, Priyanka, Ameenah, Theresa, Daisy, Tina, Shahista, Noura, Eman, and Hamzeh for encouraging me and helping me in difficult times. They have been great source of motivation and backing in my PhD journey. Special thanks to Sandeep, Aunty Gifty, Wayne, and Pastor Eric for their constant support and motivation.

Table of Contents

List of Figures

List of Tables

Glossary

A

Actors represent those who are directly or indirectly involved in public policy making.

Actors' Involvement captures the different ways in which actors participate in public policy making.

Active Involvement is a type of actors' involvement in which actors actively contributes to the formulation of public policies.

Ad-hoc in this thesis corresponds to a convoluted public policy making process which is guided and influenced by the actors.

Agenda Setting Phase is a phase of generic policy cycle that captures the issue, root cause of issue, and government's decision on placing a policy on their agenda.

Asynchronous Communication is a type of communication patterns that is being used by actor to communicate with each other. Using asynchronous communication, sender can continue to work while awaiting the response from the receiver.

B

Blockchain is a distributed ledger technology which provides secure transactions without involving centralised trust party.

Broadcast refers to a situation where more than one actor is communicated at the same time in a policy making process

Broadcast-inform is one of the modes of broadcast. In this mode, multiple actors are provided the information simultaneously without waiting for their response.

Broadcast-Asynchronous sends asynchronous request to multiple actors

Broadcast-Synchronous sends synchronous request to multiple actors

Business Process is a set of tasks to achieve common business goals.

C

Case Management is a paradigm that has specifically been designed to support Knowledge-intensive Processes.

Citizens are actors representing residents.

Communication explains the interaction among policy making actors which is carried out to formulate public policies.

Consultation represents a scenario when multiple actors discuss a common point in a policy and presents arguments and counter arguments to reach a common ground.

Constructs specify the metrics against which evaluation of the policy cycle provenance framework has been carried out.

D

Decision-makers here refers to those who are directly involved in public policy creation and approval.

Decision-making Phase is a generic policy cycle phase where only decision-makers participate in order to decide whether the policy should commence to the implementation phase.

Delay in a public policy making process incurs when completion time of a task exceeds the expected task completion time.

Documents are the artefacts which are used and generated during public policy making processes.

Dynamic depicts the run-time changes in public policy making processes.

E

Evaluation Phase is the generic policy cycle phase that captures the process for assessing the implemented policies

Evidence-Based Policy making place emphasis on rational public policy decision-making.

F

Fine-level constructs details the constructs from the perspective of actors, tasks, actors' involvement, and communication.

G

Goals represents an objective or a state that a system strives to achieve.

Goal-based Approach is proposed in this thesis to work in amalgamation with the network-based approach. It specifies high-level view that policy making process works towards achieving.

Goals' Measures specify the conditions for the fulfilment of goals.

Government represents a body who is responsible for handling the policy process. Government manages involvement of other non-government actors in public policy process. It is responsible for making the final decision. Government here refers to local as well as national government.

H

High-level Construct covers the metrics from the higher-level abstraction i.e. at the generic policy cycle level

I

Implementation Phase is one of the phases of generic policy cycle which focuses on putting the public policy in operation

Information is one of the actors' involvement types where actor only consume the information being provided by the government. They do not actively participate in public policy making.

Inform Communication is one of the communication patterns where an actor informs others but require no response from them

Initiation is one of the actors' involvement types where actors starts the policy for governments' consideration.

K

Knowledge-intensive (also called knowledge working) refers to setups/environments where actors play a significant role for deciding the next action in a process.

L

Linear explains how the policy progresses from one phase to another. In this case, the phases are executed in sequential order of generic policy cycle

Local council is a body that implements the high-level policies of national government.

Loopback (linear) explains how the policy progresses from one phase to another. In this case, a loopback occurred among adjacent phases.

Loopback (Non-linear) explains how the policy progresses from one phase to another. In this case, loopback is not linear and adjacent phases are not involved in the loopback

M

Meta-model is a model that specifies the structure of models. It describes the syntax for a set of models.

Model is an abstraction of a system under investigation. It represents a system from a specific viewpoint. It is an instance of meta-model that respects the syntax of meta-model.

N

National Government sets out the high-level policy descriptors.

Network-based Approach is a process-agnostic approach that has been proposed in this thesis to capture public policy processes. The approach does not require beforehand process orchestration.

Non-government represent a set of third-party organisations who influence and participate in public policy making but are not authorised to make the final decision or manage the overall public policy making process.

O

Objective Evidence is the rational evidence that shapes public policy making and guides actors' decisions.

Organisation Structure is the high-level structure of the organisation which is used in the provenance framework for control and ratification purposes

P

Phases refers to the stages of generic policy cycle.

Policy Cycle is the general policy making framework.

Policy Cycle Provenance (PCP) Framework has been developed in this research to enable tracking of public policy processes.

Policy Cycle Provenance Proof-of-Concept implements the proposed framework to demonstrate working and evaluation of the framework.

Policy Formulation Phase is one of the phases of the generic policy cycle that focuses on identifying policy alternatives (i.e., options) for addressing the identified public policy issue.

Policy makers are the actors who formulate public policies. Policy makers and decision–makers are used interchangeably in this thesis.

Policy Scenarios refers to policy examples that are employed to test the provenance framework. Example of policy scenarios include scenarios such as invitation to citizens to participate in formal consultation. Another example entails no participation of citizens in the formal consultation process.

Process-agnostic Approach refers to an approach which does not depend on the process definition for devising policies.

PROV is W3C standard for representing inter-operable provenance [PROV, 2019].

Provenance Use-cases specifies how the policy making provenance data can be used.

Provenance represents contextual information about policy making processes which involve subjective and objective evidence, operational evidence, activities, and stakeholders involved in creating public policies. It is used to create on-the-fly processes, find delays in PPM processes, compare processes, show cohesive roles of actors, ideas, and institutions in policy cycle phases, and explain relationship among different policy concepts of policy cycle.

Public Policy Making is the process for creating public policies

S

Sequential Communication is one of the communication patterns where actor hands-over the responsibility to the other actor.

Smart Cities in this thesis refers to incorporation of Smart Governance principles in public policy processes.

Smart Governance is defined as one of the aspects of smart cities. The smart governance helps in conception of smart cities. This thesis considers involvement of non-government bodies and citizens in public policy processes. It also considers transparency as a smart governance objective in this thesis.

Stakeholders refers to all those actors who are directly or indirectly involved in public policy making. Stakeholders and actors are used interchangeably in this thesis.

Subjective Evidence points to all that data that is generated by actors such as actors' input, suggestions, and comments.

Synchronous Communication is one of the communication patterns where actors sends request to others. The sender cannot continue to work unless a response is received.

T

Token represent an entity of a network-based approach that captures the public policy making task details**.**

W

Workflow is a process-based technology that automate the business/scientific processes.

Acronyms

Asyn: Asynchronous communication

BI: Broadcast Inform

BS: Broadcast Synchronous

BA: Broadcast Asynchronous

CM*:* Case Management

CMM*:* Common Meta-model

EBPM: Evidence-based Policy Making

GUI: Graphical User Interface

HLC: High-level Construct

IS: Inform Synchronous

KiP: Knowledge-intensive Process

MAS: Multi-agent System

PMM*:* Phase Meta-model

PPM*:* Public Policy Making

PCP Framework*:* Policy Cycle Provenance Framework

Syn: Synchronous Communication

Chapter 1 : Introduction

1.1 Background and the Thesis Research Problem

Provenance, which is referred to as the history that is associated with the creation of any artefact (Ram and Liu, 2009), has its application in multiple domains. In scientific applications, for example, it is used to record the procedure of experiments. Likewise, businesses record provenance of their business processes for the purpose of monitoring performance and ensuring transparency.

In a similar way to that of other domains, Public Policy Making (PPM) also needs provenance management. Public policies significantly impact all aspects of our lives therefore developing a correct policy and exploring what went wrong is important (Cairney, 2012). Public Policy researchers (e.g., Wu et al, 2017; Bridgman and Davis, 2004) relate the managing of policy processes to successful PPM and effective policy results. The authors further elaborate this by emphasising the understanding of policy processes for positive policy consequences and for producing quality policies. Further study (Höchtl et al, 2016) shows that previously collected policy data can improve the implementation of future policies. For example, the analysis of budgetary patterns of past policies can provide useful insight of budgets for new policies.

Public policies are shaped by diverse evidence such as the review of scientific input, the analysis of previous policies, and discussion and consultation among stakeholders etc. Philip et al (2007) found that the lack of contextual information of evidence presents difficulties for ex-ante and ex-post public policy assessment. The authors stressed the significance of provenance for recording the context of PPM. Unavailability of provenance makes it difficult to understand the context of evidence thus raising challenges of policy data reusability, data quality and reliability, robustness of findings, and generalisation of policy results. Furthermore, Scherer et al (2015) pressed on provenance to understand the interrelations of aspects of policy cases. The absence of provenance creates a gap in how variable policy concepts can be woven into a policy case; this limits the analysis and usage of public policy data for other purposes such as for simulations. Provenance of PPM is also found to be useful to enhance efficiency and operational performance (Sajjad, 2014).

Provenance also plays a pivotal role in encouraging citizens' participation in PPM (Chun and Cho, 2012). Having no provenance of PPM processes can reduce the scale of trust and participation of public in PPM. Chun and Cho investigated a tool called "Cyber Policy Forum", which is used by the Seoul Metropolitan Government (SMG) to facilitate citizens' participation. Their analysis uncovered the decrease in the use of the tool; one of the reasons being explored by the researchers is that absence of provenance of the decision-making processes of government reduces public participation. The reason is that public does not have a complete picture of government's decision-making processes, therefore, it is not known to them how their inputs were considered which can negatively impact their trust on a government.

Seemingly, public policy is a document, but it is a result of rigorous analysis and discussion. The evidence that is used and/or generated for the creation of public policies therefore undergoes inspection from various PPM stakeholders. Thus, to understand the context of evidence (that is created and/or used), the interrelation of policy aspects, and for acknowledging citizens and other stakeholders' input, the provenance of a complete PPM process is required. Only with a full picture of the PPM process, can it be known how policies were shaped by evidence and where stakeholders' inputs were sought in the process. Furthermore, with the growing recognition of Evidence-based Policy Making (EBPM), provenance can provide an insight of the evidence that was used for policy creation (Kay, 2011). EBPM encourages rational PPM rather than ideologically driven policies. Managing provenance of PPM processes can enhance a government's operational performance in respect of acknowledging where the policy process is, tasks which have already been performed, and with whom the current task resides. The positive or negative outcomes of policies are often defined in relation to peoples' satisfaction and acknowledgment of government's policy (Chun and Cho, 2012). The key ingredients for the general acceptance of public policies among community are transparency, trust, and their involvement in Public Policy Making (PPM) processes (Wu et al, 2017; Sajjad, 2014; Chun and Cho, 2012). The trust among the stakeholders, transparency in government operations, and understanding of the PPM processes can be facilitated by managing PPM provenance data.

Given the nature of the problem, an in-depth investigation was carried out in this thesis to uncover any existing solutions for tracking PPM. The study revealed that currently no PPM provenance tracking support is available. The detailed research investigation by Sajjad (2014) further shows that public policy decision-makers needs to adopt various means for managing the PPM evidence and the human resources. Furthermore, policy makers need communication with various stakeholders both collocated and distributed and having no one platform in place presents an overhead of searching different means for collaboration. What is needed here is a **holistic platform** for managing PPM operation and provenance data and for facilitating stakeholders' participation.

The key finding from the investigation of PPM domain revealed its **knowledge-intensive** nature. The Knowledge-intensive Processes (KiPs) refers to human-centric processes where the next action is based on human knowledge. The literature survey shows that at present no solution has been designed for PPM given its knowledge-intensive setup. The KiPs of PPM demands a solution where decision-makers be given a liberty to guide the PPM processes.

The Knowledge-intensive nature of PPM induces adhoc-ness i.e., decision of action at any point of public policy process is guided and influenced by the stakeholders. Thus, processes are variable for public policies which limits the productivity of humans in respect of searching and locating the required evidence. The exploration further uncovered the dynamic nature of PPM in respect of complex task specification, variable role assignment to policy tasks, and run-time decision of the next required action. Due to the adhoc-ness and dynamicity attached to PPM, an **adaptable solution** is required.

Lastly, the literature shows governance being an integral part of PPM (Höchtl et al, 2016). Therefore, it is significant to consider governance practices in PPM provenance solution design. This research considers urban level policies and good governance contributes to smart cities. Therefore, what is required is consideration of **smart governance** principles in PPM for acknowledging smart cities.

In summary, the thesis investigation anticipated a need for an holistic platform that should (i) manage (i.e. capture, store, connect, and retrieve) the provenance data (ii) connect distributed and collocated stakeholders so different collaboration means/sources can be avoided which saves their time and resources (iii) consider inherent knowledge-intensive nature of PPM processes and (iv) be able to adapt with the changing needs of PPM processes. Therefore, this research presents a Policy Cycle Provenance framework (PCP Framework) for managing the lifecycle of PPM processes, see Figure 1.1 for policy cycle. The PCP framework is further elaborated in the subsequent section.

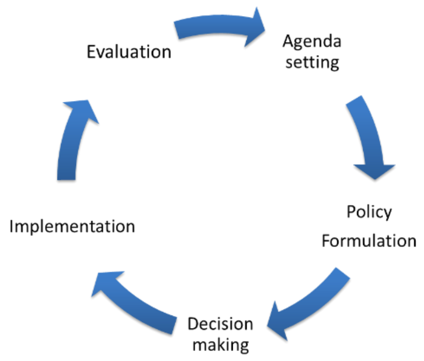

Figure 1.1: Generic Policy Cycle (Howlett et al, 2009)

1.2 The Proposed Solution

Given the significance of capturing public policy processes, the PCP framework (presented in Chapter 4, Section 4.3) is designed in this research to manage all the operations pertaining to PPM. It provides a holistic platform for all stakeholders (including government, non-government, and citizens) involved in policy creation process. The provenance related to PPM has been defined using EBPM rationale (Discussed further in Chapter 2, Section 2.2.2). The PPM provenance includes the process details, evidence that is used and produced during operations, stakeholders' input including citizens' involvement, and communication among other stakeholders. The PPM provenance is explained by 7Ws of provenance i.e. what, how, who, which, when, where, and why data (Ram and Liu, 2009). As provenance can be represented in several ways, this presents a problem for it to be understandable among stakeholders. Therefore, a standard provenance model called PROV (2019) model is employed for explaining the PPM provenance.

In this research, the generic policy cycle (Figure 1.1) has been employed for systematic representation and management of provenance information. The different stages of policy cycle compartmentalise the provenance information. Using the different stages, the provenance data can be easily understood and queried. Further details of the generic policy cycle and its suitability is covered in Chapter 2 (Section 2.2.1).

As stated in the previous section that PPM is inherently dynamic and ad hoc in nature. To address this, a model-driven approach is employed for the design of PCP framework. The model-driven approach facilitates dynamicity by exhibiting varying levels of abstraction and in building adaptable solutions (Da Silva, 2015). The PCP framework specifies the policy cycle provenance details in meta-models. The meta-models are used for generation of various provenance model at the time of capturing the provenance and during the provenance usage. Details of how meta-models are used in the PCP framework is presented in Chapter 4 (Section 4.3). The meta-models are covered in Chapter 5. Detailed overview of a model-driven approach is given in Chapter 2 (Section 2.6).

For provenance tracking, various technologies were examined. For example, Sajjad (2014) offered a philosophical stance on the suitability of a workflow technology for PPM. However, further investigation (Hauder, 2016; Van Der Aalst, 2013) uncovered challenges associated with solely using a process-based approach for knowledge-working environments. Literature demonstrates the requirement of an approach where humans are given liberty to define the next action depending on the established needs and requirements rather than defining a process that guides human actions. Detailed literature on the assessment of current process capturing technologies is covered in Chapter 2 (Section 2.5).

Considering the KiPs of PPM, a process-agnostic approach termed as a *network-based approach* inspired from IP packet switching (Warf, 2018), has been developed where users' actions and activities are captured as discreet activities. With a network-based approach, a goal-based approach is used to direct and control the process. The goal-based approach facilitates in efficient representation and querying of the provenance information. The proposed approach is presented in Chapter 4 (Section 4.4).

The conception of public policy making is changing with the emergence of new governance paradigms for the realisation of smarter cities (Barns. S, 2017; Rodríguez-Bolívar, 2015). PPM demands the inclusion of various stakeholders including citizens' participation in overall policy process. Therefore, the smart governance practices adopted for this research includes inclusion of diverse stakeholders, active citizens' participation, and transparency in the process. Further discussion is presented in Chapter 2 (Section 2.2.3)

To conclude, the PCP framework is based on a model-driven and a process-agnostic approach following smart governance principles. The collected provenance provides a base to understand and gain insight on the already executed processes. For capturing provenance, the PROV model (PROV, 2019) is used, which facilitates capturing details as per the provenance standard and assists in provenance interoperability. The collected provenance is used to: ***construct on-the-fly process*** i.e. the process that created policies, identify any **delays** caused during task execution, ***compare processes*** of past policies in order to suggest a process for a new policy, study ***different types of evidence (including context) being used during policy creation***, demonstrate intertwined ***roles of actors, ideas, and institutions in different phases of policy cycle,*** and to acknowledge relationship among ***different policy concepts of policy cycle (Figure 1.1).*** Further details of PPM provenance use cases are covered in Chapter 2 (Section 2.3).

Evaluation of the PCP Framework includes the appraisal of a model-driven, network-based approach, and the smart governance principles. The evaluation entails an assessment of the framework against case studies (detailed in Appendix B) which are centred on the top-down (policymakers plays a central role in policy formulation and implementation), bottom-up (citizens and other non-government officials plays a central role), and hybrid (government works together with other actors) approaches to PPM. PCP framework is assessed by evaluating the suitability of a model-driven and a network-based approach for PPM in smart cities. Furthermore, the effectiveness of the provenance framework has been explored by acquiring users' input on the PCP framework. Provenance use cases were assessed to evaluate the usefulness of the captured provenance.

The evaluation results show that a PCP framework is appropriate to document PPM processes because the framework successfully adapts when variable PPM processes are tested. Furthermore, the results exhibit that human knowledge guides the process which depicts the ability of the framework to support knowledge-intensive nature of PPM. The results further show that the captured provenance is effectively used to create on-the-fly process, compare processes, calculates delays, and provide insight of policy cycle phases. The results depict that PCP framework addresses the gaps that have been identified from the detailed literature investigation which makes it suitable for tracking PPM processes. The evaluation and results are covered in Chapter 6.

1.3 Research Aim, Objectives, and Outcomes

This section elaborates the aim, the objectives, and the outcomes of this research.

1.3.1 Research Aim

The aim of this research is to

Develop a provenance framework that assists to track the knowledge-intensive public policy making processes and adapt with the changing needs.

1.3.2 Research Objectives

The research aim has been accomplished by fulfilling the following objectives.

(i) To deepen the understanding of PPM process and its context in smart cities, evidence-based policy making, potential provenance use-cases for public policy devising, technologies' support for knowledge-intensive setup of public policy making, and suitability of a model-driven approach.

(ii) To push the boundary of knowledge in terms of designing a PPM provenance framework for smart cities.

(iii) To evaluate the research though an assessment of (a) the suitability of a model-driven approach (b) a process-agnostic approach (c) supporting diverse stakeholders (smart governance objectives) and (d) the impact on the public policy making community in terms of effectiveness.

1.3.3 Research Outcomes

Given the research objectives, the outcomes of this research are as follows:

(i) **Objective-1** of this research aims to produce a holistic view of previous research in the domain of PPM including an understanding of the different models of PPM, of which the policy cycle is a part. The study covers the changing dynamics of governance and its impact on public policy formation, the current research on provenance, and investigation on the suitability of technologies for knowledge-intensive and dynamic setups. Finally, it helps in identification of provenance use-cases and uncovers the impact on the policy making and research communities.

(ii) **Objective-2** of this research leads to the identification of research gaps in respect of tracking PPM processes in smart cities. The gaps serve a foundation to design a provenance framework, by using a generic policy cycle. The provenance framework employs a model-driven approach to capture diverse and dynamic PPM processes and a process-agnostic approach to address the challenges of a knowledge-intensive setup of PPM.

(iii) **Objective-3** of this research provides an evaluation approach for testing the provenance framework. The evaluation assesses the suitability of a model-driven and process-agnostic approach for tracking policy making processes. An end-user evaluation further provide insight into the effectiveness and usefulness of the provenance framework. The evaluation prove/disprove the research hypothesis.

From the research aim and objectives, the research hypothesis and questions were formulated which are covered in the subsequent section.

1.4 Research Hypothesis and Questions

'A model-driven and a process-agnostic approach can be used to design a provenance framework to capture generic public policy making cycle in smart cities to support evidence-based policy making and to facilitate adaptable and knowledge-intensive processes'

To support the stated research hypothesis, the following research questions are designed:

RQ1) Why is the provenance of public policy making processes in smart cities necessary?

To answer this question, an extensive literature review has been carried out in the domain of PPM, EBPM, and smart cities' context of PPM. This research question is answered in chapter 2.

RQ2) Why is a model-driven approach appropriate for managing public policy making provenance data?

To answer this question, Chapter 2 (Sections 2.4 and 2.6) investigates the suitability of a model-driven approach for PPM provenance capturing.

RQ3) To what extent can a model-driven based provenance framework support provenance tracking in a generic policy lifecycle?

As part of the model-driven approach, PPM meta-models are designed. Chapter 4 covers the design approach to meta-modelling; however, the details of meta-models are covered in chapter 5. Lastly, evaluation in chapter 6 test the proposed approach to assess if it can support top-down, hybrid, and a bottom-up approach to PPM.

RQ4) Can a process-agnostic approach be suitable for a provenance framework to capture a knowledge-intensive public policy making processes in smart cities?

To find a solution to this question, various process support technologies were investigated (Chapter 2, Section 2.5). The analysis of existing technology in the light of PPM domain requirements results in developing a process-agnostic approach called a network-based approach. This question is thus answered by designing, implementing, and evaluating a network-based with a goal-based approach for the purpose of provenance tracking. Chapters 4 and 6 answers this question.

1.5 Research Constraints and Assumptions

The research constraints for this research are identified in relation to time and other resource constraints. The research constraints are as follows:

Research Constraint 1: Policy making is a diverse domain, spanning across a number of interdisciplinary areas. Designing detailed meta-models or models for each policy area or domain is not practical in the limited time span of this research. Therefore, modelling of a policy domain is out of the scope of this research and meta-models intend to cover only the policy making processes. The meta-models sketch the process and data details related to the PPM and not the vocabulary of any policy domain.

Research Constraint 2: The proof-of-concept is implemented for the purpose of testing the proposed approach. Developing a full-fledge system entails additional resources such as training, additional human resources, deployment, and maintenance of the software. In addition, it is also time taking. Therefore, the testing of the proposed solution is done by using the developed proof-of-concept.

Research Constraint 3: For evaluation, improvised data is used for research evaluation, though it is very appealing to test the generality of the provenance framework in some practical environment. However, this is a very time-consuming task as PPM is a long process and therefore it is not practical to test the system in the production environment. The system is tested by considering the case studies (Appendix B) collected from the literature and evaluation metric defined in chapter 6.

Research assumptions are defined for this research to simplify the conception, implementation, and evaluation. The assumptions are as follows:

Research Assumption 1: PPM involves a diverse set of stakeholders (distributed and collocated). Their participation depends on the policy under consideration. To simplify the implementation and evaluation, it is assumed that actors fall under three categories (i) Government: it includes local as well as national government (ii) Non-government: it include third parties, companies, NGOs, and other bodies who give input in PPM and (iii) Citizens which include general public.

Research Assumption 2: Due to diverse set of stakeholders' involvements in PPM, data governance is challenging. In this research, it is assumed that the operations pertaining to PPM is responsibility of government i.e. managing of PPM processes and capturing of provenance data.

Research Assumption 3: For the consideration of policy on governments' agenda, often different means such as campaigns, social media, or other media are used to put pressure on government. This research does not consider this and assumes that PCP framework comes into play once the policy has been taken into account by government.

Research Assumption 4: It is assumed that all the tasks being managed by PCP framework is being input by the actors.

Research Assumption 5: As stated before, a goal-based approach is used in amalgamation with the network-based approach. Due to uncertainty attached with the PPM processes, the goals for each policy may change. Therefore, manual input of goal is assumed for the policies.

Research Assumption 6: During PPM, consultation is carried out among actors to discuss a common point. It is assumed that consultation process is automated i.e. actors will hold the consultation using the PCP framework. This not only saves time but also help to connect distributed stakeholders.

Research Assumption 7: To manage the operations of implementers, it is assumed that Non-government use PCP framework' Implementation workspace to capture their implementation operations. Further details of implementation workspace are covered in Chapter 5 (Section 5.3.4).

Research Assumption 8: For managing the overall policy process, PCP framework maintain the default organisation structure for different policy types. Depending on the policy need and knowledge at hand, organisation structure can change during policy creation.

Research Assumption 9: PPM is inherently a long-driven process. Defining delays in PPM processes is however somewhat subjective and often unavoidable. In this thesis, delay in a PPM process incurs when completion time of a task exceeds the expected task completion time.

1.6 Thesis Structure

The thesis structure is as follows: **Chapter 2** discuss the literature on provenance and its role in PPM, smart cities and PPM, discussion on existing technologies for capturing provenance and a model-driven based solution. The literature forms the base for identifying the research gaps. Chapter 2 also looks for potential future directions. **Chapter 3** details research strategy, reason of using qualitative approach, and the research methodology. **Chapter 4** covers the PCP Framework in detail that is based on the identified research gaps in chapter 2. This chapter also covers the meta-models design approach and the aims and objectives of each policy cycle phase. **Chapter 5** presents the modelling and development of the PCP framework. Chapter 5 also provides an overview of implementation details of a proof-of-concept. **Chapter 6** covers the evaluation framework in detail. This chapter also discuss the evaluation process. Test cases are designed and tested against the proof-of-concept. It also critically discusses the results and findings in the context of research questions and research hypothesis. Last, **Chapter 7** provide a conclusion by discussing the contribution and impact of this research. The research questions posed in chapter 1 are critically assessed to evaluate the research hypothesis. Finally, the thesis closes with the research limitations and possible future research directions.

Chapter 2 : Literature Review

This chapter focuses on identifying research gaps by critically reviewing the literature on provenance, public policy making (PPM), significance of provenance in PPM, smart cities with a focus on smart governance and its relationship with public policies, suitability, and requirement of a model-driven approach for PPM provenance management, existing provenance capturing mechanisms/technologies, and any current technologies being suitable for managing PPM provenance data. Figure 2.1 shows the structure of this chapter.

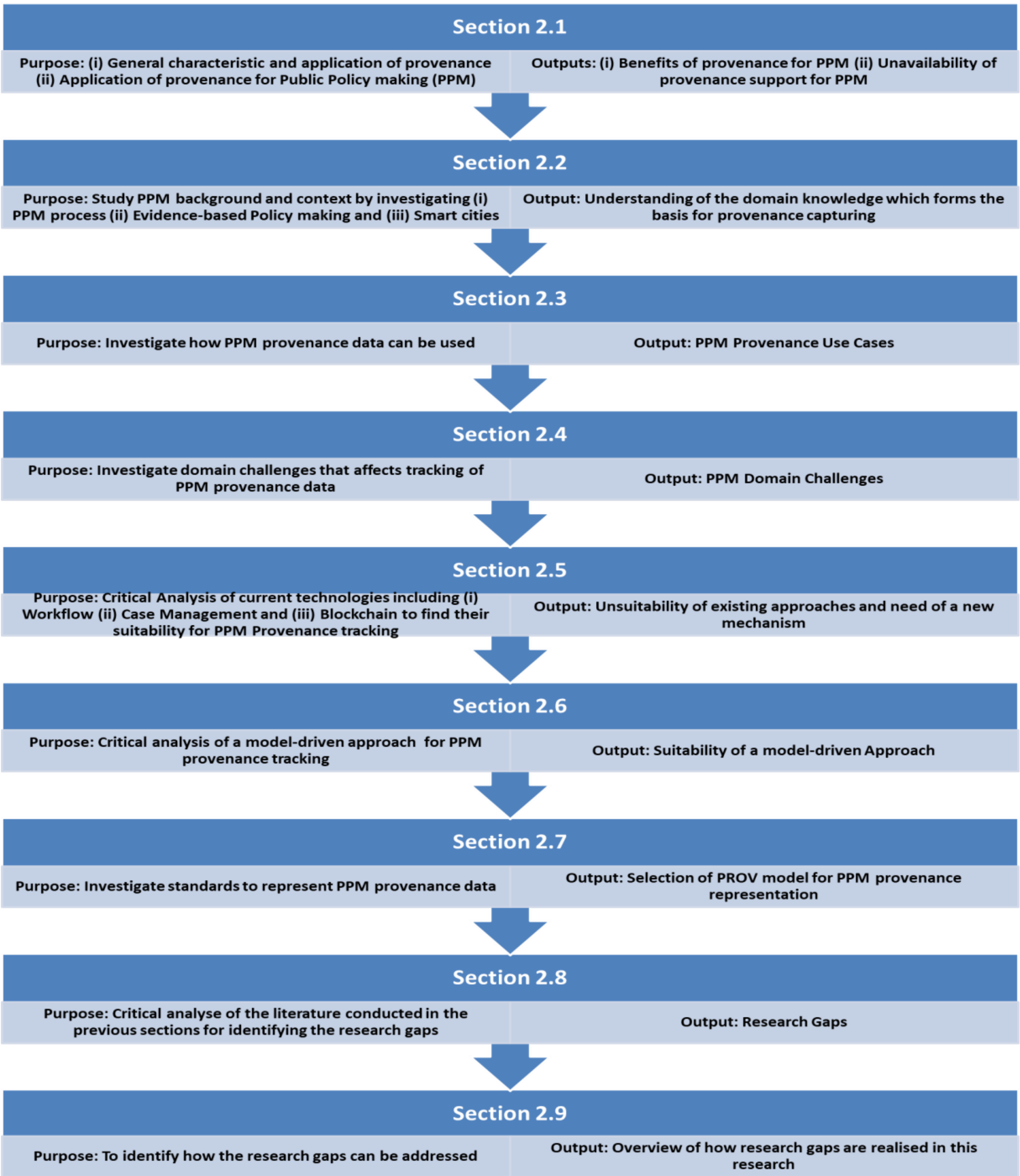

Figure 2.1: Structure of the Chapter

2.1 Characteristics and Applications of Provenance

Provenance refers to the history that is associated with the creation and modification of any artefact (Ram and Liu, 2009). It has for many years been employed to maintain information regarding the ownership/history of objects (Bose and Frew, 2005; Cheney J et.al, 2007). The term provenance was first coined in the work of arts where provenance determined originality, ownership, and value associated with an art object (Lakshmanan et al, 2011). Provenance captures what, how, who, which, when, where, and why data (Ram and Liu, 2009), also called the 7Ws of provenance. Inherent characteristics of provenance include ensuring data quality, verification and validation of information, reliability and accuracy, integrity, authenticity, validation of attribution of data, ownership, and transparency and trust in the system (Huynh et al, 2013; Liu et al., 2013; Zeng et al, 2011; Margo and Smogor, 2010).

The characteristics of provenance have attracted its application in various other scientific and business applications (Lakshmanan et al, 2011). For example, in experiments at the Large Hadron Collider (LHC) which was carried at CERN (Branson et al, 2014) provenance was captured to assist distributed scientists/researchers in their research. It has been employed by biomedical researchers to explore Alzheimer's disease (McClatchey et.al, 2014). Provenance also plays a significant role when conducting a scientific experiment where experimental data and processes are required for re-computation, verification, and the repetition of experiments (Bose and Frew, 2005). Similarly, business setups exploit provenance to monitor performance goals, to track the business process, or to ensure transparency in their operations (Lakshmanan et al, 2011). Furthermore, provenance provides a rich source of information which can be used to perform analytics (Huynh et al, 2013; Liu et al., 2013; Zeng et al, 2011; Margo and Smogor, 2010). In recent years, provenance has been widely employed for process mining of business processes (Van der Aalst, 2011).

The various applications and benefits of provenance has also attracted attention in government setups. For instance, Ding et al (2010) developed Data-Gov wiki to facilitate accessibility of government data for public consumption; in their work the provenance is inferred during the process of conversion, enhancement, and usage of government data. Provenance has also been used to track consumption, production, and mash-up of open government data (Ding et al, 2011). Furthermore, an important role of government is establishing public policies and it has been found that provenance can play a significant role for PPM (Chun and Cho, 2012; Chorley et al, 2008).

Chun and Cho (2012) proposed transparent policy decision making model to capture PPM processes. The model has been proposed for integration with existing Cyber Policy Forum; it proposes an idea of using provenance to make PPM processes visible for public to help them make informed decisions. The authors also argue that provenance can help citizens to gain insight of where their input has been incorporated in PPM process which can help to build their trust on government and policies.

In Scherer et al (2015), a model-driven and agent-based public policy modelling was done to simulate public policy scenarios. This work focuses on assisting public policy community to gain in-depth understanding of public policies. The framework takes existing or new public policy and background documents as an input and generates simulation models. Provenance is maintained during the process of transformation; the provenance maintains traces from models to documentary evidence. It provides transparency in the process of creating policy models thus, adding trust in the process of creating the models.

In recent years, Evidence-based Policy making (EBPM) has gained wider recognition among policy researchers (Dunn, 2018; Parkhurst, 2017; Sutcliffe, 2005). In this regard, Chorley et al (2008) developed provenance and argumentation solution for social scientists involved in evidence-based policy assessment. The detailed investigation being performed by scientists is captured to inform public policies; the provenance helps policy makers to have in-depth understanding of scientists' investigation and help them to base their decisions on the findings.

Analysis

The current literature demonstrates the potential of provenance for public policies. The literature (Chun and Cho, 2012) signifies the need of provenance for capturing PPM processes, but its implementation and practical application has not been realised. Further, what provenance data the PPM process entails has not been investigated by the authors. Similarly, provenance solution (Chorley et al, 2008) was designed for social scientists but how their findings are employed in PPM process is not the focus of their research. Scherer et al (2015), captures provenance during the process of transformation of public policy scenarios to simulation. However, how policy scenarios were formulated and why certain decision were taken is not captured by the provenance.

The current usage of provenance for public policies does not consider capturing the process that results in policies; further, it has not been investigated that what is the potential PPM provenance data. The need of it has been realised by the current researchers (Wu et al, 2017; Sajjad, 2014) who argues that previous policies serve as an evidence base for guiding other policies and managing information of previous policies can result in better policy outcome. Furthermore, it has been acknowledged that PPM processes influence the quality of policies (Bridgman and Davis, 2004). Therefore, tracking provides an evidence-base that can be used for analysing and improving policy creation process. The provenance tracking of PPM also provides insight of actors who were involved in the process and evidence that was employed for policies' creation. Furthermore, PPM is commonly a lengthy process (Sajjad, 2014) and decision makers need to search and locate previous executed tasks and decisions during the process of policy formulation. Managing PPM provenance data addresses the challenge of locating previously executed tasks, saving time, guiding further actions, and providing transparency (Höchtl et al, 2016; Hill, 2013; Waller et al., 2009).

Given the potential of capturing PPM processes, this research focuses on managing PPM provenance information (i.e., capturing, storing, and retrieval/usage of provenance information) and defining the details of PPM provenance data. For the purpose of this thesis, provenance represents contextual information about public policy making processes which involves usage and generation of subjective and objective evidence, operational evidence, tasks carried out during PPM, and stakeholders involved in creating public policies.

For designing a solution, it is significant to investigate in detail what is PPM process and what are the challenges of tracking and managing public policies. The subsequent sections, thus, sketches the PPM process details and associated challenges.

2.2 PPM Background and Context

This section provides background knowledge on PPM from the three perspectives as shown in Figure2.2.

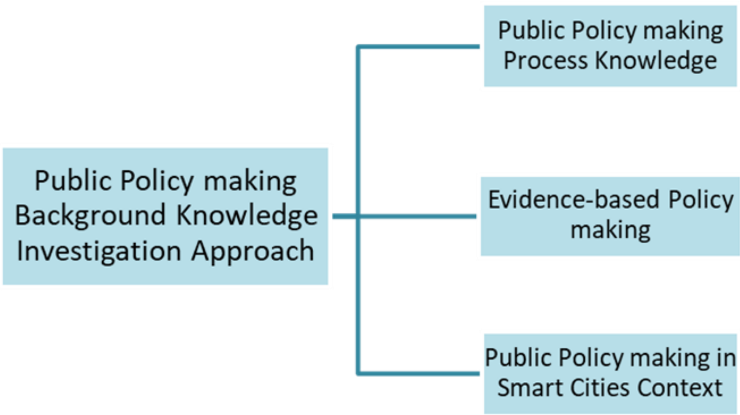

Figure 2.2: Background knowledge on PPM

Seemingly, public policy is a document stating plan, rules, and regulations. In reality, however, this piece of document is produced after a rigorous process of analysis and discussion. Knill and Tosun (2012) defines public policy as a course of action (or non-action), result of mutual concessions of actors, taken by government or legislature regarding a specific issue. How concession among actors is achieved and how government undergoes the process of policy creation requires in-depth investigation of PPM. Thus, Section 2.2.1 covers PPM process. An extensive literature exists in the domain of public policies such as policy content (i.e., how a specific policy emerged), different policy analysis approaches and techniques, or policy theories such as economic or social theories (Hill, 2014; House et al, 2008; Weimer and Vining, 2004); as the focus of this research is investigating the PPM process details therefore other areas of public policy research is out of scope of this research. To explore the evidence that is used during the process of policy formulation, this section (2.2.2) provides an insight of EBPM.

It has been found that governance principles form an integral part of public policies (Höchtl et al, 2016) therefore it is significant to investigate emerging governance principles. As public policies are generally formulated at the different levels of government including international, national, and local (at urban and rural level) (Anderson, 2014), this research focuses on urban level policies (further discussed in Section 2.2.3) and therefore the governance principles are investigated at city level. The conception of smart cities in the recent years define the governance practices at the city level; as public policies play an important role in smart cities' development (Meijer and Bolívar, 2016) therefore, this section also discusses PPM in the context of smart cities.

2.2.1 PPM Process

Policy making informs how policy is made (Hill, 2014). PPM is not only about context, tasks, process steps but it is also an instituitional arrangement, role of elites, and politics. The literature reports diverse theoratical approaches to PPM (Anderson, 2014; Hill, 2014); of which Open systems framework (Hofferbert, 1974), institutional rational approach (Kiser and Osterm, 1982), policy stream approach (Kingdon, 1984), advocacy coalition framework (Sabatier and Jenkins, 1988), and a stage model (Howlett et al, 2009; Lasswell, 1956) are the most prominant ones (Sajjad, 2014).

The open systems framework by Hofferbert (1974) stresses that historic geographical conditions, socio-economic composition, mass political behaviour, governmental institutions and elite behaviour directly or indirectly affects the policy output. This approach is suitable for cross-sectional comparison but offers several limitations including overstated role of elites, conflict between government institutes, and lack of interaction within intergovernmental policy communities (Sabatier, 1991).

The Institutional Rational Choice theory (Kiser and Osterm, 1982) focuses on effects of institution arrangement and individual actions on government policy decisions. The theory discusses the impact of institution on individual actions whose behaviour is guided by the instituition rules. Given the instituitional arrangement, the high-level body set rules for the lower level. Sabatier (1991) reports several limitations icluding neglicance of cruicial factors affecting societal effects and policy decisions and argues that focus on individual institute makes multitude of institutes unmanageable.

An interesting approach to agenda setting and policy formulation entitled 'The Policy Stream Approach' was proposed by John Kingdon (1984). The approach conceptualised PPM as three streams (i) a problem stream: focuses on information about real world problems and study the effects of past actions of government (ii) a policy stream: community, such as researchers, specialists, who analyse problem and formulate alternatives and (iii) a political stream: focuses on the outcome of elections, legislative leadership contests etc. The author argues that PPM process is not sequential and the three streams are unrelated but converge by 'a window of opportunity'. The approach goes beyond rigid institutionalism, incorporate view on policy communities, and give substantive information on problems and alternatives. The approach however, is not without limitations (Sabatier, 1991) such as large distance between ‘policy’ and ‘political’ streams and no priorities among the streams.

Sabatier and Jenkins (1988) developed a conceptual framework called ‘Advocacy Coalition Framework’ which views policy change as a result of three factors (i) interaction of advocacy coalition within policy community (including both public and private organistation) (ii) external changes and (iii) the effects of stable systems such as constitutional rules on the resources. The framework faced several criticisms (Sabatier, 1991) such as central focus on belief system and neglicance of actors’ interests, long time period for policy change, and variable institutional arrangements that makes the PMP process more difficult to execute.

PPM is inherently complex and presents many analytical difficulties. To simplify PPM, Lasswell (1956) proposed PPM as a process having a set of interrelated stages (Howlett et al, 2009). This sequence of stages is called policy cycle (Howlett et al, 2009; Jann and Wegrich, 2007). The model of Lasswell formed the basis for other researchers (Simmons et al, 1974; Brewer, 1974; Lyden et al, 1968); of these, Brewer (1974) proposed the simpler version of policy cycle and improves the Lasswell work by expanding problem beyond the confines of government. The work further clarified the policy stages and established the idea of policy process as an ongoing cycle. Based on Brewer work, several researchers have proposed different stages of policy cycle (Howlett et al, 2009; Peters and Pierre, 2006). Howlett et al (2009) discusses generic policy cycle that resulted from past researches; the generic policy cycle encompasses following stages (i) agenda setting (ii) policy formulation (iii) decision-making (iv) policy implementation and (v) policy evaluation (Figure 1.1, Chapter 1).

Stage model offers various benefits over the other PPM approaches (Sajjad, 2014). For example, it clearly specifies PPM as distinct tasks (Howlett et al, 2009) and clarifies the role of actors in PPM (Parag, 2008; Sobeck, 2003). As policy process is disintegrated into stages, these stages can be investigated exclusively or in relation to other stages. Another significant advantage of this model is that it is applicable at all socio-legal and spatial levels of government ranging from local to international level (Howlett et al 2009; Billings and Herman, 1998). Policy cycle also supports intertwined role of actors, ideas, and institutions (Howlett et al, 2009).

The stage model, however, is subjected to criticism by several researchers. Jenkins-Smith and Sabatier (1993) and Howard (2005) discuss that PPM process is ad hoc but policy cycle can be misinterpreted to solve policy problem in a systematic and linear way. Howlett et al (2009) argues that applicability of policy cycle model at different units/levels of government is unclear. Sabatier (1991) criticises that it is not evident that what drives a policy cycle from one stage to another. Lastly, Everett (2003) points out that policy cycle does not explicitly specify the content of policy.

Analysis

Public policy formation can be achieved using any of the previously discussed approaches. Open systems framework, institutional rational approach, policy stream approach, and advocacy coalition framework, however, do not focus on mutual understanding of issues by public and policy makers (Sajjad, 2014). These approaches generally cover the roles of elites, institution arrangement etc for PPM. However, this research is focused on PPM process thus, the aim of this thesis is to adopt a model that mainly focuses on process description. Furthermore, with the changing governance dynamics, citizens' involvement in PPM is becoming prominent (Wu et al, 2017). The influences of policy makers in the decisions are prevalent in these approaches (Sajjad, 2014). Therefore, a PPM approach is required that could incorporate citizens' input and participation. Public policy cycle on contrary provides room for citizens' participation in the PPM process; it also focuses on defining detailed PPM process (Sajjad, 2014).

For provenance tracking, policy cycle is appropriate because it provides a systematic yet holistic view of PPM processes (Howlett et al, 2009). Given the different stages of policy cycle, provenance data can inform which stage the data is attached to; this provides a systematic approach for management of the provenance information. Furthermore, policy cycle has also been realised appropriate to support EBPM (Sutcliffe, 2005). Using the cycle, the evidence can be easily positioned which assists in management and querying of provenance information. Policy cycle is also used for other projects such as FUPOL but not for managing the provenance information.

Though, Everett (2003) criticises that content in policy cycle is not explicitly specified but this provides an opportunity to introduce flexibility to address the ad hoc environment of PPM. Policy cycle addresses the criticism it faces regarding linear process by facilitating various loops (Figure 2.3) in the process (Prpić, et al, 2015). The policy cycle also provides a liberty to exploit it at any level of government. This research focuses on urban level policies and the policy cycle is an appropriate choice.

Given the potentials of generic policy cycle for provenance management, this research considers the policy cycle for provenance organisation. The subsequent section provides an overview of the type of evidence used in policy creation and discusses the evidence from the perspective of policy cycle stages.

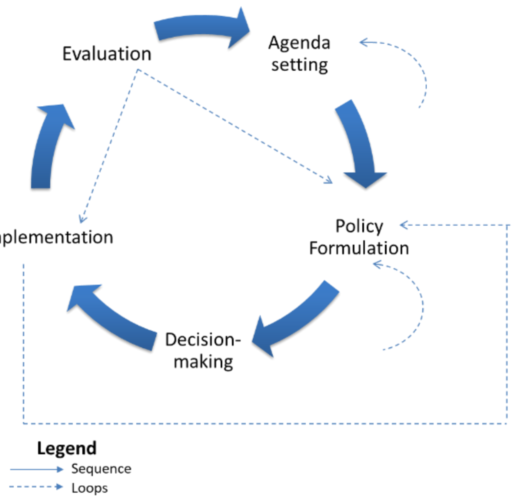

Figure 2.3: Loops in Generic Policy Cycle

2.2.2 EBPM

EBPM gained political recognition in Blair's Administration (UK) (Kay, 2011). The EBPM replaces ideologically driven PPM with rational decision-making (Sutcliffe, S., 2005). It is a systematic approach to consider valid and representative evidence to guide PPM. Use of evidence results in informed decision-making which produces better outcome (Sutcliffe, 2005; Parkhurst, 2017).

What evidence PPM constitutes is certainly a point of discussion; this is because PPM requires a wide breadth of evidence in the process. Furthermore, it is challenging to conceptualise all the evidence that is employed for PPM. Attempts have been made to define the evidence that is used in PPM. Generally, it entails evaluation of previous policies, experts' knowledge, existing research, consultation output, scientific input, and/or existing research. According to Chorley et al (2008), evidence include primary research, stakeholders' opinion analysis, simulation modelling, elicitation of public perception and beliefs, cost benefit analysis, and judgement of quality of methods used for gathering and analysis of information. Shaxson (2005) further argues wide and dynamic evidence in respect of different policies. Furthermore, stages of generic policy cycle incorporate different evidence. Table 2.1 lists evidence associated with policy cycle phases (Sutcliffe, 2005).

Table 2.1: Evidence Associated with Public Policy Cycle Phases

Policy Cycle Phases	Possible Associated Evidence
Agenda setting	This phase requires evidence in respect of identification of new issues, collecting evidence to demonstrate its significance, and discussion with relevant stakeholders to decide if it is supposed to be taken on agenda
Policy Formulation	Understanding of different options i.e., alternatives of the specific situation. Employing different investigation methods for an in-depth understanding of an issue.
Decision making	Formal consultation and informed decision of a policy that will be implemented
Implementation	It includes operation evidence i.e., practical experience when a policy is put into action
Evaluation	Evaluation procedure to evaluate the policy. Here, majorly the focus is on objective evidence. It also includes communication that whether the policy will continue or not.

Analysis

The different forms of evidence associated with PPM phases indicate the sort of evidence that is used for creation of public policies. Evidence related with each phase suggests that both objective and subjective evidence is used for policy creation. Provenance management therefore require capturing of objective as well as subjective evidence. Objective evidence here involves any facts or findings that are used to shape decisions and subjective data entails actors' input in the policy process. This clearly demonstrates that EBPM considers facts and objective analysis but do not deny the political nature of PPM and acknowledges actors' interpretation and input.

To explain the provenance, identification of concepts for each phase is needed. For example, Table 2.1 shows that in *agenda setting* **issues** are identified and discussed for further consideration whereas in *policy formulation* stage, **policy option i.e., solutions** are identified. This signifies that PPM provenance should be as such that can explain the details of what incurred at each phase. In addition, provenance also needs to explain the relation among phases.

Evidence alone is not the only factor that influence PPM, it is also influenced by governance principles (Höchtl, 2016). The growing needs and demand of smart cities also require shifts in governance practices as discussed in the subsequent section.

2.2.3 PPM in the Context of Smart Governance

Public policies are framed at different levels of government. For example, in England policies are created at National and Local level. The national policy is a framework which local authorities, along with the local community, employs to prepares a local document which then sets out a vision and framework for future development (Plain English guide to the Planning System, 2015; Cirianni et al, 2013). As discussed previously, generic policy cycle can be exploited at different levels which implies that there exists no substantial difference in its use at variant levels of government. For this research, urban level policies are considered. With the growing challenges of urbanisation, the need for smart cities has been identified (Harrison and Donnelly, 2011) and public policies plays a significant role in smart cities conception (Meijer and Bolívar, 2016). The smart cities have attracted considerable attention in the context of urban development policies. Therefore, the relation of smart cities and public policies is explored.

To-date, there is no single concrete definition of smart cities and it has consequently been conceptualised differently by various researchers (e.g., Scholl, 2014; Harrison and Donnelly, 2011; Nam and Pardo, 2011). By having no established definition, the interpretation is left to the body responsible for managing and controlling operations of cities. However, this ambiguity also gives the liberty to define the smart city project that targets cities' specific objectives, concerns, and problems (Angelidou, 2014). Governance principles being an integral part of public policies (Höchtl et al, 2016) signals exploring the 'smart governance' aspect of smart cities in this research. Furthermore, consideration of smart governance practices by government results in 'smart government' (Rodríguez-Bolívar, 2015).

Similar to smart cities, for smart governance and smart government, there is no one standard definition. Various researchers (Rodríguez-Bolívar, 2015; Mellouli et al, 2014; Scholl, 2014; Nam and Pardo, 2014; Sindhu et al, 2010) have proposed understandings of the concepts. From the research investigation, the relationship between smart governance and smart government is abstracted as shown in Figure 2.4. The focus of this research is on the circles presented in 'orange'. Figure 2.5 shows the relationship between smart government, smart governance, and public policies.

Figure 2.4: Relation between Smart Government and Smart Governance

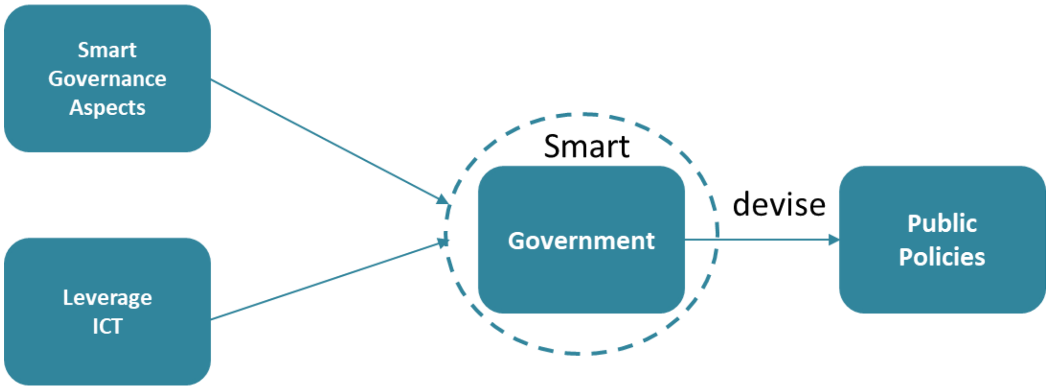

Figure 2.5: Focus of this Research in the Context of Smart Cities

Figure 2.4 shows that smart governance aspects considered for this research are involvement of different stakeholders, citizens' participation, and transparency. Involvement of stakeholders pertain to non-government organisations that directly or indirectly influence policies. With the changing governance dynamics, citizens' involvement in public policy decisions has become more prevalent thus, it is significant to employ this aspect of smart governance. Lastly, provenance information by default introduce transparency in public policy creation. Figure 2.5 further shows that smart governance facilitates smart government who in return can devise better policies. Given this, following definitions (Scholl, 2014) of smart cities, smart government, and smart governance are considered for this research.

i) **Smart Governance** forms a foundation for smart government and facilitates inclusion and collaboration with various stakeholders. Furthermore, it ensures transparency in the process.
ii) **Smart Government** is defined as the use of ICT for management and the implementation of policies and employing of smart governance principle for realising smart government.
iii) **Smart Cities** for this research refer to the use of smart governance principles for the achieving smart cities.

Given diverse stakeholders and citizens' involvement, it is significant to explore if there exists dedicated or specialised policy making process for smart cities and is discussed in the following sub-section.

2.2.3.1 PPM Cycle for Smart Cities

To-date there exists no specialised PPM process for smart cities (Parycek and Pereira, 2017). What differentiates smart cities policy making from traditional PPM is the use of smart governance principles and the use of ICT for supporting policy creation (British Standard Institute, 2019; Kourtit et al, 2017; Parycek and Pereira, 2017). Section 2.2.1 analysed that generic policy cycle is suitable for policy devising at different levels of government. As polices for smart cities falls under the umbrella of urban polices which are devised at the local level of government therefore a generic policy cycle can be adopted to devise policies for smart cities.

For producing better policies, actors play a vital role and is thus focused on by smart governance. Investigation regarding actors' participation in PPM process was therefore carried out and is covered in the next sub-section.

2.2.3.2 Actors' Role in Policy making

Traditionally, PPM is the responsibility of government officials (Chun and Cho, 2012). However, in the recent years, the dynamics of governance is shifting towards more open policy processes where citizens and street-level bureaucrats have a profound role in the PPM process (Howlett et al, 2009). This has resulted in emergence of various PPM approaches such as top-down, which is a traditional approach to devise policies, bottom-up, where street-level bureaucrats play a signification role in policy formulation and implementation, and a hybrid approach which amalgamates both top-down and bottom-up approaches.

What precisely defines top-down, bottom-up, and a hybrid approach however largely depends on the issue and the institutional context (Homsy et al, 2019; Hottenstein, 2017; Wu et al, 2017; Howlett et al, 2009). Generally, top-down centres on government decisions and the realisation of those decisions in the implementation phase (Howlett et al, 2009). A bottom-up approach involves diverse actors (including those who implement or are impacted by the policy) in the policy formulation stage; this has become more acceptable with approaches such as co-production and co-deciding (Edelenbos and Klijn, 2005). Wu et al (2017) place those scenarios in a bottom-up approach in which the implementers are given authority to change the programs with changing needs; the authors further discuss the decentralised approach being followed by implementers in the policy implementation. Lastly, the hybrid approach benefits from both the top-down and bottom-up approaches. In the hybrid approach, policy formulation and implementation are carried out by considering the goals of decision-makers and of street-level bureaucrats (Hottenstein, 2017).

Given the different perspectives on the approaches this research takes the following perspectives for the three approaches: (i) Top-down: an approach where government plays a central role and defines a policy and implementation approach. Citizens and other stakeholders may be involved during policy formulation but may not necessarily contribute in decision-making and do not promote participation (Bartoletti and Faccioli, 2016) (ii) Bottom-up approach: where citizens, street-level bureaucrats play a central role in policy formulation, reformulation, and implementation (iii) Hybrid Approach: where both government and street-level bureaucracies are actively involved in decision-making; for example, implementers may decide of some modification during policy implementation.

Analysis

The study of smart cities uncovered the role of varying stakeholders and citizens in policy creation. This suggests that provenance should capture their participation and the activities/task they carry out during PPM. It has also been found that provenance framework should be holistic and be able to support top-down, bottom-up, and hybrid approach to PPM. This research is not in a position to comment which approach (top-down, bottom-up, or hybrid) should be adopted by government. Although, Hill (2013) has extensively debated on these different paradigms. For this research study, however, this is not important because what approach may be suitable for one policy may not be appropriate for the other. Therefore, this study does not comment on which approach is more efficient.

From the literature covered in this section and the previous section (i.e., section 2.1), several provenance use-cases have been identified. Discussion of the use cases are presented in the following section.

2.3 PPM Provenance Use-Cases

PPM provenance use cases are detailed below:

i) Bridgman and Davis (2004) report good policies being a result of a rigorous process. The provenance can uncover the process that was followed for policy creation. The provenance can help to ***construct on-the-fly process***.

ii) The collection of various policy scenarios provides a foundation to ***compare processes*** of past policies in order to suggest a process for a new policy. Furthermore, provenance can lead to better understanding of the actual process (Carata et al, 2014).

iii) Sajjad (2014) discuss that policy formulation is a long-driven process. Howlett et al (2009), further report that policy making is not a straightforward process and often require revision and/or redoing during the process, this may incur delays during policy formation. Provenance information can reveal ***where the delay occurred*** and can provide insight about how it could be avoided in the future.

iv) Existing literature (Section 2.2.1) showed the complexity that is associated with studying a policy process. The stage model (Figure 1.1, chapter 1) simplifies the process. Sutcliffe (2005) further reports that stages have different evidence needs. How the different evidence is employed in a policy cycle can be studied by investigating past policies thus highlighting the need or provenance information gathering. Provenance data can show ***different types of evidence being used in different stages***.

v) Given a stage-based model of policy cycle, the stages of the cycle can be investigated alone or in terms of the relationships between one or all other phases (Howlett et al, 2009) to provide an understanding of the process. Provenance can reveal these ***stage details and the relationships amongst them***.

vi) The provenance can demonstrate intertwined ***roles of actors, ideas, and institutions in different phases of policy cycle*** (Howlett et al, 2009). This also acknowledges actors of their participation in various policies (Sajjad, 2014).

The literature uncovers the need for PPM provenance management. However, the design of a provenance framework entails an in-depth investigation of the PPM domain. This is important because the framework cannot be devised without having a clear understanding of the environment for which it is developed. The investigation of the policy domain revealed challenges which are supposed to be addressed in the provided solution. Given this, the next section discusses the PPM domain and related challenges. From the challenges, requirements for PPM provenance solution are identified which are also covered in the following section.

2.4 PPM Domain Challenges

PPM domain challenges are as follows:

i) **Challenge-1:** PPM is predominantly a ***knowledge-intensive*** (also called knowledge working) environment where humans play a key role in determining what is supposed to be done at any point of time. In knowledge intensive setups, human knowledge and information plays a central role and processes a secondary role. Given this fact, the process of developing policy rests on human knowledge and experience. The knowledge working nature of PPM induces additional challenges such as adhoc-ness and non-determinism, beforehand process orchestration, and complex task specification (Gong and Janssen, 2017; Hauder, 2016; Hallsworth, 2011).

ii) **Challenge-2:** PPM process is inherently ***ad hoc and non-deterministic*** in nature (Howlett et al, 2009). This implies that task specification and assignment to actors depends on the policy. For example, when issue is reported it can undergo discussion with stakeholders within government or with non-government officials. It may also require objective analysis such as statistical analysis. This suggests that the next piece of action largely depends on tacit knowledge of decision-makers. Moreover, Wu et al (2017) discuss that the order of execution of activities varies among policies, thus increasing challenges of defining a well-defined process.

iii) **Challenge-3:** Although, PPM and business processes share some similar characteristics but ***task specification*** in PPM is not straightforward and is rather complex (Sajjad, 2014). This is because PPM is largely non-deterministic and the decision of the next activity or its nature largely depends on human interpretation and policy requirements. This signifies that the knowledge working nature of public policies presents challenges for definitive task specification. The involvement of the large number of stakeholders further aggravates the complexity of task specification.

iv) **Challenge-4:** In comparison to business processes, ***a large number of actors*** participate in the policy creation process. The new modes of governance such as co-production, co-design, and co-management have introduced various new actors into PPM thus increasing complexity (Wu et al, 2017; Capano et al, 2015). This leads to process orchestration challenges since actors' involvement in policies vary, due to different policy needs, which aggravates the intricacy of tasks assignment and specification. Knill and Tosun (2012) argues that one of the reasons of complex role specification in PPM is that, often a government at the run-time creates a dedicated group/unit for a policy.

v) **Challenge-5:** Another prominent challenge that is related to PPM is the length of public policy making (Sajjad, 2014). Unlike business processes, PPM is a ***long-driven process*** and different policies may take different times for their conception and completion.

vi) **Challenge-6:** Given the unpredictable and knowledge intensive nature of PPM, ***beforehand process orchestration*** is not a viable option for devising policies (Gong and Janssen, 2017). This is because policies vary and may not require input from the same set of stakeholders or need the same set of activities as do other policies. Furthermore, given a large number of stakeholders (collocated and distributed including citizens) and their diverse and conflicting perceptions; emerging data/evidence; a long-driven process; and diverse and intersecting policy domains, it is challenging to collaboratively define and orchestrate a process.

Analysis

The domain challenges show the predominant effect of the knowledge working nature of PPM on other reported challenges. This is because creation of policies is largely dependent on humans' interpretation which is shaped by their ideas, judgment, objective evidence, and/or by discussion with other actors. Given this fact, the process is guided by the actors which introduces adhoc-ness and dynamicity in the process. The large number of human involvement and variable needs of policies contribute to the challenges of defining a well-defined process and explicitly specifying the process tasks. Having a large number of actors often introduce conflicts which results in delays during the process.

The provenance framework should be designed as such that the challenges are taken into account. This suggests exploring of any current technologies or mechanisms that are designed for addressing the dynamic and knowledge intensive setups. The next sections, therefore, focus on existing technologies or solutions and review their strengths and limitations in handling previously identified policy devising challenges.

2.5 Provenance Capture Mechanisms and Technologies

The existing provenance management systems have, generally, been designed by considering the requirements of scientific processes, business processes of private sector, and administrative tasks of public sector (Lin, C, 2008; Sajjad, 2014). The existing provenance systems details workflow technology as the most acceptable and popular solution for tracking processes. Furthermore, the philosophical stance of appropriateness of workflow technology for PPM is discussed by Sajjad (2014). Therefore, this section firstly covers (Section 2.5.1) workflow technology. Section 2.5.1 also covers the critical assessment of workflow technology which has been carried out in the light of the PPM domain challenges presented in the previous section. Second, case management (CM) is discussed in Section 2.5.2; these tools are specifically designed for knowledge intensive setups. Section 2.5.2 also critically analyses the suitability of case management tools. Lastly, Blockchain technology and critical appraisal is covered in Section 2.5.3.

The critical analysis for all the technologies is carried out from the perspective of KiPs. The extensive research by Hauder (2016) identified five main research challenges for KiPs i) knowledge worker empowerment ii) knowledge storage and extraction iii) authorisation and role management iv) theoretical foundation and v) data integration. This research focuses on first and second challenge and thus critical analysis only discuss the technologies from the perspective of these KiPs challenges.

2.5.1 Workflow Technology to Capture Provenance

A Workflow is a collection of repeatable coordinated tasks which represents a well-defined complex process (Haller et al, 2005; Mukherjee et al, 2004). A Workflow Management System (WFMS) is a generic information system designed to support modelling, execution, management, and monitoring of workflows (Haller et al, 2005). WFMS for business process manages and organises business processes (Thalheim and Tropmann-Frick, 2013). Before going into details of workflows, it is significant to mention here Process-aware Information Systems (PAISs). PAISs supports business processes by exploiting comprehensions from management science and information technology. A typical example of PAIS are WFMSs (Van Der Aalst et al, 2009). This implies that WFMSs are process-aware (Thalheim and Tropmann-Frick, 2013).

Workflow technology is widely exploited by scientific and business community. This work investigates both to draw conclusions regarding its suitability for public policy processes.

2.5.1.1 Provenance Capture of a Scientific Process

The need for tracking a scientific process has been identified due to the large amount of data that is generated and transformed by the execution of a scientific process. The outcome of scientific experiments is the result of transformations, manipulations, and algorithms/processes involved. A number of provenance tracking systems have been designed since provenance provides insight into how the scientific experiments were executed and how data was transformed; this signify scientific processes being ***data oriented.*** The data-oriented model is when provenance is specifically gathered about data product and its transformation (Simmhan et al, 2005). The provenance systems for scientific processes are therefore designed by keeping in view the data-oriented model (Lin, 2008).

The provenance of scientific data assist to reproduce an experiment (Davidson et al, 2007; Ludäscher et al, 2006).

Scientific workflows have been largely employed by the scientific community to orchestrate the process and to track its provenance. The provenance captured by scientific workflows usually includes the execution environment, input data, job logs, intermediary outputs, and the final output produced during the execution of a workflow. A number of provenance management systems including Chimera (Foster et al, 2002), Vistrail (Scheidegger et al, 2008), Taverna (Hull et al, 2006), myGrid (Stevens et al, 2003), Pegasus (Deelman et al, 2005), Kepler (Altintas et al, 2006), REDUX (Barga and Digiampietri, 2008), SciCumulus (de Oliveira et al, 2010), Pipeline-centric Provenance (Groth et al, 2009), SWIFT(Freire et al, 2008) are designed to support and capture the scientific workflows.

Section 2.4 reports human input at all points of PPM operations; these operations are not intended for manipulation and transformation of data to produce a certain output. This is unlike the scientific workflow where data is subjected to various operations to produce a specific output (Sajjad, 2014; Simmhan et al, 2005). Given this fact, scientific experiments can be repeated using the same procedure/process/algorithm. Repetition in PPM is not required to produce the known and calculated outcome. Furthermore, unlike scientific processes, PPM processes are control flow oriented (Sajjad, 2014) which makes the data-oriented nature of scientific workflow unsuitable to track PPM processes. Thus, this research does not further investigate any workflow technology which is specifically designed for managing scientific processes.

2.5.1.2 Provenance Capture of a Business Process

The potentials of workflow technology have paved the way for automating the business processes (Van der Aalst, 2013). Unlike scientific workflows, business workflows tend to be control flow oriented (also called process oriented) (Lin, C, 2008; Kiepuszewski et al, 2003). The control flow perspective models a sequence of activities and their order of execution to describe the flow of work through an organisation (Van der Aalst, 2013). In such setups, the data (modelling decisions, data creation etc.), resource (modelling roles, authorisation, organisational units etc.), time (modelling deadlines, duration etc.), and the functional (modelling activities) perspectives rests on the control flow perspective (Van der Aalst, 2013; Kiepuszewski et al, 2003).

Business process provenance focuses on organised way of collecting information to reconstruct what has happened in a process or organisation (Van der Aalst, 2013; Curbera et al, 2008). Generally, the provenance management systems for business process capture information such as business activities, data that was produced, resources involved etc. Research shows that the provenance of business data is required to demonstrate the events that incurred during process execution and give insight into cause-and-effect relation (Curbera et al, 2008). Provenance in business domain is deemed useful to provide audit trial for regulatory purposes (Simmhan et al, 2005). Provenance of business processes has also been exploited for mining (Van der Aalst, 2011).

Similar to business processes, PPM is also control flow oriented. Furthermore, it has been found from the literature (Sajjad, 2014) that PPM is closely related to business processes which suggests the use of business process-based workflow for tracking public policies. In this regard, this research only focuses on business workflows and investigation of their suitability for PPM.

2.5.1.3 Suitability of Business Workflow Technology to Track Public Policy making Processes

As discussed, in section 2.5.1.2, numerous workflow technologies exist for managing the business processes but what is important here is to evaluate the approach that underpins the business process workflows. Therefore, an enquiry of the types of business workflows is carried out. The different types are assessed in correspondence to the challenges and requirements identified in Section 2.4.

A study of business processes uncovers three main categories namely structured, semi-structured, and ad hoc (Burkhart and Loos, 2010; Dustdar et al, 2005; Huth et al, 2001; Kiepuszewski et al, 2000). The types reflect the processes that the workflow technology manages. Structured workflows entail well-defined process orchestration at design time whereas semi-structured and ad hoc workflows are subject to change at the run-time.

There exists no obvious boundary in semi-structured and ad hoc workflows and the terms are often used interchangeably. For example, Voorhoeve and Van der Aalst (1997) terms ad hoc workflows as those that are defined at design time but changes at run-time. This explanation however is employed by Rito (2011) to define semi-structured workflows. According to Burkhart and Loos (2010), ad hoc workflows specify execution path at the run-time and semi-structured falls somewhere between ad hoc and structured processes. Furthermore, some researchers (e.g., Huth et al, 2001) defined adhoc workflows as short-lived processes and semi-structured as those that are predetermined and recurrent. The variable conceptualisation, of ad hoc and semi-structured, offers no clear demarcation of the two categories. Therefore, to avoid confusion, this research termed such processes as ***flexible workflows***.

Structured workflows frame the well-defined process (hence called tightly framed processes (Van der Aalst, 2013)) at the design time and is suitable for setups where changes are infrequent (Stavenko et al, 2013). Nevertheless, from the challenges identified in section 2.4, it has been deduced that tracking a PPM process requires a system that can address the ad hoc and non-deterministic nature of PPM. The major drawback of structured workflow is limited users' influence on processes as they are defined during the modelling phase. The policy making challenges, however, draw attention towards large human involvement and inappropriateness of beforehand process orchestration which certainly suggests non-suitability of structured workflows (Pesic et al, 2007). Given this, only semi-structured and ad hoc workflows are investigated in detail in this research.

Extensive research has been done in the past to introduce flexibility in the workflows (Pesic et al, 2007; Zhao and Liu, 2007; Adams et al, 2006; Rinderle et al, 2004; Weske, 2001; Agostini De Michelis, 2000; Van der Aalst et al, 2000; Meng et al, 2000; Ellis et al, 1998, Voorhoeve and Van der Aalst, 1997; Ellis et al, 1995). This research discusses the recent advancements only (i.e., the last 10 years from 2009-2019).

Nahabedian et al (2019) employs declarative approach to facilitate dynamic reconfiguration of business processes. A declarative approach specifies constraints/rules (pre and post conditions) to drive model enactment and it describes dependency relationship between tasks rather than framing the sequence of actions. Changes in business processes result in workflow transition from the old towards the new. However, when the transition should take place remain an issue; this is because many workflow instances could be running on old business process models. Given the challenge, this work is centred on defining business process reconfiguration strategy. The strategies are specified as per business/organisation regulations and policies.

Vasilecas et al (2016) proposed rule and context based dynamic business processes modelling and simulation. In their approach, the business process is not defined as a sequence of activity; rather, predefined business rules and context are used to select the relevant activity at the run-time. Further, the authors employ goal-based approach to select an activity and guide the business process. The goals here are based on business rules which specify the target that a business process needs to achieve. Furthermore, log of all business processes is maintained to guide the future processes.

A case-based approach for workflow adaptation has been proposed by Minor et al (2014). Their approach records all changes in previous workflows' execution as cases. The cases record the reason of changes, what caused the change, and the workflow instance prior to adaptation. For realising change in the workflow, previous case(s) from the pool of available cases is automatically chosen. Only the closely related cases are chosen for workflow adaptation. The authors support processes adaptation at build time and run-time. The approach addresses the overhead related to process definition from the scratch.

Thalheim and Tropmann-Frick (2013) proposed the idea of generic workflows to derive the current workflows instance which represents the current situation. The generic workflows define a high-level view of possible workflows and serves as a template for the generation of another workflow at the run-time. The constraints specified in the generic workflow forms the basis for deriving current workflows.

Dadam and Reichert (2009) developed a software technology, called ADEPT2, for managing clinical processes. The clinical environment required process adaption with changing organisational needs. The use cases identified for devising workflow solution for the given challenge were i) facilitate ad hoc changes at the process instance level i.e., at the run-time and ii) enable process schema evolution at the design level. The ADEPT2 technology supports adhoc as well as process specification level changes and users can change the process by adding, moving, or deleting activities. The system also ensures that defined changes are correct.

Van Der Aalst et al (2009) developed a WFMS entitled DECLARE which uses a declarative approach to support loosely structured processes. The authors combined YAWL (highly structured workflow system) and DECLARE (loosely structured) to support structured processes with the loosely structured processes.

In addition to the aforementioned work, dynamicity in business process remained focus of research among many other researchers. For example, Hildebrandt and Mukkamala (2011) facilitates automatic construction of workflow given business requirements. Multiple researches have been done where business rules are used to induce dynamicity in the business processes (Milanovic al, 2011; Mejia Bernal et al, 2010; Hermosillo et al, 2010). Xiao et al (2011) represent dynamicity by introducing different versions of business processes. Adams (2010) facilitates run-time construction of business processes from reusable tasks.

Analysis

From the literature it seems that workflow technology is an attractive approach to track PPM. However, for the adoption of the workflow technology, it is important to assess the technology as per the requirements of PPM. Section 2.4 reports large contribution of knowledge-intensive nature of PPM on other identified challenges. Therefore, workflow technology is assessed to find its suitability for knowledge working environments.

The knowledge working setup, such as PPM, induces unpredictable and emergent circumstances (Hauder, 2016). It has been found from the research that workflow technology is not suitable for such systems (Hauder, 2016; Van Der Aalst, 2013). One of the contributing factors for the inflexibility of the workflow technologies is routing; this is because routing states what should be done rather than what can be done. As flexible workflows are also based on routing, so the mentioned problem is still prevalent in them thus making them infeasible to use for KiPs of PPM (Hauder, 2016). Moreover, PPM processes (covered in section 2.4) are unstructured and the workflow/business process community (e.g., Van Der Aalst, 2013) explored the inaptness of workflow technology for such processes.

The declarative approach to model business processes for KiPs seems an appropriate approach (Vaculín et al, 2011). However, PPM is not guided by any organisation or business rules rather, PPM is directed by stakeholders' values, interests, their understanding, as well as their conflicts with others. This added complexity of PPM processes makes it challenging to use a declarative approach.

As stated in Section 2.4, PPM processes are emergent (i.e., course of actions are revealed as process progresses), non-repeatable (i.e., each process instance looks slightly different), and unpredictable (i.e., subsequent action is situation-specific) which are also characteristics of KiPs. The researchers (e.g., Reichert and Weber, 2012) argues loose specifications for the given challenges.

Given the above-mentioned challenges, workflow is not an attractive approach for capturing PPM. Harrison-Broninski (2005) further reports the unsuitability of process-based approaches (such as workflows) for knowledge working setups. The author argues that knowledge-intensive environments require a solution where humans can guide processes whereas existing business process employs a process to direct humans.

The Knowledge-intensive-Processes (KiPs) of PPM instigated further research in such processes. Davenport (2005) concludes the importance of KiPs in today's organisations. Almost half of all papers on research in KiPs were published in 2012 and later thus indicating growing research interests in solutions that support KiPs (Hauder, 2016). In the past couple of years, researchers and practitioners have proposed various case management approaches to support of KiPs. Section 2.5.2 thus covers recent developments in case management (also called case handling) to find the suitability for PPM provenance tracking.

2.5.2 CM

CM is a process support paradigm for knowledge-intensive setups (Mutschler et al, 2008). The idea was first conceptualised by Van der Aalst et al (2001) to support flexible and knowledge intensive business processes. Heterogeneous terms (Hauder, 2016) are used for CM (Gerstner, 2014; Kaan et al, 2006), e.g., dynamic CM (Le Clair and Moore, 2009), production CM (Meyer et al, 2014), adaptive CM (Cano et al, 2015; Huber et al, 2015; Hauder et al, 2014; Motahari-Nezhad and Swenson, 2013), and emergent CM (Böhringer, 2010), case handling (Van der Aalst et al, 2005). This section firstly provides CM/ case handling overview then adaptive, dynamic, and emergent CM are briefly discussed.

When compared to workflow technology, CM offers less rigid process and a declarative approach is used to define execution paths (Reijers et al, 2003). The central concepts of CM are cases and data which are unlike the control flow structures of workflows (Hauder, 2016; Van der Aalst et al, 2001). The other potential differences from workflow technology are (i) no context tunnelling as knowledge workers can access the entire case data (ii) authorisation and distribution of work is kept separate and knowledge workers have additional roles in addition to execution of activities. For example, a knowledge worker can skip, or redo activities (iii) based on available data, decide the activity and (iv) knowledge workers can access and view the data independent from the activities i.e., view data before or after the corresponding activities (Van der Aalst et al, 2005). Figure 2.6 shows CM meta-model taken from Van der Aalst et al (2005).

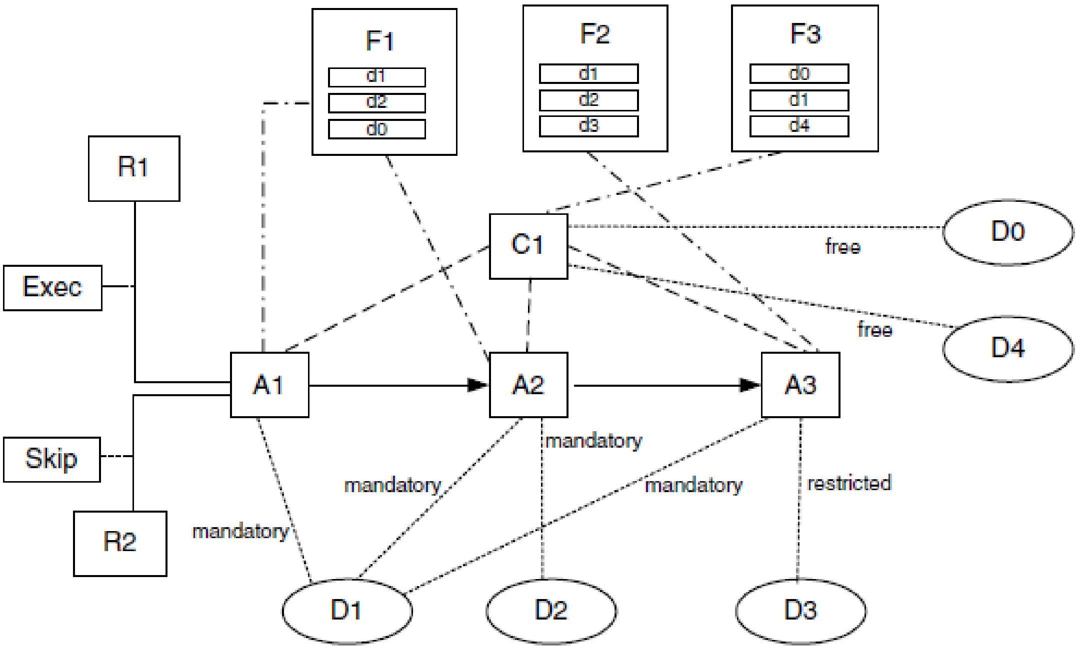

Figure 2.6: CM Meta-model (Van der Aalst et al, 2005)

C1 in the meta-model represent case definition that consists of activities (A1-A3), Data (D0-D4), forms (F1-F3), and roles (R1 and R2). Association between data definition and activity definition is represented by dotted line. Forms contain the most relevant data objects for the activities and case. The *mandatory* data objects in the meta-model represents the data that is relevant for a given activity but can be used by other activities also (e.g., D1 can be employed by A1 and by A2). If D1 is consumed by A2 then A1 can be skipped thus adding flexibility in the process. The *restricted* data object is tied to only one specific activity (e.g., In the figure D3 is relevant to A3 only). *Free* data objects are linked to case and can change anytime. Roles are attached to activities who can perform various actions on the activity such as execute or skip (other roles are presented in (Van der Aalst et al, 2005).

As mentioned previously, CM is data-oriented thus, data drives the process. Based on the available data, the activity is selected for the execution. Data here serves as a post-condition for an activity. For instance, A1 will be executed when D1 will be available. The detailed explanation of the meta-model with the relevant example is covered by Van der Aalst et al (2005).

The idea of the CM has attracted various research. Le Clair and Moore (2009) underpins the concept of dynamic CM which is defined as a highly structured, collaborative, dynamic, and KiP driven by events and responses from the case handler. The authors present various CM use cases including audit tracking, fraud, and electronic freedom of information act, biotechnology regulatory events, homelands security, and multichannel support for customers, and reducing customers' frustration over the phone and website. The various use cases provide an overview of potentials of CM however, the authors do not present any new mechanism or software solution. Dynamic CM facilitate knowledge workers to adapt on existing case structures which were defined at the build-time (Zensen and Küster, 2019).

Böhringer (2010) proposed emergent CM for ad hoc processes. The emergent CM is the bottom-up approach for managing ad hoc processes. The bottom-up approach facilitates users to assign activities and artefacts to a certain case. It is unlike top-down approach (Van der Aalst et al, 2005) which according to the status of the case provide users an information about a set of activities. The proposed bottom-up approach is based on the microblogging (e.g., twitter) and activity streams. The users publish information on microblog which is enhanced by joint activities (by other users) called activity stream. Once the artefacts and activities are placed in an activity stream then they can be assigned to cases. The idea is inspired from social systems such as twitter where a user tweets and others re-tweet or hashtag. Re-tweets and hashtags are activities' streams that are placed in a case and undergoes further discussion.

Swenson (2010) exploited both structured and emergent approaches of the processes. Software developers or modelling experts exploit a top-down approach to define structured processes and a bottom-up approach that gives liberty to knowledge workers to adapt where needed. The adaptive CM is emergent and slightly structured thus making it suitable for organisations that need to adapt their processes.

Herrmann and Kurz (2011) presents an adaptive CM solution to support loosely structured KiPs. Their approach consists of three subsequent phases (i) execution (ii) control and (iii) overarching case adaptation. The *execution* phase entails instantiating the given case. As per specific needs, adaptations can be realised before or during the case execution. Smaller adaptations can be directly incorporated while complex ones need case manager's approval. On finishing case execution, control phase starts. The *control* phase entails assessment of the executed case based on its efficiency and goal realisation. The final phase *overarching case adaptation* involves identification and selection of improvements which are generalised in the case template. The generalised templates represent the best practices that are maintained in the template library. Motahari-Nezhad and Swenson (2013) proposed an adaptive CM approach that focuses on end-users to define cases and does not rely on software developers for the task; the central idea of adaptive CM is also data-centric.

Cano et al (2015) developed an adaptive CM called Linkcare for improving healthcare services. The system provides (i) patient-centred management (ii) improved patient accessibility and (iii) manage coordination of health-care professionals. Linkcare identifies patients-centric cases by considering their health condition and medical needs. The health professionals evaluate the cases and generate work plans accordingly. Linkcare moreover, is an adaptive CM which facilitate changes during the execution of work plans.

Hauder (2016) also worked on the adaptive CM and extended Hybrid Wiki with new structuring concepts which are required to support KiPs. The researcher developed a software solution to facilitate non-expert end-users in collaborative and emergent structuring of KiPs and they require no knowledge about existing process modelling notation. The software assists users to modify and structure content and processes on Wiki. The solution, in specific, has been designed to empower knowledge workers. The adaptive CM system facilitates end-users to define cases at runtime while also giving modelling experts an autonomy to maintain templates for cases. *Work plans* are employed by knowledge workers to describe the course of actions using lightweight structuring concepts. The lightweight structuring process enable users to structure the process without needing high-level modelling knowledge such as Case Management Modelling Notation (CMMN). The work plans add adaptive and emergent behaviour while the structure is added in the software solution by *work template*. The work template represents reusable knowledge and best practises in an organisation including data and process structures. It also manages constraints among tasks in a work plan; this is used to specify restrictions in a system in parallel to the freedom that is provided to the knowledge workers through work plans. Different roles are defined to manage the KiPs such as *Authors* who can modify the structuring of KiPs in work plans and *Tailors* who require modelling capabilities to maintain the schema for KiPs i.e., manage work template. Furthermore, the software solution explicitly captures expertise of knowledge workers which assists in delegation and execution of tasks by the most appropriate knowledge worker.

Zensen and Küster (2019) created a java-based lightweight, adaptable, and platform independent application for CM which is based on CMMN. The framework supports major CMMN elements which is unlike Camunda (Camunda, 2019) that only supports partial CMMN and a part of business process engine. The framework enables developers to systematically implement CM applications based on CMMN model. A CM application can be assembled by reusing the components of the framework.

A notion of production CM was put forward by Meyer et al (2014). The production CM framework combines features of static process management and adaptive CM. In their approach, the complex business process is represented into smaller process components. The process components model a normal path (with no exceptions and additional details), exceptional behaviour(s), and a path with additional details. The data objects are used as pre and post conditions for process components. Based on the user requirement, an appropriate path is chosen at the run-time for the execution. Moreover, multiple process components can be used to fulfil one user request.

Analysis

The literature shows that CM systems have specifically been designed by considering the knowledge-intensive environments thus suggesting the appropriateness of their adoption. However, CM may not necessarily address all the challenges related to change (Van der Aalst et al, 2005) therefore, it is important to evaluate the aptness of CM for PPM.

As discussed previously, central concept of CM are cases which incorporate roles, processes, and data. The specification of cases for PPM also requires identification of the mentioned case elements. In PPM, cases can be specified using phases of the policy cycle. However, the cases in PPM cannot be disjoint rather, they require relation with each other. Constraints such as execution order among phases needs to be taken into account. For instance, implementation phase cannot be executed before the agenda setting phase.

Second, there are feedback loops in policy making (see Figure 2.3 and further elaboration in Chapter 4, Figure 4.18) which suggests explicit specification of connection among cases. Furthermore, unlike in CM, loop-back does not mean redo or erase what has been done previously; it may require addition of new knowledge while keeping the previous discussions. The reason being that prior knowledge might have led to the current state of the policy process. This adds to the complexity as it introduces high cohesion among potential PPM cases.

There can also be scenarios where cases may execute in parallel. For instance, phase 1 of the policy cycle can execute in parallel with phase 2. In such a scenario, there is a requirement for continuous collaboration between cases thus aggravating the complexity.

The case definition entails process specification. As mentioned in section 2.4, task specification in PPM is rather complex. Furthermore, policies require different processes and variable set of actors. This induces challenges of defining case processes for PPM. Furthermore, activities in a case may need a link with activities of another case. Parallel execution of cases further exacerbates the complexity of defining relation between the activities of different cases. For every policy, new activities, and process (for cases) orchestration is needed. This induces an overhead in terms of processes definition and re-definition for different policies. Storage of cases is also complicated because management of links among cases is required whilst storing.

Adaptive CM introduces more flexibility than dynamic and production CM due to declarative approach. However, analysis on workflow shows the associated challenges of using the declarative approach for PPM.

Le Clair and Moore (2009) further discusses the benefit of CM for scenarios where a case is managed from start to finish by a knowledge worker. The author argues that CM provides benefit in terms of reducing the workforce. However, the scale of PPM processes and actors is significantly higher than other business setups. PPM processes require input from diverse stakeholders thus less workforce is practically not possible.

Emergent case management (Böhringer, 2010) seems suitable, but PPM is rather complex, and blogging may limit the interaction of decision makers. The policy makers need to be communicated directly for their input on the required policy. Furthermore, due to security concerns often policies are not shared with all stakeholders.

The different CM approaches including dynamic, production, and adaptive tells that their use is based on the KiPs application need. The detailed analysis of CM and its adoption for PPM shows that CM provides additional challenges in respect of case definition, storage, process and business rules specification, and role resolution. Therefore, for this research, CM is not considered for capturing PPM provenance data.

In the recent years, Blockchain technology is gaining momentum for tracking provenance information. Therefore, following sub-section briefly discuss the Blockchain technology from the perspective of provenance capture and storage.

2.5.3 Provenance Capturing Using Blockchain Technology

The foundation of Bitcoin Blockchain was laid by Satoshi Nakamoto in 2008. Blockchain is a distributed ledger technology which enables secure transactions without requiring centralised trust party (Tasatanattakool and Techapanupreeda, 2018). Blockchain is also an interactive environment for building distributed applications (Ramachandran and Kantarcioglu, 2018).

Decentralised storage of data by different peers in the network forms the basis of Blockchain technology. Transactions form a basic unit of work in Blockchain. The transaction is recorded as a block but in some cases, it is recorded into an auxiliary data store rather than in the block. Several connected blocks are called chain. A new block is only added in the chain if there is a consensus among peers in the network. Each block holds the cryptographic hash function of the previous block in the chain, timestamp, nonce, and the transaction data. How the business among parties is conducted is supported by a Smart Contract. It is an executable code that define rules and constraints in the contract; the Smart Contracts are deployed on the shared ledger and it works on the application-level. When any event that is outlined in the contract is triggered then the code is executed. An example of Smart Contract is managing the activities of purchase order (Di Ciccio et al, 2018; Hull et al, 2016).

The promises of Blockchain have opened new opportunities for managing processes. Weber et al (2016) showed the feasibility of Blockchain technology for business process management. The work implements automated translation of BPMN model to Smart Contract. López-Pintado et al (2019) developed a dynamic role binding approach to support organisational perspective of process on Blockchain. The researchers proposed a dynamic role binding approach to help actors dynamically nominate others for the execution of an activity. Sturm et al (2019) used Blockchain-based workflow management engine to support inter-organisational business processes. The researchers used Smart Contract to execute business process on the Blockchain. The Smart Contract implements and enforce the control-flow specified in BPMN. For every business process execution, smart contract is deployed first. Their approach fills the gap of trust in cross-organisational processes. Hull (2016) adopted artefact-centric approach (as in case management) to design Business Collaboration Language (a form of Domain Specific Language) to create Smart Contracts in Blockchain.

Blockchain provides reliable, shared, trusted, transparent, non-repudiate-able data repositories (Di Ciccio et al, 2018; Hull et al, 2016). These feature of Blockchain facilitate trustworthy data provenance collection. SmartProvenance is a distributed Blockchain based scientific data provenance system that validates the changes before storing them (Ramachandran and Kantarcioglu, 2018). The authors use Blockchain as a medium for trusted provenance storage. SmartProvenance is built using Smart Contracts for automated generation and verification of scientific provenance data. SmartProvenance consists of two types of Smart Contracts i) Document Tracker Contract which maintains access control policies for provenance trial generation of a particular document ii) Vote Contract to implement the voting policy among peers and the contract is invoked when any change in the document is required. Given the costly storage on Blockchain, the work stores documents in the Cloud.

Di Ciccio et al (2018) employed Blockchain for inter-organisational business process traceability. The Blockchain provides an underlying infrastructure for business process management. Smart Contracts are used to manage business processes on the Blockchain. Smart Contracts here act as a container for business processes and holds control-flow logic of workflows. During process execution, a new block is added to the chain whenever an activity is executed. A block contains transaction details (i.e., executed activity), smart contract address, sender address, and transaction data containing the hash code.

Neisse et al (2017) analysed the feasibility of Blockchain technology for General Data Protection Regulation (GDPR). The researchers proposed several solution design choices for data accountability and provenance management of GDPR. Three main roles are defined for GDPR that includes person, organisation with whom person is registered, and other organisations with whom person data is shared by the organisation. Customised contracts are used for different people.

Provenance is tracked when person data is accessed and when the data is forwarded to other organisations. The solution is designed from the perspective of person privacy who can access how his/her data has been exploited by organisations with whom she/he has registered and with whom the data is shared.

In addition to business setups, Blockchain-based provenance management has also paved its way for other setups such as for the Cloud computing. Liang et al (2017) proposed a blockchain based data provenance system called ProvChain for cloud auditing. ProvChain track changes that occur in cloud environments and generate events that corresponds to users' actions. The user events are recorded as transaction on the Blockchain.

Analysis

The current literature shows future trends of Blockchain for business processes management. The recent advancements of Blockchain-based business processes management reveal that it requires provision of application-level details on Blockchain**.** This implies that for capturing the PPM processes using the Blockchain technology, an application is required that specifies all details of PPM process alongside control flow and data flow constraints. Though, Blockchain records provenance information as blocks but for the management of PPM, transaction details need to be encoded in Smart Contracts.

Blockchain, in addition, provides verification of data before recording it which provides trust among peers. The security and trust features of the technology makes it very attractive to use for provenance management. This research, however, does not focus on the security aspects of provenance data. Here the focus is on the application level that represents and captures PPM provenance information. Furthermore, it has been found by the researchers (e.g., Sturm et al, 2019) that Blockchain aptness for any domain require an in-depth investigation. This research only focuses on designing an approach for PPM provenance management. Therefore, the assessment of Blockchain for PPM provenance management is out of scope of this research and Blockchain technology is therefore not considered for this research.

Based on the technologies' assessment in this section, Section 2.8 discusses critical gap analysis and introduce a possible solution approach in Section 2.9. Section 2.4 showed the dynamic nature of PPM which recommends a flexible and an adaptive approach for provenance tracking. A model-driven approach has proven to be suitable for such settings (Bocanegra et al, 2016), therefore, the next section evaluates its relevance for PPM provenance capturing.

2.6 A Model-driven Approach for PPM Provenance Management

A model-driven approach provides different views of a system with varying levels of abstraction (Seebacher and Maleshkova, 2018). A model-driven approach captures a system as reusable models.

It facilitates in common understanding among technical and non-technical stakeholders as it separates business logic from platform implementation. Central concepts of a model-driven approach are models and meta-models. A model is an abstraction of a system under investigation which shows properties of interest from a specific viewpoint. It represents simplified and partial view of a system. The model captures entities and relationship among them. A meta-model is also a model that defines the structure of models. Models are instance of meta-models (Bocanegra et al, 2016; Da Silva, 2015).

A model-driven approach facilitates complex techniques such as meta-modelling, model transformation, and automatic code generation from the models. Numerous terms are used for model-driven such as Model-driven Architecture (MDA), Software Factories, Domain Specific Language (DSL) Engineering. The terms are generically classified as Model-Driven Engineering (MDE). Irrespective of the adopted term, all of them share common concepts (Da Silva, 2015). In general, a model-driven approach captures models and enables transformation of models to code or to other models (Bocanegra et al, 2016).

A model-driven solution facilitates changes at the run-time and is suitable for uncertain situations. Using the approach, the adaptive behaviour can be introduced at different stages of software engineering process including requirements, design, software process, and quality assurance process (Bocanegra et al, 2016). Furthermore, it enhances productivity, maintainability, reusability, and interoperability (Corredor et al, 2012).

A model-driven approach has shown a great deal of promise for assisting government in various cities' operation. Bucchiarone and Cicchetti (2018) developed and implemented a model-based solution for Smart Journey Planning. The variable situations such as weather conditions, available mobility resources, urban environment conditions, new sustainability rules by government, or change in government policies often results in refinement of the current city journey planner. To address the aforementioned gaps, a model-driven based flexible journey planner for smart cities is designed by the authors.

A model-driven based solution for smart spaces is conceptualised by Corredor et al (2012). With the advancement of technology, smart spaces can be designed by exploiting various smart objects such as mobile phones, specific sensors. However, connecting the heterogeneous technologies for the purpose of smart space remains a challenging task. Therefore, the proposed model-driven based solution provides common language and procedures for connecting Internet of Things and Web of Things for the purpose of designing smart spaces.

The benefits surrounding a model-driven approach has also attracted its application in modelling of public policies (Scherer et al, 2015). In the given research work, the existing or new policy documents are used to produce policy simulation models. For a given policy case, the policy analyst chooses the relevant policy aspects to be modelled; the conceptual models are based on the meta-models which encompasses structure and behaviour elements. The structural elements are actors, objects, attributes, and relationship between model elements. Behaviour elements encode dynamic aspects such as actions of an actor and pre and post conditions. The output of the policy model's simulation are narrative scenario texts and visualisation of numerical data.

Loukis (2007) developed an ontology for supporting public policies design, implementation, and evaluation. The ontology structures electronic Government-to-Government collaboration. The main objective of the ontology is creation, management, indexing, retrieval, reuse, and exchange of knowledge. It is a horizontal ontology which can be used with vertical ontologies such as education, planning, tourism etc.

Analysis

The literature shows the aptness of a model-driven approach for dynamic and ad hoc setups. As identified in section 2.4, PPM is inherently dynamic and uncertain which indicates the suitability of a model-driven approach for capturing PPM provenance data. Though, the current survey (Scherer et al, 2015) shows that its potential has already been realised in public policies, but it has not been exploited for PPM processes. The current literature uses existing or new policies for simulation purposes in order to aid the understanding of the already devised policies. However, how the policies came into existence is not the focus of their work. Furthermore, their work is confined to policy devising phase only and implementation and evaluation is not considered by them.

Loukis (2007) encode concepts of policy devising, implementation, and evaluation. However, the paper does not represent provenance specificities that could show the operations that were carried out for PPM. The ontology does not explicitly specify the details of generic policy cycle phases such as stakeholders' information, the communication among stakeholders, participation of different stakeholders in each policy phase, phase related concepts, transition (linear flow and loop -backs) among policy cycle phases, and relation of the phases with each other. As PPM is a KiP therefore, the meta-models must also specify the concepts related to tracking approach that is being adopted. The current work does not consider the aforementioned details thus, signifying the gap in the existing work.

2.7 PROV Model for PPM Provenance Representation

The literature covered in the previous sections show the requirement of provenance for managing PPM. The provenance representation requires a standard notation to provide a common understanding among various actors/stakeholders. At present, the widely adopted provenance standard is PROV (2019) model. PROV representation covers entities (any physical, digital, or conceptual thing), people, and processes that produces a piece of provenance data. This research adopts PROV model for representing PPM provenance data (discussed further in Chapter 5, Section 5.4). The focus here is not to extend the existing PROV model or design any new standard rather the aim here is to employ the existing standard. Detailed description of PROV is presented in PROV (2019).

2.8 Critical Gap Analysis

Findings from the literature review reveals a list of research gaps that points to the need of PPM provenance framework. The research gaps are centred on (i) a provenance tracking facility for PPM (ii) PPM provenance in smart cities context (iii) support mechanism for KiPs of PPM and (iv) a model-driven approach to manage and capture dynamic and ad hoc PPM processes. Discussion for each gap is detailed below.

Section 2.1 shows that ***a provenance tracking facility is currently not available for PPM*** **(Research Gap-1)**. Section 2.3 states several PPM provenance use cases which highlights the benefits that are associated with PPM provenance capturing. The use cases show that provenance for PPM requires the tracking of process details including how and why the process was carried out, timestamp and duration for completion of a task to show the delays in a process, data representing different evidence being used in policy creation, phase details and their relation, and actors' participation. Due to lack of such facility, the government rely on various means for managing the PPM data (Sajjad, 2014).

The several use cases demand holistic representation of provenance information to find out what happened during PPM process. The literature shows a lack of holistic PPM provenance framework that can enable and manage policy creation process, store provenance information, and relate all the provenance concepts, facilitate actors' communication with each other, and helps in PPM provenance querying.

Section 2.2 clearly demonstrated the relation of governance with PPM. The detailed discussion shows that ***no support is currently available that can manage the provenance data of PPM in Smart Cities context*** **(Research Gap-2) i.e., no smart cities' solution for decision-making is available.** As covered in Sections 2.2.3 and 2.4, humans play a significant role in PPM. Furthermore, the changing governance dynamics demand input from diverse stakeholders. Therefore, the provenance framework needs to support the various policy making models including top-down, bottom-up, and hybrid. These models encourage varying levels of stakeholders' participation in the process.

From the process of investigating the literature on generic policy cycle (Section 2.2.1), it was explored that the cycle is a valuable resource for managing policy-making provenance data (which includes process details also). The generic policy cycle provides additional layer of abstraction for the policy-making process i.e., whatever, the process and approach (top-down, bottom-up, or hybrid) is being followed falls under the umbrella of general description of the generic policy cycle. This makes the evidence more useful and yet meaningful for the decision-makers. Given the potentials of generic policy cycle, this research employs it to manage the policy-making processes.

The need for PPM provenance framework uncovered the need of a provenance capturing mechanism. For this, Section 2.5 covered in detail the various technologies/mechanism that are currently available for supporting processes and capturing provenance data. The current mechanisms/technologies have been critically assessed in relation to the PPM challenges and requirements (presented in Section 2.4). The detailed investigated and analysis showed that the available approaches cannot be fully adopted for PPM. Thus, **Research Gap-3** is ***currently no provenance capturing mechanism for KiPs of PPM is available*****.**

The PPM domain challenges in section 2.4, uncovered the dynamic and ad hoc nature of PPM. The provenance being generated by policies may not be same and varies as per policies' needs. Furthermore, the collected provenance may require analysis and understanding from different standpoints. Section 2.6 shows the suitability of a model-driven approach for such settings. Therefore, this research aims at ***using a model-driven approach to support evidence-based policy making in smart cities which has not been realised before*** **(Research Gap-4)**. For representing the provenance, PROV model has been adopted.

In summary, this research focuses on provision of a provenance framework for supporting PPM. The diverse evidence being associated with the PPM is organised and managed using a generic policy cycle. The provenance framework supports the diverse actors and facilitate top-down, bottom-up, and hybrid approaches to PPM. The dynamic and ad hoc nature of PPM is supported by a model-driven approach. Finally, a new provenance capturing mechanism for PPM has been designed.

2.9 Realisation of Research Gaps

The Research Gap-1 produces the PPM provenance framework that provides a holistic view of provenance capturing, storage, and querying. The framework is presented in Chapter 4 (Section 4.3). To answer the Research Gap-2, the provenance framework facilitates participation of different stakeholders, further covered in Chapter 4 and Chapter 5. Furthermore, case studies are designed as such that different modes of stakeholder participation are covered (Appendix B). For Research Gap-3, a novel process-agnostic approach for PPM provenance framework has been proposed and is presented in Chapter 4 (Section 4.4). The associated meta-model details of the approach are covered in Chapter 5. Research Gap-4 is addressed by using a model-driven approach for the provenance framework. The meta-models are specified which captures the diverse evidence, actors' details, communication among actors, and details of a process-agnostic approach. How the meta-models are used in the provenance framework is detailed in Chapter 4 whereas Chapter 5 covers the meta-models. The technical and implementation details of the framework is covered in chapter 5. Chapter

6 is dedicated to evaluation of the provenance framework which has been designed in response to the research gaps identified in Section 2.8.

2.10 Summary of the Chapter

This chapter covered in detail the state of the art from the perspective of PPM, smart cities and PPM, provenance and its role in PPM, process capturing technologies, and a model-driven based solution. The detailed analysis demonstrated the need for a PPM provenance tracking facility in the context of smart cities. As public policies are impacted by the governance principles. Therefore, smart governance aspects of smart cities are considered. To capture provenance, detailed investigation of current technologies was carried out which signified the need of a new approach to facilitate KiPs of PPM. Lastly, the literature uncovered the suitability of a model-driven approach for supporting the dynamic and ad hoc setup of PPM. Based on the research gaps identified in this chapter, Chapter 4 presents the provenance framework which is based on a model-driven and a process agnostic approach. Furthermore, this chapter looked for various future directions which are presented in Chapter 7. Before going into the details of the provenance framework, it is significant to discuss how the research was carried out therefore Chapter 3 discusses the research design and the research methodology.

Chapter 3 : Methodology and Research Design

To respond to the early-identified research gaps, this chapter presents the research approach which defines a systematic way to inform the objectives of this research. The research approach explains the research phases, their sequence, and their output. This chapter first, provides an overview of research approach (Section 3.1). Second, a research strategy is elaborated in Section 3.2. A brief overview of research method is presented in Section 3.3. Last, for a systematic research process, research methodology followed in this research is sketched in Section 3.4.

3.1 Research Approach

There are two types of research approaches known as deductive and inductive research. Deductive research is a top-down approach. It is concerned with the formulation of a research hypothesis based on existing theory and then designing a research strategy to test the hypothesis. It is moving from a broad spectrum to a specific conclusion. In comparison to deductive research, inductive research is a bottom-up approach. In this approach, a researcher starts from a specific observation and develops a general conclusion or theory (Kothari, 2009). In the context of this research, a deductive approach has been employed as the focus of this research is not to build a theory. Deductive research follows the following steps (i) theory (ii) hypothesis (iii) observation and (iv) confirmation or rejection of a research hypothesis. This research investigated in detail current studies to identify gaps in knowledge and to understand the public policy processes. Based on the collected data, a hypothesis has been proposed which is tested by various experiments (the details of which are covered in chapter 6, section 6.6).

The research approach is a map that guides a researcher in approaching research, however, to construct knowledge and to test a hypothesis it is significant to identify a research strategy. The research strategy for this research is elaborated in the following section.

3.2 Research Strategy

A research strategy specifies the plan to conduct research and to select a research method for data collection and analysis. There exist various research strategies for data collection including case studies, interviews, experiments, survey, longitudinal studies, ethnography, action research, and grounded theory (Saunders et al., 2007; Cavaye, 1996).

To design a provenance framework for tracking public policy processes, an in-depth investigation of the Public Policy Making (PPM) environment is needed. The modelling of provenance requires the identification of PPM concepts and the relationships among them. Thus, to capture these details, an appropriate selection of a research strategy is important. Existing research strategies include qualitative, quantitative, and a mixed method approach. A quantitative approach (Creswell, 2009; Hakim, 2000) is suitable for research studies which require statistical verification. A qualitative approach (Creswell, 2009; Denzin and Lincoln, 1994) on the other hand is appropriate for in-depth inspection of phenomenon under investigation; it examines phenomena in its natural setting and acknowledges the influence of culture, people, or environmental forces on the phenomenon under study. Mixed method research includes both qualitative and quantitative approaches (Creswell, 2009).

PPM is largely guided by humans; thus, it is important to acknowledge the human involvement in the public policy devising. Detailed enquiry is also needed to uncover the concepts and their relationships for the modelling. Therefore, it is significant to understand the PPM setup. Given the nature of the problem, a qualitative approach is an appropriate consideration. This is because modelling PPM requires a rigorous analysis of the policy environment to understand and model various concepts and their associations. Due to required details for the modelling, a quantitative approach has not been considered in this research.

Based on the research strategy, the next section elaborates the research method adopted for this research.

3.3 Research Method

The qualitative approach provides an in-depth investigation, but it is not without limitations (Miles and Huberman, 1994). For example, individual perceptions are captured in qualitative research to understand the phenomenon which certainly are influenced by researchers' own understanding; this leads to a potentially biased interpretation of the data. Furthermore, the qualitative approach also presents a limitation in respect of generalisation of the findings. To address these concerns, a multi-method approach (also called triangulation) is adopted to reduce possible errors and fallacy in the constructed knowledge. For this research, a literature review and experts' suggestions were considered for the data collection. How the literature review and experts' suggestion were sought in this research is covered in the subsequent section.

3.4 Overview of Methodology and Research Design

This section elaborates the research methodology for this research. The methodology guides to conduct research in a systematic manner. The research methodology for this research is presented in Figure 3.1.

The research methodology comprises six phases which include (i) inception (ii) proposition (iii) design and modelling (iv) poof-of-concept development (v) evaluation and (vi) communication.

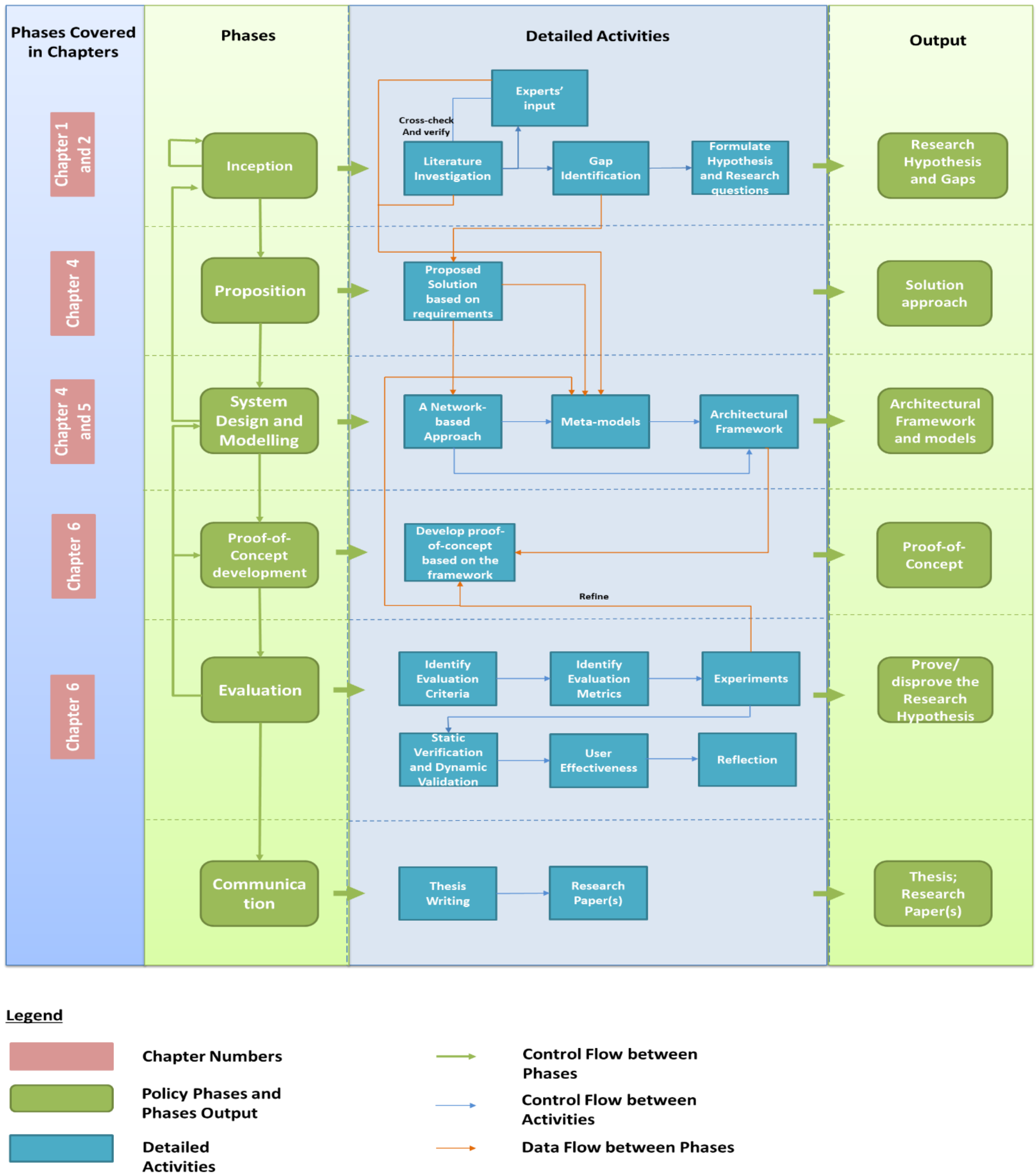

Figure 3.1: The Research Methodology

Inception

This research started with the literature survey which was carried out with the purpose of exploring the research gaps. Literature review was conducted in the domain of PPM, provenance, existing application of provenance in PPM, current mechanisms and technologies to support PPM, smart cities, and a model-driven approach.

An in-depth study related to PPM setup was carried out to derive initial observations which clarified the challenges associated with PPM (Chapter 2, Section 2.4). As discussed in the previous section, multiple sources of data were used for building knowledge. Therefore, the initial observations that were drawn up regarding the PPM environment were verified with public policy experts. Furthermore, the input of experts that were recorded was also cross-checked with the literature. This was done to address the possible biases of interviewees and the researcher. Furthermore, this phase shows the self-loop, this is because firstly the observations were drawn from the literature, and then the experts' suggestions were sought for data collection. The observations were cross verified; the process that was followed for cross-verification is as follows: (i) first the literature review was conducted to uncover important details regarding policy environment (ii) the experts' suggestions were collected and (iii) experts' suggestions were further explored using the literature.

The benefit attached to performing literature review before the interviews is the identification of relevant interviewees. A careful selection was made regarding the group of people that were interviewed in order to gather information about PPM. To ensure the fairness, policy experts were chosen who had prior knowledge of developing policies. PPM experts were therefore thought to be appropriate candidates for this exercise. For this research, the number of experts considered were three. The chosen experts have several years of experience in PPM and were mainly involved in devising land-use policies. The experts have experience of many years of working in the local councils. To ensure anonymity and confidentiality of the interviewees, their details have not been included in this thesis.

Based on the study, research gaps were identified (Chapter 2, Section 2.8). The literature and the research gaps assisted in the formation of the research hypothesis and the research questions (presented in Chapter 1 (Section 1.4)).

Proposition

The observations drawn in the previous stage helped in the identification of the characteristics of the provenance tracking system. This stage proposed the solution's needs based on the identified research gaps. Further details are covered in chapter 4.

System Design and Modelling

The proposition phase sets the foundation for the system design and modelling. Based on the detailed investigation, a process-agnostic approach and a model-driven approach were used to build the Policy Cycle Provenance (PCP) framework. The methodology shows that this phase loops back to the first phase. The reason for this is that for the formulation of the meta-models an in-depth investigation of PPM process is required. At this phase, the experts' input on the case studies (details covered in evaluation phase of the methodology) were sought. The input helped in further improving the meta-models. The data collected from the mentioned sources helped in cross-verification and enhanced overall understanding. A systematic approach for designing meta-models was followed and is further discussed in chapter 4 whereas chapter 5 covers meta-models in detail.

Proof-of-Concept Development

Based on the system architecture, the proof of concept was developed. The implementation was done using Java language and the PROV toolbox. Details are covered in Chapter 5, Section 5.6.

Evaluation

This phase outlines the approach for evaluating the research hypothesis. For this evaluation, a criterion was formulated which defines the metrics for the evaluation. The metrics used for evaluation are: (i) completeness which tests capability of the framework to support generic policy cycle (ii) correctness to evaluate actual and expected results and (iii) repeatability to test the framework under same/similar cases. Evaluation criteria is presented in Chapter 6 (Section 6.1). To evaluate the PCP framework, three hypothetical case studies (Appendix B) were employed; reasons for using the hypothetical case studies are covered in Chapter 6 (Section 6.2).

Researchers (e.g., Bonoma, 1985) have argued that a single case study is best applied for theory generation. For theory testing and/or exploration, multiple case studies are ideal. According to Yin (2003), single case study is more useful when (i) testing a critical case for a well-established theory (ii) single case is representative (iii) a phenomenon previously developed is not available to research and (iv) longitudinal study.

For this research, a single case study is not appropriate because the focus here is not to build a theory, the study is not longitudinal, and the PCP framework is not conceptualised previously by other researchers. Furthermore, the PCP framework is not developed for a specific policy therefore testing multiple PPM scenarios is needed to evaluate the generality of the framework. A single case study is not representative enough to test and explore the diverse capabilities of the PCP framework. What is required is to assess the capability of the framework by evaluating it against multiple PPM scenarios. However, there exist diverse policy types and the literature (on public policies) does not document their processes; secondly, PPM processes are not constant (discussed in Section 2.4). Moreover, identification of all possible scenarios for different policy types is practically not possible in the time span of this research. Therefore, this research adopts the most widely reported generic approaches i.e., top-down, bottom-up, and hybrid approach. These approaches depict disparate PPM scenarios, and they are independent of any policy type.

For the evaluation of the PCP framework, three case studies (presented in Appendix B), that articulate top-down, bottom-up, and a hybrid approach to PPM, are identified using a detailed investigation of the literature. These hypothetical case studies are built by studying policy making processes that have been reported in the literature (such as in Tsohou et al, 2012). The scenarios are elaborated using detailed literature on the three approaches (i.e., top-down, bottom-up, and hybrid) and the overall public policy making setup. As PCP framework is independent of any policy type therefore, it does not account that which policy type is used for evaluation. Moreover, the aim of this research does not state the evaluation requirement for a specific policy. Thus, for the purpose of this research, urban planning policies are employed for testing the framework.

Case study-1 is *Shops opening hours' extension policy*. This scenario has been formulated using the case study presented in Tsohou et al (2012). Some modification, as per literature, in the case study has been made to test various capabilities of the framework (Dunn, 2018; Wu et al, 2017; Höchtl et al, 2016; Sonntagbauer et al, 2015; Khan et al, 2014; Anderson, 2014; Cairney, 2012; Hill, 2013; Anderson, 2014; Cairney, 2012; Knill and Tosun, 2012; Kingdon and Stano, 2011; Howlett eta al, 2009, Macintosh, 2004; Weimer and Vining, 2004).

Case study-2 is *Air Quality Improvement policy* is devised using the literature (Dunn, 2018; Hottenstein, 2017; Wu et al, 2017; Höchtl et al, 2016; Sonntagbauer et al, 2015; Khan et al, 2014; Anderson, 2014; Cairney, 2012; Hill, 2013; Anderson, 2014; Tsohou et al, 2012; Cairney, 2012; Knill and Tosun, 2012; Kingdon and Stano, 2011; Howlett et al, 2009; Macintosh, 2004; Weimer and Vining, 2004).

Finally, **Case study-3** is *Local Planning Policy in Decision-making* captures the bottom-up approach to PPM (Wu et al, 2017; Höchtl et al, 2016; Sonntagbauer et al, 2015; Rodríguez-Bolívar, 2015; Scholl, 2014; Khan et al, 2014; Nam and Pardo, 2011; Howlett et al, 2009; Edelenbos and Klijn, 2005; Macintosh, 2004).

The evaluation process is further elaborated in detail in chapter 6. The evaluation results in proving/disproving the research hypothesis.

Communication

This phase focuses on thesis writing and research publications.

3.5 Summary of the Chapter

This chapter presents the research design and methodology. For this research, Deductive approach and Qualitative research strategy are found appropriate. Further, this chapter gives the detailed outline of the research methodology. How multiple information sources are used in this research is discussed in detail. The systematic process led to the formulation of the provenance framework,

References

Adams, M., Ter Hofstede, A.H., Edmond, D. and Van Der Aalst, W.M. (2006) Worklets: A service-oriented implementation of dynamic flexibility in workflows. *In OTM Confederated International Conferences" On the Move to Meaningful Internet Systems" (pp. 291-308). Springer, Berlin, Heidelberg.*

Adams, M. (2010) Dynamic workflow. In: (2010) *Modern Business Process Automation* [online]. Springer, pp.123-145.

Agostini, A. and De Michelis, G. (2000) Improving flexibility of workflow management systems. In: (2000) *Business Process Management* [online]. Springer, pp.218-234.

Agrawal, P., Benjelloun, O., Sarma, A.D., Hayworth, C., Nabar, S., Sugihara, T. and Widom, J. (2006) Trio: A system for data, uncertainty, and lineage. *Proc.of VLDB 2006 (Demonstration Description)* [online].

Altintas, I., Barney, O. and Jaeger-Frank, E. (2006) Provenance collection support in the kepler scientific workflow system. *In International Provenance and Annotation Workshop (pp. 118-132). Springer, Berlin, Heidelberg.*

Anand, M.K., Bowers, S. and Ludäscher, B. (2010) Techniques for efficiently querying scientific workflow provenance graphs. *In EDBT (Vol. 10, No. 2010, pp. 287-298).*

Anderson, J.E. (2014) *Public Policymaking* [online]. Cengage Learning.

Angelidou, M. (2014) Smart city policies: A spatial approach. *Cities* [online]. 41 pp.S3-S11.

Anthopoulos L and Reddick C G. (2016). Smart City and Smart Government: Synonymous or Complementary? *Proceedings of the 25th International Conference Companion on World Wide Web, Pages 351-355.*

Barga. R and Digiampietri. L. (2008) Automatic capture and efficient storage of e-Science experiment provenance. *In Journal of Concurrency and Computation: Practice & Experience, vol. 20 Issue 5, pp. 419-429.*

Barns, S. (2018) Smart cities and urban data platforms: Designing interfaces for smart governance. *City, culture and society*, *12*, pp.5-12.

Bartoletti, R. and Faccioli, F. (2016) Public Engagement, Local Policies, and Citizens' Participation: An Italian Case Study of Civic Collaboration. *Social Media+ Society*, *2*(3), p.2056305116662187.

Billings, R.S. and Hermann, C.F. (1998) Problem identification in sequential policy decision making: The re-representation of problems. *Problem representation in foreign policy decision making*, pp.53-79.

Bocanegra, J., Pavlich-Mariscal, J. and Carrillo-Ramos, A. (2016) September. On the role of model-driven engineering in adaptive systems. In *2016 IEEE 11th Colombian Computing Conference (CCC)* (pp. 1-8). IEEE.

Böhringer, M. (2010) Emergent case management for adhoc processes: a solution based on microblogging and activity streams. *In International Conference on Business Process Management (pp. 384-395). Springer, Berlin, Heidelberg.*

Bonoma, T. V. (1985) Case research in marketing: opportunities, problems, and a process. *Journal of Marketing Research, 22, pp. 2199-208.*

Bose, R. and Frew, J. (2005). Lineage retrieval for scientific data processing: a survey. *ACM Computing Surveys (CSUR)*, *37*(1), pp.1-28.

Branson, A., McClatchey, R., Le Goff, J.M. and Shamdasani, J. (2014) CRISTAL: A practical study in designing systems to cope with change. *Information Systems*, *42*, pp.139-152.

Brewer, G.D. (1974) The policy sciences emerge: to nurture and structure a discipline. *Policy sciences*, pp.239-244.

Bridgman, P. and Davis, G. (2004) *The Australian policy handbook*. Allen & Unwin.

British Standard Institute. PAS 181: (2014) Smart City Framework—Guide to Establishing Strategies for Smart Cities and Communities. *British Standard Institute: London, UK, 2014. Available online: http://www.bsigroup.com/en-GB/smart-cities/Smart-Cities-Standards-and-Publication/PAS-181-smart-cities-framework/* [Accessed on 30 August 2019].

Bryman, A. and Bell, E. (2007) Business research methods. *USA: Oxford University Press.*

Bucchiarone, A. and Cicchetti, A. (2018) A Model-Driven Solution to Support Smart Mobility Planning. *In Proceedings of the 21th ACM/IEEE International Conference on Model Driven Engineering Languages and Systems (pp. 123-132). ACM.*

Burkhart, T. and Loos, P. (2010) Flexible business processes-evaluation of current approaches. *Multikonferenz Wirtschaftsinformatik 2010*, p.243.

Cairney, P. (2012) Understanding Public Policy: Theories and Issues. *Basingstoke: Palgrave Macmillan*.

Camunda. Workflow and decision automation platform. https://camunda.com [Accessed September 20, 2019].

Cano, I., Alonso, A., Hernandez, C., Burgos, F., Barberan-Garcia, A., Roldan, J. and Roca, J. (2015) An adaptive case management system to support integrated care services: lessons learned from the NEXES project. *Journal of biomedical informatics, 55, pp.11-22.*

Capano, G., Howlett, M. and Ramesh, M. eds. (2015) *Varieties of governance: Dynamics, strategies, capacities*. Springer.

Carata, L., Akoush, S., Balakrishnan, N., Bytheway, T., Sohan, R., Selter, M. and Hopper, A. (2014) A primer on provenance.

Cavaye, A. (1996) Case study research: A multifaceted research approach for IS. *Information Systems Journal, 6 (3), pp. 227-242.*

Chavan, A., Huang, S., Deshpande, A., Elmore, A., Madden, S. and Parameswaran, A. (2015) Towards a unified query language for provenance and versioning. *In 7th {USENIX} Workshop on the Theory and Practice of Provenance (TaPP 15).*

Cheney, J., Ahmed, A. and Acar, U.A. (2007) September. Provenance as dependency analysis. In *International Symposium on Database Programming Languages* (pp. 138-152). Springer, Berlin, Heidelberg.

Chorley, A., Edwards, P., Hielkema, F., Philip, L. and Farrington, J. (2008) Supporting provenance and argumentation in evidence-based policy assessment. *In Proceedings of the Oxford eResearch Conference.*

Chun, S. and Cho, J.S. (2012) E-participation and transparent policy decision making. *Information Polity, 17(2), pp.129-145.*

Cirianni, F., Panuccio, P. and Rindone, C. (2013) A comparison of urban planning systems between the UK and Italy: Commercial development and city logistic plan. *WIT Transactions on the Built Environment, 130,* pp.785-797.

Cobb, R., Ross, J.K. and Ross, M.H. (1976). Agenda building as a comparative political process. *American political science review, 70(1), pp.126-138.*

Corredor, I., Bernardos, A.M., Iglesias, J. and Casar, J.R. (2012) Model-driven methodology for rapid deployment of smart spaces based on resource-oriented architectures. *Sensors, 12(7), pp.9286-9335.*

Creswell, J.W. (2009) Research design: Qualitative, quantitative, and mixed methods approaches *(Vol. 4). Thousand Oaks, CA: Sage.*

Curbera, F., Doganata, Y., Martens, A., Mukhi, N.K. and Slominski, A. (2008) Business provenance–a technology to increase traceability of end-to-end operations. In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"* (pp. 100-119). Springer, Berlin, Heidelberg.

Dadam, P. and Reichert, M. (2009) The ADEPT project: a decade of research and development for robust and flexible process support. *Computer Science-Research and Development, 23(2), pp.81-97.*

Daniell, K.A., Morton, A. and Insua, D.R. (2016) Policy analysis and policy analytics. *Annals of Operations Research*, *236*(1), pp.1-13.

Da Silva, A.R. (2015) Model-driven engineering: A survey supported by the unified conceptual model. *Computer Languages, Systems & Structures, 43, pp.139-155.*

Davenport, T.H. (2005) Thinking for a living: how to get better performances and results from knowledge workers. *Harvard Business Press.*

Davidson, S., Cohen-Boulakia, S., Eyal, A., Ludäscher, B., McPhillips, T., Bowers, S., Anand, M.K. and Freire, J. (2007) Provenance in Scientific Workflow Systems. *Data Engineering Bulletin, vol. 30, no. 4, pp. 44–50.*

Deelman, E., Singh, G., Su, M.H., Blythe, J., Gil, Y., Kesselman, C., Mehta, G., Vahi, K., Berriman, G.B., Good, J. and Laity, A. (2005) Pegasus: A framework for mapping complex scientific workflows onto distributed systems. *Scientific Programming*, *13*(3), pp.219-237.

De la Vara González, J.L. and Diaz, J.S. (2007) Business process-driven requirements engineering: a goal-based approach. *In Proceedings of the 8th Workshop on Business Process Modeling (p. 15).*

De Marchi, G., Lucertini, G. and Tsoukiàs, A. (2016) From evidence-based policy making to policy analytics. *Annals of Operations Research, 236(1), pp.15-38.*

Denzin, N. Y. K. and Lincoln, Y. (1994) Handbook of Qualitative Research. *SAGE Publications, London, UK.*

de Oliveira, D., Ogasawara, E., Baião, F. and Mattoso, M. (2010) Scicumulus: A lightweight cloud middleware to explore many task computing paradigm in scientific workflows. In *2010 IEEE 3rd International Conference on Cloud Computing* (pp. 378-385). IEEE.

Di Ciccio, C., Cecconi, A., Mendling, J., Felix, D., Haas, D., Lilek, D., Riel, F., Rumpl, A. and Uhlig, P. (2018) Blockchain-based traceability of inter-organisational business processes. *In International Symposium on Business Modeling and Software Design (pp. 56-68). Springer, Cham.*

Ding, L., DiFranzo, D., Graves, A., Michaelis, J.R., Li, X., McGuinness, D.L. and Hendler, J. (2010) Data-gov wiki: Towards linking government data. In *2010 AAAI Spring Symposium Series*.

Ding, L., Lebo, T., Erickson, J.S., Difranzo, D., Williams, G.T., Li, X., Michaelis, J., Graves, A., Zheng, J.G., Shangguan, Z. and Flores, J. (2011) TWC LOGD: A portal for linked open government data ecosystems. *Web Semantics: Science, Services and Agents on the World Wide Web, 9(3), pp.325-333.*

Dunn, W. (2018). Public Policy Analysis. *6th edition. New York: Routledge, https://doi.org/10.4324/9781315181226*

Dustdar, S., Hoffmann, T. and Van der Aalst, W. (2005) Mining of adhoc business processes with TeamLog. *Data & Knowledge Engineering*, *55*(2), pp.129-158.

Edelenbos, J. and Klijn, E.H. (2005) Managing stakeholder involvement in decision making: A comparative analysis of six interactive processes in the Netherlands. *Journal of public administration research and theory*, *16*(3), pp.417-446.

Ellis, C., Keddara, K. and Rozenberg, G. (1995) August. Dynamic change within workflow systems. In Proceedings of conference on Organizational computing systems (pp. 10-21). ACM.

Ellis, C.A., Keddara, K. and Wainer, J. (1998) Modeling workflow dynamic changes using timed hybrid flow nets. In Workflow Management: Net-based Concepts, Models, Techniques, and Tools (WFM'98), Proceedings (pp. 109-128).

Everett, S., 2003. The policy cycle: democratic process or rational paradigm revisited?. *Australian Journal of Public Administration*, *62*(2), pp.65-70.

Foster, I., Vockler, J., Wilde, M. and Zhao, Y. (2002) Chimera: A virtual data system for representing, querying, and automating data derivation. In *Proceedings 14th International Conference on Scientific and Statistical Database Management* (pp. 37-46). IEEE.

Freire, J., Koop, D., Santos, E. and Silva, C.T. (2008) Provenance for computational tasks: A survey. *Computing in Science & Engineering, 10(3), pp.11-21.*

Gerstner, M. (2014) Empowering End-users to Support Knowledge-intensive Processes with the Case Management Model and Notation. *Master's thesis. Technische Universität München. Germany.*

Gong. Y and Janssen, M. (2017) The Impact of Social Business Process Management on Policy making in e-Government. *In Proceedings of the 50th Hawaii International Conference on System Sciences, Waikoloa Village, HI, USA.*

Groth. P et al. (2009). Pipeline-centric provenance model. *In the Proceedings of the 4th Workshop on Workflows in Support of Large-Scale Science (WORKS'09).*

Haas, P.J. and Springer, J.F. (2014) Applied policy research: Concepts and cases. *Routledge.*

Hakim, C. (2000) Research Design: Successful Designs for Social and Economic Research. *2nd, Routledge, London.*

Haller, A., Oren, E. and Petkov, S. (2005) Survey of workflow management systems. *Technical Report 2005-05-02, DERI.*

Hallsworth, M. (2011). Policy Making in the Real World: Evidence and Analysis. *Institute for Government: London, UK.*

Hammad, R. (2018) A hybrid e-learning framework: Process-based, semantically-enriched and service-oriented. *Doctoral dissertation, University of the West of England.*

Harrison-Broninski, K. (2005) Human Interactions: The Heart and Soul of Business Process Management: How People Reallly Work and How They Can Be Helped to Work Better. *Meghan-Kiffer Press: Tampa, FL, USA.*

Harrison. C and Donnelly. I. A. (2011) A theory of smart cities. *In the 55th Annual Meeting of the International Society for the Systems Sciences, 17-22.*

Hauder, M., Pigat, S. and Matthes, F. (2014). Research challenges in adaptive case management: a literature review. In *2014 IEEE 18th International Enterprise Distributed Object Computing Conference Workshops and Demonstrations* (pp. 98-107). IEEE.

Hauder, M. (2016) Empowering End-Users to Collaboratively Structure Knowledge-Intensive Processes. *Doctoral dissertation, Technische Universität München.*

Hermosillo, G., Seinturier, L., and Duchien, L. (2010) Using Complex Event Processing for Dynamic Business Process Adaptation. *In 2010 IEEE International Conference on Services Computing, pages 466–473.*

Herrmann, C. and Kurz, M. (2011) Adaptive case management: Supporting knowledge intensive processes with it systems. *In International Conference on Subject-Oriented Business Process Management (pp. 80-97). Springer, Berlin, Heidelberg.*

Hildebrandt, T.T. and Mukkamala, R.R. (2011) Declarative event-based workflow as distributed dynamic condition response graphs. *arXiv preprint arXiv:1110.4161.*

Hill, Michael. (2013). The public policy process (6th ed). *Harlow: Pearson.*

Hill, Michael. (2014) Studying Public Policy: An International Approach. *1st ed., Bristol University Press, Clifton, Bristol, UK; Chicago, IL, USA, 2014, pp. 295–328. JSTOR, www.jstor.org/stable/j.ctt1t898hb.30.*

Höchtl et al. (2016) Big data in the policy cycle: Policy decision making in the digital era. *Journal of Organizational Computing and Electronic Commerce, 26:1-2, 147-169, DOI: 10.1080/10919392.2015.1125187.*

Hofferbert, R. (1974). The Study of Public Policy. *Indianapolis, IN, Bobbs-Merrill.*

Holland, D.A., Braun, U.J., Maclean, D., Muniswamy-Reddy, K.K. and Seltzer, M.I. (2008) Choosing a data model and query language for provenance. In *Proceedings of the 2nd International Provenance and Annotation Workshop (IPAW'08)*. Springer.

Homsy, G.C., Liu, Z. and Warner, M.E. (2019) Multilevel governance: Framing the integration of top-down and bottom-up policy making. *International Journal of Public Administration*, *42*(7), pp.572-582.

Hottenstein, K.N. (2017) Protecting the Teaching and Learning Environment: A Hybrid Model for Human Subject Research Public Policy Implementation. *Journal of Research Administration, 48(2), pp.26-36.*

House, J.S., Schoeni, R.F., Kaplan, G.A. and Pollack, H. (2008). The health effects of social and economic policy: the promise and challenge for research and policy. *Making Americans healthier: Social and economic policy as health policy, pp.3-26.*

Howard, C. (2005) The policy cycle: a model of post-Machiavellian policy making?. *Australian Journal of Public Administration, 64(3), pp.3-13.*

Howlett, M., Ramesh, M. and Perl, A. (2009) Studying public policy: Policy cycles and policy subsystems *(Vol. 3). Oxford: Oxford University Press.*

Huber, S., Fietta, M. and Hof, S. (2015) Next step recommendation and prediction based on process mining in adaptive case management. *In Proceedings of the 7th International Conference on Subject-Oriented Business Process Management (p. 3). ACM.*

Hull, D., Wolstencroft, K., Stevens, R., Goble, C., Pocock, M.R., Li, P. and Oinn, T. (2006) Taverna: a tool for building and running workflows of services. *Nucleic acids research*, *34*(suppl_2), pp.W729-W732.

Hull, R., Batra, V.S., Chen, Y.M., Deutsch, A., Heath III, F.F.T. and Vianu, V. (2016) Towards a shared ledger business collaboration language based on data-aware processes. *In International Conference on Service-Oriented Computing (pp. 18-36). Springer, Cham.*

Huth, C., Smolnik, S. and Nastansky, L. (2001) Applying topic maps to ad hoc workflows for semantic associative navigation in process networks. In *Proceedings Seventh International Workshop on Groupware. CRIWG 2001* (pp. 44-49). IEEE.

Huynh, T.D., Ebden, M., Venanzi, M., Ramchurn, S.D., Roberts, S. and Moreau, L. (2013) Interpretation of crowdsourced activities using provenance network analysis. In *First AAAI Conference on Human Computation and Crowdsourcing*.

IP - RFC 791 - https://tools.ietf.org/html/rfc791 [Accessed June 3, 19].

Jann, W. and Wegrich, K. (2007) Theories of the policy cycle. *Handbook of public policy analysis: Theory, politics, and methods, 125, pp.43-62.*

Jenkins-Smith, H.C. and Sabatier, P.A. (1993). The study of public policy processes. *Policy change and learning: An advocacy coalition approach, pp.1-9.*

Kaan, K., Reijers, H.A. and van der Molen, P. (2006) Introducing case management: opening workflow management's black box. *In International Conference on Business Process Management (pp. 358-367). Springer, Berlin, Heidelberg.*

Karvounarakis, G., Ives, Z.G. and Tannen, V. (2010) Querying data provenance. *In Proceedings of the 2010 ACM SIGMOD International Conference on Management of data (pp. 951-962). ACM.*

Kay, A. (2011) Evidence-Based Policy-Making: The Elusive Search for Rational Public Administration. *Australian Journal of Public Administration, 70(3), pp.236-245.*

Khan, Z. (2009) Bridging the Gap between Business Process Models and Service-Oriented Architectures with Reference to the Grid Environment. *PhD Thesis, University of the West of England (UWE).*

Khan, Z et al. (2014) ICT enabled participatory urban planning and policy development: The UrbanAPI project. *Transforming Government: People, Process and Policy, 8 (2). pp. 205-229. ISSN 1750-6166.*

Kiepuszewski, B., Ter Hofstede, A.H.M. and Bussler, C.J. (2000) On structured workflow modelling. *In International Conference on Advanced Information Systems Engineering (pp. 431-445). Springer, Berlin, Heidelberg.*

Kiepuszewski, B., ter Hofstede, A.H. and van der Aalst, W.M. (2003) Fundamentals of control flow in workflows. *Acta Informatica, 39(3), pp.143-209.*

Kingdon, J. (1984) Agendas, Alternatives and Public Policies. *Little Brown, Boston.*

Kingdon, John W., & Stano, Eric. (2011). Agendas, alternatives, and public policies (Updated 2nd ed, Vol. Longman classics in political science). Boston: Longman.

Kiser, Larry L., and Elinor Ostrom. (1982) The Three Worlds of Action: A Meta-theoretical Synthesis of Institutional Approaches. *In Strategies of Political Inquiry, ed. Elinor, Ostrom. Beverly Hills, CA: SAGE Publications, pp. 179–222.*

Knill, C. and Tosun, J. (2012) Public policy: A new introduction. *Macmillan International Higher Education.*

Kothari, C.R. (2004) Research methodology: Methods and techniques. *New Age International.*

Kourtit, K., Nijkamp, P. and Steenbruggen, J. (2017) The significance of digital data systems for smart city policy. *Socio-Economic Planning Sciences, 58, pp.13-21.*

Lakshmanan, G.T. et.al, (2011). Guest editors' introduction: Provenance in web applications. *IEEE Internet Computing, 15(1), pp.17-21.*

Lasswell, Harold D. (1956) The Decision Process: Seven Categories of Functional Analysis. *College Park: University of Maryland Press.*

Le Clair, C. and Moore, C. (2009) Dynamic Case Management—An Old Idea Catches New Fire. *Forrester Research.*

Liang, X., Shetty, S., Tosh, D., Kamhoua, C., Kwiat, K. and Njilla, L. (2017) May. Provchain: A blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability. *In Proceedings of the 17th IEEE/ACM international symposium on cluster, cloud and grid computing (pp. 468-477). IEEE Press.*

Lin, C., Lu, S., Lai, Z., Chebotko, A., Fei, X., Hua, J. and Fotouhi, F. (2008) Service-oriented architecture for VIEW: a visual scientific workflow management system. In *2008 IEEE International Conference on Services Computing* (Vol. 1, pp. 335-342). IEEE.

Liu, Q et al. (2013) Data Provenance and Data Management Systems. *In Data Provenance and Data Management in eScience, Springer Berlin Heidelberg.*

López-Pintado, O., Dumas, M., García-Bañuelos, L. and Weber, I. (2019) Dynamic role binding in blockchain-based collaborative business processes. *In International Conference on Advanced Information Systems Engineering (pp. 399-414). Springer, Cham.*

Loukis, E.N. (2007) An ontology for G2G collaboration in public policy making, implementation and evaluation. *Artificial Intelligence and Law, 15(1), pp.19-48.*

Ludäscher, B., Altintas, I., Berkley, C., Higgins, D., Jaeger, E., Jones, M., Lee, E.A., Tao, J. and Zhao, Y. (2006) Scientific workflow management and the Kepler system. *Concurrency and Computation: Practice and Experience, 18(10), pp.1039-1065.*

Lyden Fremont J., George A. Shipman, and Robert W. Wilkinson. (1968) Decision- flow analysis: A methodology for studying the public policy-making process. *In P.P. Le Breton, ed., Comparative Administrative Theory. Seattle: University of Washington Press, 155-68.*

Macintosh, A. (2004) Characterizing e-participation in policy making. *In Proceedings of the 37th Hawaii International Conference on System Sciences, Big Island, HI, USA, 5–8 January 2004.*

March, J.G. and Olsen, J.P. (1996) Institutional perspectives on political institutions. *Governance, 9(3), pp.247-264.*

Margo. D and Smogor. R. (2010) Using provenance to extract semantic file attributes. *In: Proceedings of the 2nd conference on Theory and practice of provenance.*

Marinho, A., Murta, L., Werner, C., Braganholo, V., Cruz, S.M.S.D., Ogasawara, E. and Mattoso, M. (2012) ProvManager: a provenance management system for scientific workflows. *Concurrency and Computation: Practice and Experience, 24(13), pp.1513-1530.*

McClatchey. R., et.al. (2014). Analysis Traceability for Biomedical Researchers. *In Computer Based Medical Systems.*

Meijer, A. and Bolívar, M.P.R. (2016) Governing the smart city: a review of the literature on smart urban governance. *international review of administrative sciences*, *82*(2), pp.392-408.

Mejia Bernal, J.F., Falcarin, P., Morisio, M. and Dai, J. (2010) Dynamic context-aware business process: a rule-based approach supported by pattern identification. *In Proceedings of the 2010 ACM Symposium on Applied Computing (pp. 470-474). ACM.*

Mellouli, S.; Luna-Reyes, L.F.; Zhang, J. (2014) Smart government, citizen participation and open data. *Inf. Polity, 19, 1–4.*

Meng, J., Helal, S. and Su, S. (2000) An ad-hoc workflow system architecture based on mobile agents and rule-based processing. *In Proceedings of the 2000 international conference on artificial intelligence (ICAI2000), Las Vegas* (pp. 245-251).

Meyer, A., Herzberg, N., Puhlmann, F. and Weske, M. (2014) Implementation framework for production case management: Modeling and execution. *In 2014 IEEE 18th International Enterprise Distributed Object Computing Conference (pp. 190-199). IEEE.*

Milanovic, M., Gasevic, D., and Rocha, L. (2011). Modeling Flexible Business Processes with Business Rule Patterns. *In 2011 IEEE 15th International Enterprise Distributed Object Computing Conference, pages 65–74.*

Miles and Huberman, (1994). Qualitative data analysis: A sourcebook. *Sage, Thousand Oaks, CA.*

Minor, M., Bergmann, R. and Görg, S. (2014) Case-based adaptation of workflows. *Information Systems, 40, pp.142-152.*

Motahari-Nezhad, H.R. and Swenson, K.D. (2013) Adaptive case management: Overview and research challenges. *In 2013 IEEE 15th Conference on Business Informatics (pp. 264-269). IEEE.*

Mukherjee, S., Davulcu, H., Kifer, M., Senkul, P. and Yang, G. (2004) Logic-based approaches to workflow modeling and verification. *In Logics for emerging applications of databases (pp. 167-202). Springer, Berlin, Heidelberg.*

Murta, L., Braganholo, V., Chirigati, F., Koop, D. and Freire, J. (2014) noWorkflow: capturing and analyzing provenance of scripts. *In International Provenance and Annotation Workshop (pp. 71-83). Springer, Cham.*

Mutschler, B., Weber, B. and Reichert, M. (2008). Workflow management versus case handling: results from a controlled software experiment. *In Proceedings of the 2008 ACM symposium on Applied computing (pp. 82-89). ACM.*

Nahabedian, L., Braberman, V., D'ippolito, N., Kramer, J. and Uchitel, S. (2019) Dynamic Reconfiguration of Business Processes. *In International Conference on Business Process Management (pp. 35-51). Springer, Cham.*

Nam, T. and Pardo, T.A. (2011) Conceptualizing smart city with dimensions of technology, people, and institutions. In *Proceedings of the 12th annual international digital government research conference: digital government innovation in challenging times* (pp. 282-291). ACM.

Nam, T.; Pardo, T.A. (2014). The changing face of a city government: A case study of Philly311. *Gov. Inf. Q. 2014, 31, S1–S9.*

Neisse, R., Steri, G. and Nai-Fovino, I. (2017) A blockchain-based approach for data accountability and provenance tracking. *In Proceedings of the 12th International Conference on Availability, Reliability and Security (p. 14). ACM.*

Parag, Y., 2008. Who governs the air we breathe? Lessons from Israel's industrialist covenant. *Journal of Environmental Policy & Planning*, *10*(2), pp.133-152.

Parkhurst, J. (2017) The politics of evidence: from evidence-based policy to the good governance of evidence (p. 182). *Taylor & Francis.*

Parycek, P. and Pereira, G.V. (2017) Drivers of Smart Governance: towards to evidence-based policy-making. *In Proceedings of the 18th Annual International Conference on Digital Government Research (pp. 564-565). ACM.*

Patton, C. V. and Sawicki, D. S. (1993). Basic Methods of Policy Analysis & Planning. *(2nd ed.). Prentice Hall: New Jersey.*

Pesic, M., Schonenberg, H. and Van der Aalst, W.M. (2007) Declare: Full support for loosely-structured processes. *In 11th IEEE International Enterprise Distributed Object Computing Conference (EDOC 2007) (pp. 287-287). IEEE.*

Peters, B.G. and Pierre, J. (2006). The three action levels of governance: Re-framing the policy process beyond the stages model. *Handbook of public policy*, pp.13-30.

Philip, L., Chorley, A., Farrington, J. and Edwards, P. (2007) Data provenance, evidence-based policy assessment, and e-social science. In Third international conference on e-social science.

Plain English guide to the Planning System. (2015) *Department for Communities and Local Government.*

PostgreSQL, 2019. https://www.postgresql.org/ [Accessed July 20, 2019]

PROV, 2019. https://www.w3.org/TR/prov-primer/ [Accessed: July 14, 2019].

ProvToolbox, 2019. https://lucmoreau.github.io/ProvToolbox/ [Accessed: February 15, 2019]

Prpić, J., Taeihagh, A. and Melton, J. (2015) The fundamentals of policy crowdsourcing. *Policy & Internet, 7(3), pp.340-361.*

Ram, S and Liu, J. (2009) A new perspective on Semantics of Data Provenance. *First International Workshop on the role of Semantic Web in Provenance Management (SWPM).*

Ramachandran, A. and Kantarcioglu, M. (2018) SmartProvenance: a distributed, blockchain based dataprovenance system. *In Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy (pp. 35-42). ACM.*

Reichert, M. and Weber, B. (2012) Enabling flexibility in process-aware information systems: challenges, methods, technologies. *Springer Science & Business Media.*

Reijers, H.A., Rigter, J.H.M. and van der Aalst, W.M. (2003) The case handling case. *International Journal of Cooperative Information Systems, 12(03), pp.365-391.*

Ribino, P., Lodato, C. and Cossentino, M. (2017) Modeling business rules compliance for goal-oriented business processes. *In Workshop on Enterprise and Organizational Modeling and Simulation (pp. 83-99). Springer, Cham.*

Rinderle, S., Reichert, M. and Dadam, P. (2004) Correctness criteria for dynamic changes in workflow systems—a survey. *Data & Knowledge Engineering, 50(1), pp.9-34.*

Silva, A.R. (2011). A blended workflow approach. In *International Conference on Business Process Management* (pp. 25-36). Springer, Berlin, Heidelberg.

Rodríguez-Bolívar, M.P. (2015). Smart Cities: Big Cities, Complex Governance? In *Transforming city governments for successful smart cities* (pp. 1-7). Springer, Cham.

Ruiz, M., Costal, D., España, S., Franch, X. and Pastor, O. (2015) GoBIS: An integrated framework to analyse the goal and business process perspectives in information systems. *information systems, 53, pp.330-345.*

Sabatier, Paul A. (1991) Toward Better Theories of the Policy Process. *Political Science and Politics, 24(2), pp. 147–56.*

Sabatier, P. and Jenkins H. S. (1988) An advocacy coalition framework of policy change and the role of policy-oriented learning there in. *Policy Sciences, 21, pp. 129-168.*

Sajjad, F. (2014) An Analytical Framework for Workflow Technology Adoption in Local Government Authorities. *Brunel Business School, Brunel University: London, UK.*

Saunders, M., Lewis, P. and Thornhill, A. (2007) Research Methods for Business Students. *4 th edn, Prentice Hall.*

Scheidegger, C., Koop, D., Santos, E., Vo, H., Callahan, S., Freire, J. and Silva, C. (2008) Tackling the provenance challenge one layer at a time. *Concurrency and Computation: Practice and Experience*, *20*(5), pp.473-483.

Scherer, S. et al. (2015) Evidence Based and Conceptual Model Driven Approach for Agent-Based Policy Modelling. *Journal of Artificial Societies and Social Simulation.*

Scholl, H.J.; Scholl, M.C. (2014) Smart Governance: A Roadmap for Research and Practice. *In Proceedings of the iConference 2014, Berlin, Germany, 4–7 March 2014; pp. 163–176.*

Seebacher, S. and Maleshkova, M. (2018) A model-driven approach for the description of blockchain business networks. *In Proceedings of the 51st Hawaii International Conference on System Sciences.*

Shaxson, L. (2005). Is your evidence robust enough? Questions for policy makers and practitioners. *Evidence & Policy: A Journal of Research, Debate and Practice*, *1*(1), pp.101-112.

Simmhan, Y.L., Plale, B. and Gannon, D. (2005) A survey of data provenance techniques. *Computer Science Department, Indiana University, Bloomington IN, 47405, p.69*

Simmons, R.H., Davis, B.W., Chapman, R.J. and Sager, D.D. (1974) Policy flow analysis: A conceptual model for comparative public policy research. *Western Political Quarterly*, *27*(3), pp.457-468.

Sindhu, U.; Balaouras, S.; Dines, R.A.; Hayes, N.M.; Nelson, L. (2010) Helping CIOs Understand "Smart City" Initiatives: Defining the Smart City, Its Drivers, and the Role of the CIO. *Forrester Research: Cambridge, MA, USA.*

Sobeck, J. (2003) Comparing policy process frameworks: What do they tell us about group membership and participation for policy development?. *Administration & Society*, *35*(3), pp.350-374.

Sonntagbauer, S., Sonntagbauer, P., Nazemi, K. and Burkhardt, D. (2014) The FUPOL policy lifecycle. In *Handbook of Research on Advanced ICT Integration for Governance and Policy Modeling* (pp. 61-87). IGI Global.

Stavenko, Y., Kazantsev, N. and Gromoff, A. (2013) Business process model reasoning: from workflow to case management. Procedia Technology, 9, pp.806-811.

Stevens R.D et al. (2003) myGrid: personalised bioinformatics on the information grid. *In Bioinformatics, vol. 19, pp. 302-304.*

Stone, D.A. (1989) Causal stories and the formation of policy agendas. *Political science quarterly, 104(2), pp.281-300.*

Sturm, C., Scalanczi, J., Schönig, S. and Jablonski, S. (2019) A Blockchain-based and resource-aware process execution engine. *Future Generation Computer Systems, 100, pp.19-34.*

Sutcliffe, S. (2005) Evidence-based policy making: What is it? How does it work? What relevance for developing countries?(No. Folleto 1427).

Swenson, K. D. (2010) Mastering the unpredictable: How adaptive case management will revolutionize the way that knowledge workers get things done. *Meghan-Kiffer Press.*

Tasatanattakool, P. and Techapanupreeda, C. (2018) Blockchain: Challenges and applications. *In 2018 International Conference on Information Networking (ICOIN) (pp. 473-475). IEEE.*

Thalheim, B. and Tropmann-Frick, M. (2013) Mini story composition for generic workflows in support of disaster management. *In 2013 24th International Workshop on Database and Expert Systems Applications (pp. 36-40). IEEE.*

Tsohou, A., Lee, H., Al-Yafi, K., Weerakkody, V., El-Haddadeh, R., Irani, Z., Ko, A., Medeni, T.D. and Campos, L.M. (2012) Supporting Public Policy Making Processes with Workflow Technology: Lessons Learned From Cases in Four European Countries. *International Journal of Electronic Government Research (IJEGR)*, *8*(3), pp.63-77.

Vaculín, R., Hull, R., Heath, T., Cochran, C., Nigam, A. and Sukaviriya, P. (2011) Declarative business artifact centric modeling of decision and knowledge intensive business processes. *In 2011 IEEE 15th International Enterprise Distributed Object Computing Conference (pp. 151-160). IEEE.*

Van der Aalst, W.M., Basten, T., Verbeek, H.M.W., Verkoulen, P.A. and Voorhoeve, M. (2000) Adaptive workflow. *In Enterprise Information Systems (pp. 63-70). Springer, Dordrecht.*

Van der Aalst, W.M. and Berens, P.J.S. (2001) Beyond workflow management: product-driven case handling. *In Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work (pp. 42-51). ACM.*

Van der Aalst, W.M., Weske, M. and Grünbauer, D. (2005) Case handling: a new paradigm for business process support. *Data & Knowledge Engineering, 53(2), pp.129-162.*

Van Der Aalst, W.M., Pesic, M. and Schonenberg, H. (2009) Declarative workflows: Balancing between flexibility and support. *Computer Science-Research and Development, 23(2), pp.99-113.*

Van der Aalst, W.M.P. (2011) Process Mining: Discovery, Conformance and Enhancement of Business Processes. *1st ed.;Springer: New York, NY, USA.*

Van der Aalst, W.M. (2013) Business process management: a comprehensive survey. *ISRN Software Engineering.*

Vasilecas, O., Kalibatiene, D. and Lavbič, D. (2016) Rule-and context-based dynamic business process modelling and simulation. *Journal of Systems and Software, 122, pp.1-15.*

Verbert, K. et al. (2011) Dataset-driven research for improving recommender systems for learning. *in the 1st ACM International Conference on Learning Analytics and Knowledge, 10-13 Oct 2011, Hongkong, pp.44-53.*

Voorhoeve, M. and Van der Aalst, W. (1997) Ad-hoc workflow: problems and solutions. In *Database and Expert Systems Applications. 8th International Conference, DEXA'97. Proceedings* (pp. 36-40). IEEE.

Waller, P., Morris, R.M., Simpson, D. and Hazell, R. (2009) *Understanding the Formulation and Development of Government Policy in the context of FOI*. Stationery Office.

Warf, B. (2018) Packet Switching. The SAGE encyclopedia of the Internet. Publisher: *Sage Reference. ISBN: 1-4739-6035-5, 978-1-4739-6035-0.*

Weber, I., Xu, X., Riveret, R., Governatori, G., Ponomarev, A. and Mendling, J. (2016) Untrusted business process monitoring and execution using blockchain. *In International Conference on Business Process Management (pp. 329-347). Springer, Cham.*

Wegrich, K. and Jann, W. (2006) Theories of the policy cycle. In *Handbook of public policy analysis* (pp. 69-88). Routledge.

Weimer. D. L and Vining. A. R. (2004). Policy analysis : concepts and practice. *4th Edition. Upper Saddle River, NJ ; [Harlow] : Pearson Prentice Hall, 2004.*

Weske, M. (2001) Formal foundation and conceptual design of dynamic adaptations in a workflow management system. *In Proceedings of the 34th annual Hawaii international conference on system sciences (pp. 10-pp). IEEE.*

World Health Organization (2005) Health service planning and policy making: a toolkit for nurses and midwives. Manila: WHO Regional Office for the Western Pacific.

Wu, X., Ramesh, M., Howlett, M., Fritzen, S. (2017). The public policy primer: managing the policy process. *London: Routledge, https://doi.org/10.4324/9781315624754*

Wylot, M., Cudre-Mauroux, P. and Groth, P. (2015) Executing provenance-enabled queries over web data. *In Proceedings of the 24th International Conference on World Wide Web (pp. 1275-1285). International World Wide Web Conferences Steering Committee.*

Xiao, Z., Cao, D., You, C. and Mei, H. (2011) Towards a constraint-based framework for dynamic business process adaptation. *In 2011 IEEE International Conference on Services Computing (pp. 685-692). IEEE.*

Yin, R. (2009) Case Study Research: Design and Methods. *London, Sage Publications.*

Yin, R.K. (2003) Case Study Research Design and Methods. *Sage Publications, Inc., Thousand Oaks, CA.*

Zeng et al. (2011). A Method to Build and Analyze Scientific Workflows from Provenance Through Process Mining. *In Third USENIX Workshop on the Theory and Practice of Provenance (TaPP'11), pages 1-5. USENIX, Berkeley, CA.*

Zensen, A. and Küster, J.M. (2019) A Java-Based Framework for Case Management Applications. *In International Conference on Business Process Management (pp. 107-124). Springer, Cham.*

Zhao, X. and Liu, C. (2007) Version management in the business process change context. *In International Conference on Business Process Management (pp. 198-213). Springer, Berlin, Heidelberg.*