DATA SCIENCE FOR A CITY - SOUTH AFRICAN LOCAL GOVERNMENT

Delyno du Toit

2018-04-28

Contents

Chapter 1

Title page

Declaration English - Abstract (max. 500 words) Afrikaans - Opsomming (max. 500 words) Acknowledgements Table of Contents List of Figures List of Tables Main Body – e.g. Chapter 1, Chapter 2, etc. Reference List Addenda (e.g. Addendum A, Addendum B, etc.)

Chapter 2

INTRODUCTION

2.1 Background

The public sector, more than the private sector, is confined by laws, bylaws, policies, and a risk-averse mindset. Layers of bureaucracy, inflexible rule applications, political agendas, multiple agency involvement curb progressive breakthroughs. Financial stresses are adding to local governments restrictions making them even less responsive. All the constraints limit local government to focus on outpus rather than the bigger picture and solutions. Officials get measured on how many houses are build, how quickly are potholes filled, how much lenghts of road is build rather than actual reductions in homelessness, reduction in potholes, efficient public transportation. Local government is protected against any abuse of discretion through an orientation toward compliance instead of the impact on society. Complexity and rule-driven accountability affect the way officials operates and regulate. Officials who've been with government for long periods, have slumped into a culture of dealing with problems by applying older bureaucratic structures to enforce compliance, resulting in inefficient utilisation of resources.

Yet local government has the means to completely reverse this trend toward despair. One of the opportunities presenting itself is in the form of data modernisation. With the digital tools available today, officials can revolutionize local government, making it more responsive, transparent, and cost-effective than it has ever been.

Local government is structured into compartmentalized units, with each having a unique mandate. The South African city used for this dissertation consist of 11 directorates. A directorate can be thought of as a company within the amalgamation of the city:

1. Area-Based Service Delivery
2. Assets and Facilities Management
3. Corporate Services
4. Directorate of the Mayor
5. Energy
6. Finance
7. Informal Settlements, Water and Waste Services
8. Office of the City Manager
9. Safety and Security
10. Social Services
11. Transport & Urban Development Authority

Each of the directorates have a specific mandate that they need to realize. This leads to directorates working in isolation and duplicating efforts and investments when managing their area. Although each directorate is unique, they must all work together towards a common purpose of providing a safe-, caring-, inclusive-,

opportunity- and well-run city . Big strides have been made to create a transversal approach to solving problems and to break down the silo way of working.

Daily, decisions get taken that have an impact on the city's operations, strategies, urban plans and finances. The City leadership, is dependent on data to inform their decision-making. The problem is that the data used to inform decision making resides on multiple systems and with multiple people. The systems don't talk to each other and to do performance reporting is cumbersome, let alone analyses. Accessing the data is resource intensive and time consuming. Once the data is collated more work need to be performed on it before presenting it for decision making. The data presented is open for interpretation and limited or non-existent in analysis to provide insights to improve performance and get the maximum effect with limited resources. Due to the manual interventions to synthesis the data errors creep in and more concerning different figures are being reported on for the same information.

When it comes to decision-making the city systems is restricted to reporting on the pass and by function, e.g. service requests, revenue, etc. The main system used for this is the ERP system. The reports are reflective and present data in a static format, i.e. not visual or interactive or highlighting exceptions, trends or providing insight from the data. Limited analysis is done on the data on a functional level, and on a transversal level even less. The non-transactional data is used haphazardly during the decision-making process, i.e. survey's, images, videos, sensors, external data from national and provincial government and private sector, etc.

Within the City the term "data-driven decision-making" means reporting key performance metrics based on historical data and using these metrics to support and justify business decisions. This is a good start but data volume, variety and velocity is ever-increasing. To capitalize on opportunities that can be identified, data and analytics should take on a more active and dynamic role in powering the activities of the entire organization, not just reflecting where it's been.

Given all the data there is is still an omission of data-driven decisions. Decisions should no longer be left to gut instinct. Instead, decisions and actions should be based in facts, and those facts also fuel algorithms that predict optimal outcomes. Although it's taken a while to take root, leading enterprises are finally embracing this perspective. The city's IT systems main purpose is to automate processes. Data is stored, then analyzed, often as an afterthought, to assess what had already happened. That passive approach need to give way to a more proactive, engaged model, where systems are architected and built around analytics.

The box below provide a good summary of the bottlenecks faced within the city.

```
In New York City, Michael Bloomberg took office as mayor (January 1, 2002 - December 31, 2013)
after long years of experience in the use of data, and he created a metrics-driven mayoralty.
Agencies agreed to cooperate to set up his proposed data analytics center and other interagency
data initiatives. Yet almost all of them soon asserted legal, technical, and operational obstacles
to full participation. Budget experts also pushed back, worried about costs. Lawyers cited vast
numbers of rules (most from the federal government) that prohibited sharing of data.  Within each
city agency, its chief information officer would explain how only he or she could manage the
complex legacy databases of that unit. Despite his mandate, his commitment to data, and a raft
of first-rate appointees, Bloomberg would not have succeeded in making New York City a leader in
data-driven government had he not pushed hard from the top for change. The lesson here is a bit
paradoxical: How can leaders at the top of a hierarchy create the conditions that will replace
that hierarchy with a far more open and fluid system?
```

(Stephen Goldsmith, 2014) The Responsive City

Unleashing the power of data and analytics will bring about the disintegration of the age of bureaucracy, allowing government to move from a compliance model to a problem-solving one that truly values the intelligence and dedication of its employees and the imagination and civic spirit of its citizens.

Data Science

City Dashboards

Cities Current Data Science Environment

2.2 Problem Statement

How can a South African city use the vast amount of data in their posession to make better decsions by applying data science. The dissertation will address this question by applying data science techniques to three use cases.

2.3 Definitions, assumptions, limitations

2.4 Brief Chapter overviews

The dissertation consist of five chapters.

- In Chapter 1, a compelling case is made for the application of data science within local government.
- Chapter 2 serves as the foundation on which the study is built and as a basis for discussing results and interpretations. This chapter situates the study in the context of previous research and scholarly material pertaining to data science within local government or government, presents a critical synthesis of empirical literature according to relevant themes or variables, justifies how the study addresses a gap or problem in the literature, and outlines the theoretical or conceptual framework of the study.
- Chapter 3 provides a rationale for the methodology followed, describes the research setting and sample, and describes data collection and analysis methods. The chapter provides a detailed description of all aspects of the design and procedures of the study.
- Chapter 4 present a factual reporting of the study results is presented. The chapter offers the researcher an opportunity to reflect on the study's findings, and the practical and theoretical implications thereof.
- Chapter 5 presents a set of concluding statements and recommendations. "Knowing what I now know, what conclusion can I draw?" This may include implications for practice as well as implications for future research. Findings are integrated with the theory employed in the first chapter and the body of knowledge presented in the second chapter. The chapter ends with a cogent conclusion summarizing the importance of the study findings.

Chapter 3

LITERATURE REVIEW

3.1 Introduction

3.2 Broad context theory base

3.3 Detailed (works organised by topic or idea)

3.4 Conclusion

Chapter 4

METHOD

4.1 Introduction

4.2 Research Design

4.3 Research Instruments

4.4 Data

4.5 Analyses

4.6 Limitations

4.7 Ethics

4.8 Conclusion

Chapter 5

FINDINGS AND ANALYSES

5.1 Use Case 1: Predictive Maintenance of a city's Electricity Assets

5.1.1 introduction

5.1.2 sections

5.1.3 sub conclusions

5.2 Use Case 2: Image Classification of Informal Dwellings

5.2.1 introduction

5.2.2 sections

5.2.3 sub conclusions

5.3 Use Case 3: Supervised- and Unsupervised Learning applied to a city's Service Notification

5.3.1 introduction

5.3.2 sections

5.3.3 sub conclusions

Chapter 6

CONCLUSION

6.1 Summary of Findings

6.2 Conclusions

6.3 Summary of Contributions

6.4 Future Research

Bibliography

Stephen Goldsmith, S. C. (2014). *The Responsive City, Engaging Communities Through Data-Smart Governance*. John Wiley and Sons Inc, Published by Jossey-Bass, One Montgomery Street, Suite 1200, San Francisco, CA 94104-4594.