Articles, special issue: ODSCSD

Mapping an emerging Open Data ecosystem

Michelle McLeod
The University of the West Indies, Jamaica
Corresponding author
michelle.mcleod@uwimona.edu.jm

Maurice McNaughton
The University of the West Indies, Jamaica
maurice.mcnaughton@uwimona.edu.jm

The purpose of this paper is to explore a methodological approach to understand an emergent open data ecosystem in developing countries and specifically in the tourism sector. The conceptual and methodological bases are explored using Actor Network Theory (ANT) and Social Network Analysis (SNA) to understand their application to the open data phenomenon. Thirteen tourism industry officials in the public sector of five Caribbean countries were interviewed using a research instrument derived from the Open Data Research Network (ODRN) Common Assessment Framework. The findings reveal an inter-connected emergent open data ecosystem across five Caribbean countries.

Introduction

The open data (OD) phenomenon is still a relatively new political, social and economic movement. As it evolves from its initial pre-occupation with publishing open government data portals, key stakeholders such as governments, citizens and multilateral agencies, are becoming increasingly interested in more analytic methods for evaluating the value potential and impact of open data initiatives.

McLeod, M., McNaughton, M. (2016). Mapping an emerging open data ecosystem. *The Journal of Community Informatics*, 12(2), (Special issue on Open Data for Social Change and Sustainable Development), 26-46.

Date submitted: 2015-08-16. Date accepted: 2016-05-09.

In the Caribbean, governments have been slow to embrace the open data movement[1]. As a region of small-island developing economies, open data compete with a range of socio-economic policy demands for scarce resources and political attention. Within this context, the assessment of the readiness for open data initiatives and its potential economic impact requires greater specificity, more targeted focus and a compelling narrative.

Various frameworks and mechanisms have emerged for analysing prospective open data contexts. The World Bank's Open Data Readiness Assessment (ODRA)[2] methodology provides a toolkit to support an economic and action-oriented assessment of the readiness to evaluate, design and implement an open data initiative. The Open Data in Developing Countries (ODDC) project developed a stakeholder mapping method using Visual Understanding Environment (VUE)[3] software (Open Data in Developing Countries, 2013) that facilitates the identification of issues, actors, important information and data sources and their inter-connections (see Appendix I).

The Open Data Census[4] and the Open Data Barometer[5] have become well-established global indices on open government data. The Open Data Census ranks countries and cities on the supply and degree of openness of common datasets, while the Open Data Barometer provides a more comprehensive methodology for ranking countries and regions on multiple dimensions of open data readiness, implementation and impact. These are among the more popular, established assessment mechanisms, but are predominantly descriptive in nature, supply-side oriented with the unit of analysis emphasising the individual country, city or agency. Subsequent efforts have been made to consolidate common methods and frameworks for the study of open data with the development of the Common Assessment Framework for Open Data[6].

The Caribbean Open Institute (COI) recently conducted a series of sector studies (as part of a larger research initiative 'Harnessing Open Data to Achieve Development Results in Latin America and the Caribbean'[7]) that seek to determine the institutional attributes, data sharing practices, value-opportunities and potential enablers/constraints of an emergent open data ecosystem in various sectors. An open data ecosystem can be characterised as a network of inter-connected, interdependent actors, data sources and data exchange relationships, wherein the network attributes can enable or constrain the impact of data within that ecosystem. The use of the ecosystem metaphor to describe open data has been explained (Harrison, Pardo & Cook, 2012), but there is still a need for effective analytic mechanisms to explicate the operation of such an open data ecosystem.

---

1 As at May 2015, only two Caribbean countries had signed up to the Open Government Partnership (Trinidad and Tobago, and the Dominican Republic). At the time of writing, there were no official open government data portals in the Caribbean.

2 http://opendatatoolkit.worldbank.org/en/odra.html

3 http://vue.tufts.edu

4 http://global.census.okfn.org

5 http://opendatabarometer.org/

6 http://opendataresearch.org/content/2014/646/towards-common-methods-assessing-open-data-workshop-report

7 http://idatosabiertos.org/wp-content/uploads/2014/04/OD4DProposaversionfinal11.pdf

Specific to the tourism sector, semi-structured interviews were conducted and thirteen institutional stakeholders in five Caribbean countries participated. Existing assessment methods were applied to the data collected, but lacked sufficient analytic precision to answer key research questions. Hence the study underscored the need and presented an opportunity to explore other approaches to analysing and understanding an emergent open data ecosystem in a developing country context.

This paper proposes a methodological approach for understanding an open data ecosystem as an actor-network within which data flows are facilitated through an enactment of actors. Knowledge is built through the receipt of data and information which can be shared through a network (McLeod & Vaughan, 2015). Data activity can be explained through a network mechanism of relationships (McNaughton, McLeod & Boxill, 2014) and thereby an open data ecosystem can be characterised as an actor-network where network relationships represent data supplied freely for use and re-use, and utilised for various activities, not in the least policy-making. Specifically, the application of Actor Network Theory and Social Network Analysis as conceptual and methodological approaches to examine an emergent tourism open data ecosystem was explored.

Theoretical background

This paper explores conceptual and methodological aspects of an open data ecosystem. Open data are based on the availability of machine readable data with no restrictions on use and re-use. In addition, the provision of open data has been largely the responsibility of public entities that are the repositories of collected datasets. An ecosystem can be viewed as a set of inter-connected components that are built and function around the purpose of that ecosystem. Open data ecosystems are formed through the provision of open data but the characteristics and operation of that ecosystem still need to be revealed. The ecosystem metaphor has found useful application in the open data field (Davies, 2011; Harrison et al., 2012; Lee, 2014; Van Schalkwyk, Willmers & McNaughton, 2015). Davies (2011) combines infrastructure and ecosystem metaphors to explain the impact of open data initiatives. Lee (2014) proposes eleven elements of an open data ecosystem including data audit, privacy, dataset selection, data access, licensing, high-quality data, data discovery, public agency support, data user engagement, economic reuse and evaluation. While these elements can be purposefully driven towards achieving economic, social or environmental goals, the required configuration for achieving a successful ecosystem architecture can guide interventions to improve the operation of an open data ecosystem. It is with this view that conceptual and methodological approaches have been identified in this paper.

Data sharing is enacted by a system of actors. Actors can be framed using the lens of Actor Network Theory (ANT) (Latour, 2005). ANT provides an ontological position for this work and proposes that enactment occurs through heterogeneous relationships of human and non-human entities, including for the purpose of a data actor-network, not exhaustively as suggested by Law (2007) human beings, machines, organisations and ideas. ANT proposes that humans and non-humans are actors alike (Jóhannesson, Van der Duim & Ren, 2012). An open data ecosystem actor-network comprises an assemblage of these actors and once actors can be identified, and their relationships established, an arrangement of actors can be revealed. An actor-network arrangement is based on key ANT concepts of assembling,

enacting and ordering (see McLeod, 2013). Within an actor-network arrangement, translation of an actor-network can be understood (Ren, Jóhannesson, & Van der Duim, 2012). Translations can be viewed as the processes within the network as 'actants', both human and non-human, become agents through their inter-relationships. Based on the works of Latour, Callon and Law, it is suggested that 'humans often stand powerless confronted with strong actor-networks made up of technologies, documents, habits, discourse and schemes' (Ren, Jóhannesson, & Van der Duim 2012, p. 16) and this speaks to the potential agency of non-human actors.

There is a need to map ecosystems, and also understand the composition of any gaps within the ecosystem in order to develop policies to enable ecosystem dynamics. Whilst a network can be cross-sectional based on a snapshot of actors and ties within a time period, an ecosystem is a broader concept that suggests inter-dependence of actors that emerge and leave the system based on certain interactions. The idea of an open data actor-network is largely descriptive; however this representation is enhanced with the application of Social Network Analysis (SNA), a theory and methodology that maps actors and their relationships to reveal a pattern that can influence the outcomes for those actors (Freeman, 2004). The tradition of SNA is based in sociology with mathematical foundations to represent networks in graph-theoretic terms and in that regard SNA is conceptually dissimilar from ANT. Modern day SNA involves the understanding of the antecedents and consequences of social relationships (Borgatti & Halgin, 2011). According to Borgatti and Halgin (2011) the 'theory of networks' is the antecedents of networks in terms of the selection of and opportunity for certain types of relationships and 'network theory' suggests that there are processes, which result in the network pattern. In its simplest form a network begins with two nodes and a tie, called a dyadic relationship. The network is formed as nodes are inter-related by certain associations, cognitions, events and activities (Borgatti, Everett & Johnson, 2013). The interconnections form a network pattern and SNA provides a mechanism that can measure and reveal the consequences of that network pattern. Although one shortcoming of SNA is its inability to show translations that ANT suggests occur, its representation of an existing network of relationships can contribute to understanding the key influencers and the processes within the network. Certainly, open data ecosystem mapping using SNA is a starting point for understanding the ordering of a network.

ANT is based on the premise that human and non-human actors are enacted within an interdependent system that evolves towards certain outcomes. One of the major debates surrounding ANT is whether or not it is a 'theory' in that there is a lack of a coherent framework (Mol, 2010). It is rather the case that ANT brings a new perspective on the ordering of 'things' both human and non-human. This ordering can be viewed as a 'network tracing activity' and therefore actor-network theory is based on understanding the network ordering rather than its structure (Latour, 1996). SNA focuses on the structure or pattern of the actors in the network and in that regard there have been several challenges with application of this approach. Borgatti, Brass and Halgin (2014) have summarised the main confusions, criticisms and controversies surrounding SNA as: 'all description, no theory' (p. 3), 'just methods and measures' (p. 5), 'all structure, no content' (p. 7), 'all structure, no agency' (p. 10), 'all static, no change' (p. 13), and 'all networks, no context' (p. 16). They conclude that work is needed to build on the concepts of network content, network change and network context to advance SNA.

The underlying premise in both theories, is the idea of a 'network'. ANT focuses on the workings or enactment upon which actors are recruited to evolve the network and SNA focuses on the structure of the network that results in certain outcomes. While there are distinct differences between the two bodies of knowledge a conjoining of ANT and SNA can add value to understanding the nature of an open data ecosystem, especially one that is nascent or emergent, and hence nowhere near equilibrium, as is the proposed case of the Caribbean tourism open data ecosystem. An emergent property within a network context is conceptually matched with ANT and therefore ANT provides a conceptual foundation to understand an open data ecosystem, as ANT concepts are aligned with the process or dynamic features of an ecosystem. The enactment of actors is downplayed by utilising a snap-shot network model approach as with the use of SNA. Nonetheless, the value that network structure adds is the capturing of the influences that exist within an ecosystem. Influential actors and their placement within the network are the potential enablers or constrainers of the emerging ecosystem.

Conjoining ANT and SNA can contribute to understanding the characteristics and operation of an open data ecosystem, which can also be viewed as a socio-technical system (STS). Socio-technical systems (STSs) are comprised of social and technological elements within a system. Kling, McKim and King (2003) suggest that Socio-Technical Interaction Networks (STINs) provide a basis for which there is interaction between technology that impacts human behaviour and human behaviour influences technology. In that regard Kling et al. (2003) diverge from utilising ANT both epistemologically and methodologically and take a more conservative approach regarding non-human agency and propose a STIN model that allows for proactive determination of the actors and their interactions to model a socio-technical network. Thus, a STIN approach suggests that a model of actors and interactions can be pre-determined to achieve certain goals, but this approach limits an assessment of existing actors and interactions. A broader approach should be taken to understand STSs such as an open data ecosystem and herein is the value of ANT. ANT has been applied to conceptualising STSs (Heeks & Stanforth, 2007; Tinati, Halford, Carr & Pope, 2011; Tinati, Carr, Halford & Pope, 2012) and SNA has been utilised for mapping STSs to understand information flows in open data development (Zappia, 2011). The equality of human and non-human actors in terms of agency within ANT can add value to understanding the impacts of open data ecosystems (Kuk & Davies, 2011). Kuk and Davies (2011) made a distinction between human agency and material agency (material agency combined with artifacts has been referred to as materiality) to explain an assemblage of open data complementarities. This materiality can influence the outcome of an open data ecosystem and therefore ANT can be a conceptual lens for understanding open data ecosystems.

Methodology

Research location

The research objectives of this paper is to outline conceptual and methodological approaches for understanding an emergent Caribbean tourism open data ecosystem, identify influential actors and inhibitors, and suggest opportunities for improving the ecosystem. Tourism is an important economic and social activity for the Caribbean region. Five countries in the region were selected for this research study: Antigua and Barbuda, Barbados, the Dominican

Republic, Jamaica, and Trinidad and Tobago. In 2014, tourist arrivals in the Caribbean region peaked at an estimated 26.3 million tourists and visitor expenditure reached USD29.5 billion (CTO, 2015). In terms of the number of tourist arrivals, each country is at a different stage. The Dominican Republic continues to be a major performer with 5.1 million tourist arrivals and Jamaica's tourist arrivals surpassed 2 million tourists in 2014 (CTO, 2015). In 2014, the other three countries, Antigua and Barbuda, Barbados and Trinidad and Tobago combined accounted for about 50% of the number of tourists that visit Jamaica and about 25% of the number of tourists that visit the Dominican Republic.

The tourism sector, like any other industrial sector is data driven. Tourism development requires informed policies that can propel tourism sector growth and performance. Knowledge-based tourism policy-making becomes important to achieve effective tourism development (McLeod & McNaughton, 2015). In the Caribbean tourism sector open data about tourism arrivals, tourism products and tourism service providers are essential for making strategic decisions to advance the tourism sector. The private sector also plays an important role as stakeholders of tourism products and services. An open data platform can be readily utilised to inform tourism businesses and civil society to create opportunities for investment and entrepreneurship in the tourism sector. The Dominican Republic, Jamaica, Antigua and Barbuda, and Barbados have tourism industries that make substantial contributions to those economies. For example, tourism accounts for an estimated 20% of GDP in Jamaica to an estimated 70% GDP in Antigua and Barbuda.

Data collection

The Common Assessment Framework for Open Data (Open Data Research Network, 2014) was adapted for purposes of a sector-specific study of an emergent open data ecosystem, and used as the basis for developing a research instrument for data collection. The data were collected using the common assessment framework for open data. This framework was not designed to conduct network analysis; however, through its inclusion of questions relating to sources and users of tourism data, the instrument allowed the capture of tourism data relationships of suppliers and users within the five countries. This paper reports on results obtained from the data provenance/profile, data use and data impact sections of the ODRN framework. The main tourism datasets including tourism arrivals, tourism assets and tourism service providers and any other tourism datasets were assessed using the framework. In terms of data use, information regarding the primary users as well as any potential users who should and could add value to the sector were obtained. The potential impacts were categorised based on social, environmental, political/governance, and economic/commercial dimensions and identified using the Caribbean Tourism Guiding Policy Principles for Sustainability (CTGPPS) (CTO, 2014). The policy issues used in the CTGPPS research study were: capacity management, marketing, transportation, environment, economic linkages, and health, safety and security (CTO, 2014).

The government sources of tourism data were selected for the research study. The government was selected as the official source of tourism data through which government policy can transform these data into open data. The participants included the main public sector tourism organisations such as the Ministry of Tourism and the tourism authorities in five Caribbean countries. Skype and telephone interviews were conducted and notes were taken during the interview. Thirteen questionnaires were obtained from an initial 16

invitations, an 81.3% response rate. Two key assumptions included: (i) that at least one respondent is able to represent each tourism government institution in relation to the tourism data landscape in their country; and (ii) that the three main datasets – tourism arrivals, tourism products and tourism service providers – exist in all five countries.

Data analysis

Some assessment was conducted using existing open data assessment approaches. Appendix I shows the stakeholder mapping approach, which depicts the various institutional stakeholders in the tourism sector and their connectedness to the key policy issues and supporting datasets. Based on the stakeholder mapping exercise at Appendix I, the tourism datasets were interconnected within the tourism sector data system of the countries; however there is need to refine the relative importance of these datasets. In addition, Appendix II depicts the relative ranking of the five countries in terms of the openness of Caribbean datasets including the tourism datasets using the Open Data Census. Both techniques provide a useful visual representation of aspects of the tourism data ecosystem, but are limited in terms of any further analytic insight.

A framework containing attribute categories and actors (Table 1) was developed to assist with defining and categorising the relationships among the actors. Attribute categorisation is particularly important since the same actor such as a government agency can be placed in more than one category such as source, owner and custodian of data (Figure 1). An attribute category was created and defined as 'open' (circled in Figure 1) and this category includes the public, website, no restriction, no cost and Excel format. The 'open' attribute category is not a data attribute but a combined construct of data demand and supply elements.

The actors are not an exhaustive list but are indicative of the types of responses from the participants of this research study. Data providers and users were identified using a shortened version of their name such as Immigration (Immigra), Statistical Office (StatOff) or Central Bank (CenBank). A naming approach instead of a coding approach was taken to assist with understanding the diagram. The tourism datasets, called 'data type', include tourism arrivals, tourism assets, tourism service providers, exit survey, destination studies and entertainers. Tourism data content can influence the outcomes of the system in terms of its impact on tourism policy issues and therefore it was important to include policy issues within a tourism data ecosystem to understand the more important tourism policy issues in the Caribbean region. Network relational data were derived from the questionnaires, placed into attribute categories and entered into an Excel spreadsheet for analysis using SNA software. The relational data were analysed using UCINET 6 software (Borgatti, Everett & Freeman, 2002) and NetDraw (Borgatti, 2002). UCINET 6 software has been popularised for SNA and can be obtained from Analytic Technologies[8]. The relational data were entered using a node list format with the country being the focal node and relationships were assigned to the focal node based on interview responses. For example, a linked list for country, data source, data owner, and data custodian may include: Trinidad & Tobago Immigration Tourism Authority Immigration Tourism Authority Ministry of Tourism (see Table 1).

8 Analytic Technologies - http://www.analytictech.com/

Specifically, through use of SNA, we explore the following questions about this emergent tourism open data ecosystem. Questions 1 to 3 relate specifically to the domain under study, while Questions 4 and 5 examine the efficacy of SNA techniques for evaluating an open data ecosystem.

1) Who are the influential actors in the ecosystem?
2) What are major inhibitors: (i.e. structural attributes or patterns) resulting in data exchange inefficiency that could be alleviated by using open data?
3) What opportunities exist for a digital commons-based approach to sharing data resources and approaches across countries?
4) Does the network analysis enable us to infer the degree of data openness consistent with conventional assessment methods?
5) Are there visible, observable network patterns or structural attributes that would signify the potential of a healthy open data ecosystem?

SNA was utilised to address these questions as once the network had been mapped, certain measures could be applied to understand the importance of network actors and consequences of the network pattern. Centrality is a key measure of network position and is an indicator of an actor's influence. Those actors that are more central are in an advantageous position to control network resources or the removal of the most central node can fragment the network (Borgatti, Everett & Johnson, 2013). The overall network centralisation is the extent to which actors revolve around the most central actor (Everett & Borgatti, 2005) and network centralisation supports convergence within the network as actors with similar relationships hold similar network positions.

Table 1: Attribute categories and actors framework

Data attribute categories	**Actors**	**Other attribute categories**	**Actors**
Data source	Immigration Ministry of Tourism Tourism Authority Hotel Association Central Bank	**Country**	Antigua & Barbuda Barbados Dominican Republic Jamaica Trinidad & Tobago
Data owner	Immigration Ministry of Tourism Tourism Authority Central Bank	**Primary user**	Ministry of Tourism Tourism Authority Students Researchers Planning Authority Government Agencies Businesses Public
Data custodian	Immigration Ministry of Tourism Tourism Authority Central Bank	**Should be user**	Public Researchers Government Agencies Competing Destinations Businesses Media

Data type	Tourism Arrivals Tourism Assets Tourism Service Providers Tourism Exit Survey Tourism Entertainers Tourism Destination Studies	**Could be user**	Researchers Hotel Associations Businesses Application Developers Overseas Offices Government Agencies
Data share	Ministries Teachers Caribbean Tourism Organisation (CTO) Overseas Offices	**Open**	Public Website NoRestrict NoCost Excel
Data dissemination	E-mail Regular Mail Telephone Website Reports Special Request Press Release Excel PDF	**Public sector**	Ministry of Tourism Tourism Authority Central Bank Tax Office Ministry of Trade
Data restriction	Yes Restriction No Restriction	**Private sector**	Hotel and Tourism Association Businesses Airlines
Data cost	No Cost Free	**Policy issue**	Marketing Economic Linkages Capacity Management Transparency Environment Transport Health, Safety & Security

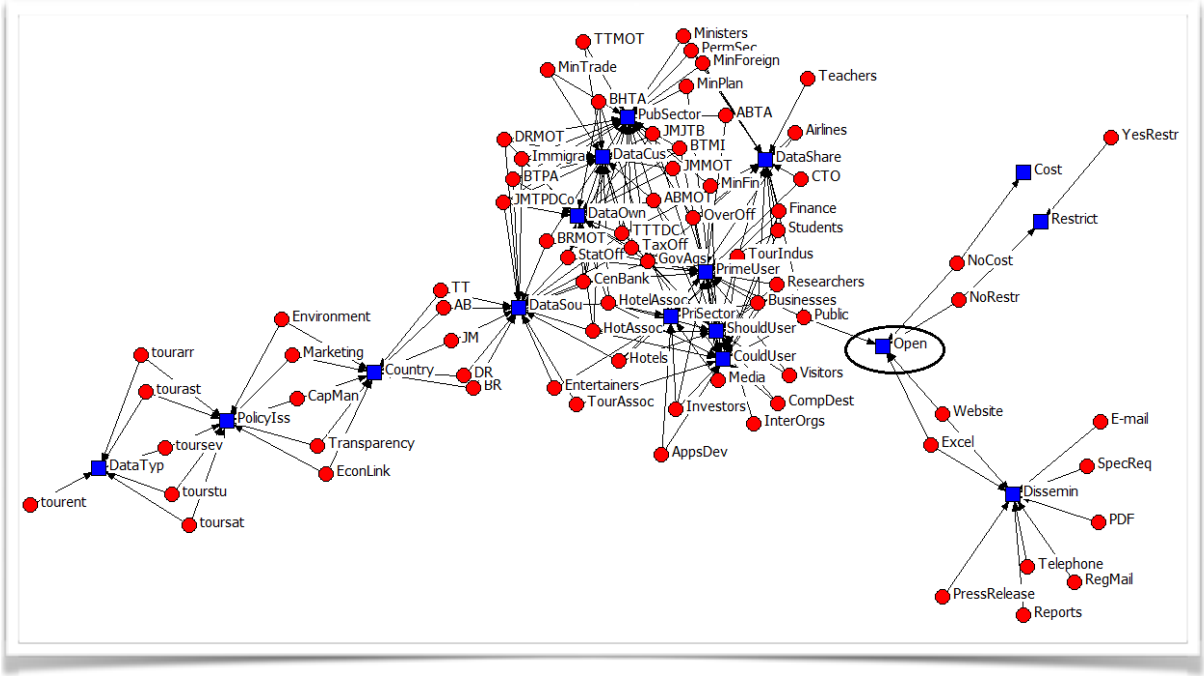


Figure 1 Attribute categories and actors

Findings

The findings section outlines influential actors, inhibitors and data sharing approaches, openness assessment and a healthy data ecosystem with opportunities for open data. Figure 2 shows the outcome of the Caribbean tourism data ecosystem network mapping.

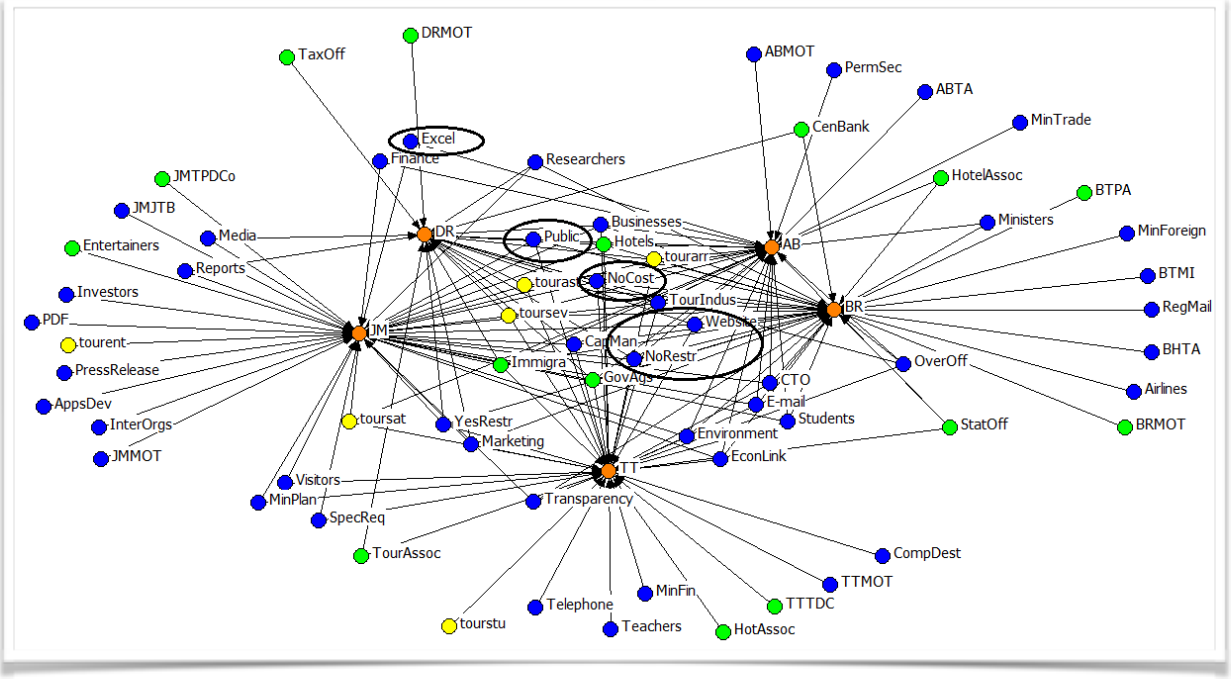


Figure 2: An emergent Caribbean tourism data ecosystem

Influential actors

Influential actors are central actors who can enact either positively or negatively and influence network effectiveness and outcomes. Centrality is measured as the degree (in terms of the number of relationships) to which actors were enacted within the network. The overall network centralisation in-degree is 1.346% and this is the extent to which all other actors revolve around the most in-degree central actor. A 'Public' actor has a degree centrality of 16 and this actor has the highest in-degree centrality value that has been normalised for comparative purposes (in this network of 72 actors, the normalised in-degree centrality is 16/71 times 100 or 22.54). The actors with high normalised in-degree centrality values are influential to the ecosystem's operation as these actors are at the core of the network and the top five include the public, the tourism industry (private sector), the statistical offices in three Caribbean countries (Antigua and Barbuda, Barbados, and Trinidad and Tobago), the various immigration departments and the tourism assets dataset (Table 2). The most prominent policy issues with higher in-degree centrality values are capacity management, marketing and the environment. The disseminative capacity can be determined by no cost, website, no restriction and e-mail as these actors (nodes) have higher in-degree centrality values than other disseminative actors (nodes) such as reports.

Table 2: Centrality values (Freeman degree centrality)

Actors	InDegree	NrmInDeg	Actors	InDegree	NrmInDeg
Public	16	22.54	GovAgs	10	14.08
TourIndus	15	21.13	Website	9	12.68
StatOff	15	21.13	NoRestr	9	12.68
Immigra	14	19.72	CapMan	9	12.68
tourast	13	18.31	CenBank	9	12.68
tourarr	12	16.90	Marketing	8	11.27
DRMOT	12	16.90	OverOff	7	9.86
TTTDC	10	14.08	ABMOT	7	9.86
toursev	10	14.08	E-mail	7	9.86
NoCost	10	14.08	Environment	7	9.86
Businesses	10	14.08	TaxOff	7	9.86

Inhibitors and data sharing approaches

Data sharing relationships are enacted through a network mechanism. Inhibitors can constrain data dissemination. As a result, there is need for both a network mechanism facilitated by disseminating actors such as a website for there to be effective data diffusion. The composition of the network in terms of the types of actors and disseminating actors can assist with clarifying any inhibitors. In this case several observations are made of the ordering of network positions and these identify potential inhibitors to data sharing activities:

- User and dissemination nodes are in proximity: from visual evidence the 'Students' node is in proximity to the 'E-mail' node, and the 'Media' node is in proximity to the 'Reports' node. A breakdown of the network revealed that there are matches between the types of users and dissemination tools among the countries and hence the reason for the network positions of the nodes. For instance, both the Dominican Republic and Jamaica indicated media as users and reports as a dissemination tool, while the other countries did not and hence the network position of the 'Media' actor and the 'Reports' actor.

- User groups are positioned in similar proximity within the network: a group of private sector actors – 'Businesses', 'Hotels' and 'Tourism Industry' – is positioned similarly and this relates to a set of common users within all five countries. There is no evidence however of a common dissemination tool that reaches this user group and therefore it is suggested that this is an opportunity for an open data initiative.

- The public sector actors are key potential enablers or inhibitors based on the strength of ties between these actors and the other actors within the emergent tourism open data ecosystem (see Appendix III). The Ministry of Tourism and various tourism authorities in the countries, particularly the ABMOT (Antigua and Barbuda Ministry of Tourism),

DRMOT (Dominican Republic Ministry of Tourism) and TTTDC (Trinidad and Tobago Tourism Development Company) are important public sector actors in the ecosystem. The dissemination policies developed within these organisations can enable or inhibit data sharing within the emergent tourism open data ecosystem.

Openness assessment and a healthy data ecosystem with opportunities for open data

The tourism data network comprises 72 actors in five Caribbean countries (Figures 1 and 2). In Figure 1, the attribute categories reveal dense relationships between the data providers (sources, owners and custodians) and users (primary, should be and could be users). In Figure 2, the overall network topology is circular in pattern with several pendants pulling away Jamaica and Trinidad and Tobago from the other three countries, Antigua and Barbuda, Barbados, and the Dominican Republic. The country actors' positions in the emergent tourism open data ecosystem are similar to the geographical location of the countries and therefore the emergent ecosystem has an assemblage that is a form of 'geo-ordering' (a term coined by authors). Also, the network lacks a coordinating actor that can make a direct intervention to reach other actors in one step. A star network allows for this reach dynamic and is therefore the most efficient network topology (Borgatti et al., 2013). In addition, the pendants are opportunities to create convergence within the data ecosystem as those countries without these actors should seek to engage these pendant entities more readily, for example investors and application developers are only enacted in Jamaica and they can also be contributing to open data ecosystem dynamics in the other countries. In addition, openness actors (Figure 2; circled), 'Website', 'Public', 'No Cost', 'No Restrictions' and 'Excel' are differently positioned in the network and this is evidence of a lack of coordination in the network and that the open data ecosystem is an emergent one.

Although, the overall network seems disorganised, the countries that have similar relationships are closer in proximity to each other on the network diagram as the geodesic distance (the shortest path between two nodes) is shorter. For example, Antigua and Barbuda, and Barbados are similar, and Jamaica is farthest. An open data initiative can create similar types of data relationships within this tourism data ecosystem and therefore improve convergence. Similar observations are made with the policy issues. Policy issues that are in proximity relate to the environment and economic linkages, while marketing and capacity management issues are apart. This network positioning is an indication for greater assessment of the data needs for the formulation of policy guidelines to achieve effective policy implementation success within the tourism sector. While tourism sector stakeholders are engaged with providing information and utilising the various tourism datasets through a formal arrangement for the benefit of the tourist destination, the network ordering suggests that the tourism assets and tourism service providers datasets are in closer proximity to capacity management issues and therefore some attention should be given to these datasets (Figure 2). This is an opportunity for an open data initiative that can contribute to the development of data to achieve the required policy goals.

Conclusions

Open data initiatives provide an opportunity for the creation of an open data ecosystem within which several data elements and users interact to achieve specific goals. Although there is a notable limitation with traditional SNA in realising the process or enactment aspect

of ANT, advances with SNA modelling and visualisation can potentially capture ANT concepts but this was not the focus of this research study. By taking an ANT perspective the interrelationships between data supply and user elements converge and provide meaningful insight about the prime influencers of an open data ecosystem. A network approach for understanding open data ecosystems can add value to the field. A network can be viewed as a coordinating mechanism in which ties are bonds that create cooperation between actors or converge and create sameness of action (Borgatti &d Halgin, 2011). An understanding of ordering using ANT and co-ordination using SNA are complementary, both conceptually and methodologically for understanding an open data ecosystem.

The enrollment of actors within an emergent open data ecosystem can effect tourism development as the existence and extent of openness attributes within the ecosystem facilitate the effectiveness of tourism policies (McLeod &McNaughton, 2015). The key actors are the prime users, including the members of the public and the tourism industry. Any open data initiative should review effective dissemination tools for these user groups. In addition, important influencers in terms of tie strength are those agencies in the public sector – the Ministry of Tourism in the countries and the tourism authorities – and therefore these agencies are required to implement and set out the policy framework respectively for a healthy tourism open data ecosystem. The resourcing of these actors becomes important to effect changes in the tourism open data landscape.

The interconnectedness of human and non-human actors can enable or constrain network outcomes. The impact of the data within the system is enabled by its ability to diffuse and the provision of content to meet policy goals. For instance, although the arrivals dataset from the demand-side can be used for marketing purposes, information from the tourism assets and tourist exit survey can also influence marketing strategy as these assess the supply-side of the sector. The interventions needed to transform the tourism system to bring value can be linked to the materiality of the non-human actors with equal emphasis placed on the resourcing and agency of human actors. This perspective is important to understand the impact of data within the ecosystem.

For developing countries, particularly those with limited resource endowments, such as the small-island developing states of the Caribbean, opportunities for open data require a more targeted sector-specific approach that can maximise the benefits and returns from the allocation of scarce financial and human capital resources to such policy initiatives. This research study explores an analytic framework for understanding the characteristics and potential consequences of an emergent open data ecosystem in the tourism sector in order to inform the targeted allocation of resources to stimulate the sector's development.

This study is exploratory and further work will re-examine the structure of the instrumentation used for data collection, to be more specifically designed to support network analysis. There is no doubt that an open data initiative with its emphasis on data sharing will support the generation of knowledge but the suitability of the tools and architectures that can support a healthy open data ecosystem for there to be continued data sharing and utilisation can be revealed through ongoing work through application of ANT and SNA concepts and methodologies. The research design can be improved by collecting data from stakeholders that use tourism data such as the hotel associations and community organisations. The instrument did not allow for naming of specific users but for the categories of users. Such

naming facilities snowballing to capture the breadth of an open data ecosystem as each actor within the ecosystem, whether human, such as a user or non-human, such as a dataset can potentially enact or influence the use of tourism data within the ecosystem.

Acknowledgments

This research study was funded by a grant from the International Research Development Centre, Canada. The authors acknowledge the contributions of Mr Ryan Skeete and Ms Sophia Lashley of the Caribbean Tourism Organization, the support of Mr Julio Orozco of the Association of Caribbean States and Ms Marcia Isa, Vice-Minister of Tourism, the Dominican Republic. Thanks is expressed to Ms Gale Yearwood of the Ministry of Tourism, Barbados for supporting documentation. The administrative support of Mr Carlos Gordon of the Mona School of Business and Management is acknowledged. The authors also express thanks to the thirteen individuals who responded to the survey.

References

Borgatti, S. P. (2002). *NetDraw Network Visualization*. Harvard, MA: Analytic Technologies.

Borgatti, S. P., Brass, D. J. & Halgin, D. S. (2014). Social network research: confusions, criticisms, and controversies. In D.J. Brass, G. Labianca, A. Mehra, D.S. Halgin, & S.P. Borgatti (Eds), *Research in the Sociology of Organizations*, 40. Bradford, UK: Emerald.

Borgatti, S. P. & Halgin, D. (2011). On network theory. *Organization Science*, 22(5), 1168-1181.

Borgatti, S. P., Everett, M. G. & Freeman, L. C. (2002). *UCINET 6 for Windows: Software for Social Network analysis*. Harvard: Analytic Technologies.

Borgatti, S. P., Everett, M. G. & Johnson, J. C. (2013). *Analyzing Social Networks*. London: SAGE.

Caribbean Tourism Organization (2014). *Caribbean sustainable tourism policy framework*. Barbados: Caribbean Tourism Organization.

Caribbean Tourism Organisation (2015). *Caribbean Tourism Review, industry update – end of year 2014*. Barbados: Caribbean Tourism Organization.

Davies, T. (2011). *Open Data: infrastructures and ecosystems*. Unpublished paper, University of Southampton.

Everett, M. G. & Borgatti, S. P. (2005). Extending centrality. In Carrington, Scott & Wasserman (eds.), *Structural Analysis in the Social Sciences, Models and Methods in Social Network Analysis* (pp. 57-76). Cambridge: Cambridge University Press.

Freeman, L. C. (2004). *The Development of Social Network Analysis: A study in the sociology of science*. Vancouver: Empirical Press.

Harrison, T. M., Pardo, T. A., & Cook, M. (2012). Creating Open Government ecosystems: a research and development agenda. *Future Internet* 4, 900-928.

Heeks, R. & Stanforth, C. (2007). Understanding e-Government project trajectories from an Actor-Network perspective. *European Journal of Information Systems*, 16 (2), 165-177. doi: 10.1057/palgrave.ejis.3000676

Jóhannesson, G. T., Van der Duim, R. & Ren, C. (2012). Introduction. In Van der Duim, Ren & Jóhannesson, *Actor-network Theory and Tourism: Ordering, materiality and multiplicity*. Oxon: Routledge.

Kling, R., McKim, G. & King, A. (2003). A bit more to it: scholarly communication forums as socio-technical interaction networks. *Journal of the American Society for Information Science and Technology*, 54 (1), 47-67.

Kuk, G. & Davies, T. (2011). *The roles of agency and artifacts in assembling Open Data complementarities*. In Thirty Second International Conference on Information Systems. China: Shanghai.

Latour, B. (1996). *On actor-network theory: a few clarifications*. Retrieved from http://www.bruno-latour.fr/sites/default/files/P-67%20ACTOR-NETWORK.pdf.

Latour, B. (2005). *Reassembling the Social: An introduction to Actor-Network Theory*. Oxford: Claredon.

Law, J. (2007). *Actor network theory and material semiotics*. Retrieved from http://heterogeneities.net/publications/Law2007ANTandMaterialSemiotics.pdf

Lee, D. (2014). Building an open data ecosystem – An Irish experience. ICEGOV2014, 27-30 October 27-30, Guimaraes, Portugal.

McLeod, M. T. (2013). A review of 'Actor Network Theory and tourism, ordering, materiality and multiplicity'. *Tourism Management*, 37, 48-49.

McLeod, M. & Vaughan, R. (2015). *Knowledge Networks and Tourism*. Oxon: Routledge.

McLeod, M. & McNaughton, M. (2015). Knowledge-based tourism policy formulation, as an application of open data in Caribbean tourism. Conference Proceedings from International Conference on Tourism, 24–27 June 2015, London, United Kingdom, (pp. 205-215).

McNaughton, M. L., McLeod, M. T. & Boxill, I. (2014). Tourism open data in Jamaica: An actor network perspective. Conference Proceedings from 3rd Interdisciplinary Tourism Research Conference, 3–8 June 2014, Istanbul, Turkey (pp. 268-274).

Mol, A. (2010). *Actor-network theory: sensitive terms and enduring tensions*. Kölner Zeitschrift für Soziologie und Sozialpsychologie, 50(1), 253-269.

Open Data in Developing Countries (2013). *Open data stakeholder mapping*. Retrieved from: www.opendataresearch.org/emergingimpacts.

Open Data Research Network (2014). *Common assessment framework for Open Data*. Retrieved from: http://opendataresearch.org/content/2014/646/towards-common-methods-assessing-open-data-workshop-report .

Ren, C., Jóhannesson, G. T. & Van der Duim, R. (2012). How ANT works. In Van der Duim, Ren & Jóhannesson, *Actor-network Theory and Tourism: Ordering, materiality and multiplicity*. Oxon: Routledge.

Tinati, R., Halford, S., Carr, L. & Pope, C. (2011). Conceptualising a web of linked actors. *Web Science*, 3–6.

Tinati, R., Carr, L., Halford, S. & Pope, C. (2012). Exploring the impact of adopting Open Data in the UK Government. In Digital Economy Conference. Aberdeen.

Van Schalkwyk, F., Willmers, M. & McNaughton, M. (2015). Viscous open data: The roles of intermediaries in an open data ecosystem. *Journal of Information Technology for Development*. doi: 10.1080/02681102.2015.1081868

Zappia, Z. (2011). *Participation, power, provenance: Mapping information flows in open data development*. University of Oxford.

Appendix I Caribbean Tourism Stakeholder Maps (five countries)

Key:

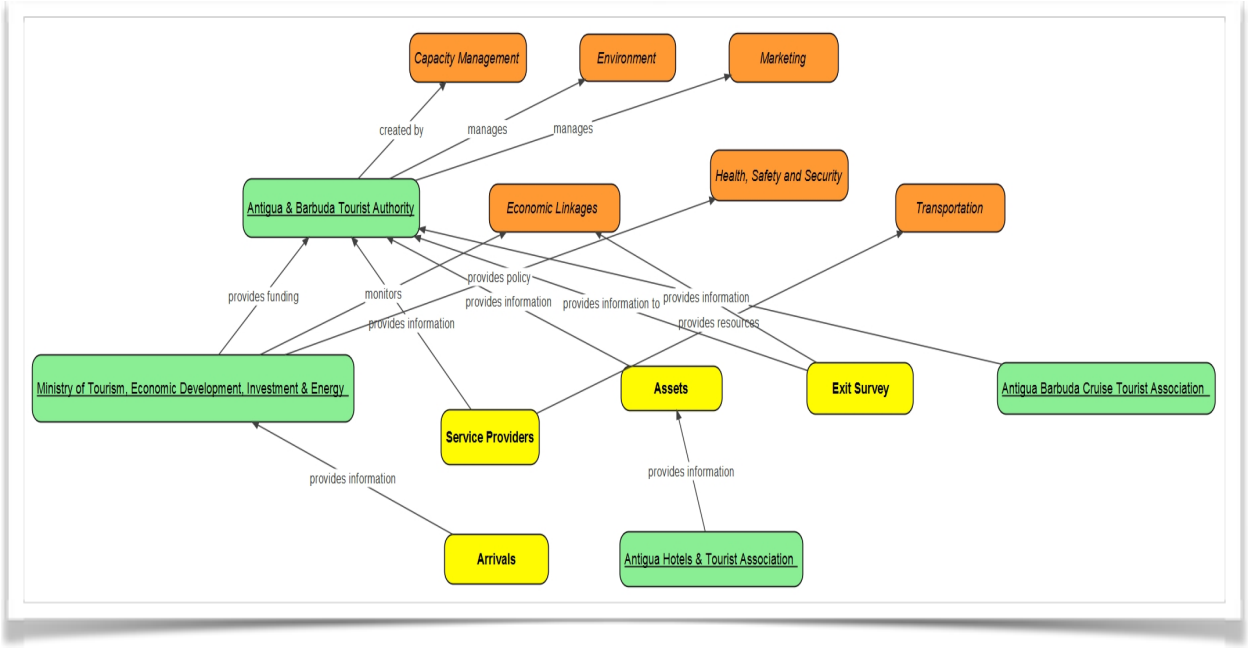


Figure 3: Antigua and Barbuda Tourism Sector Stakeholder Map

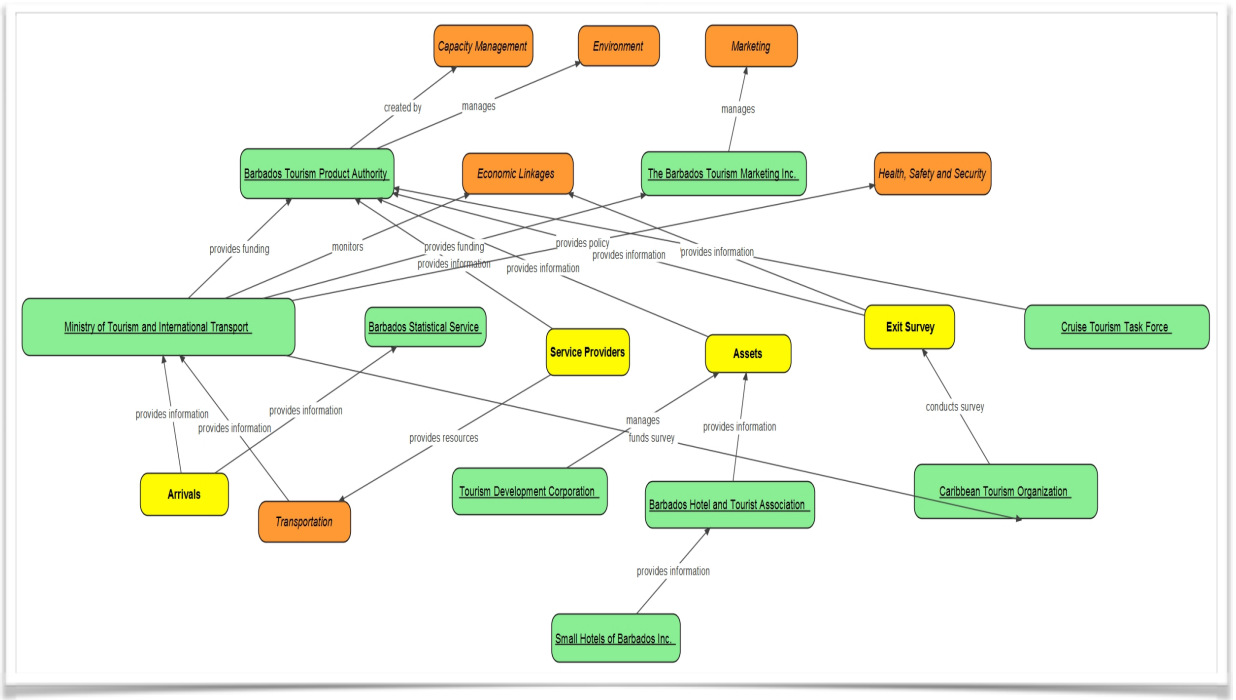


Figure 4: Barbados Tourism Sector Stakeholder Map

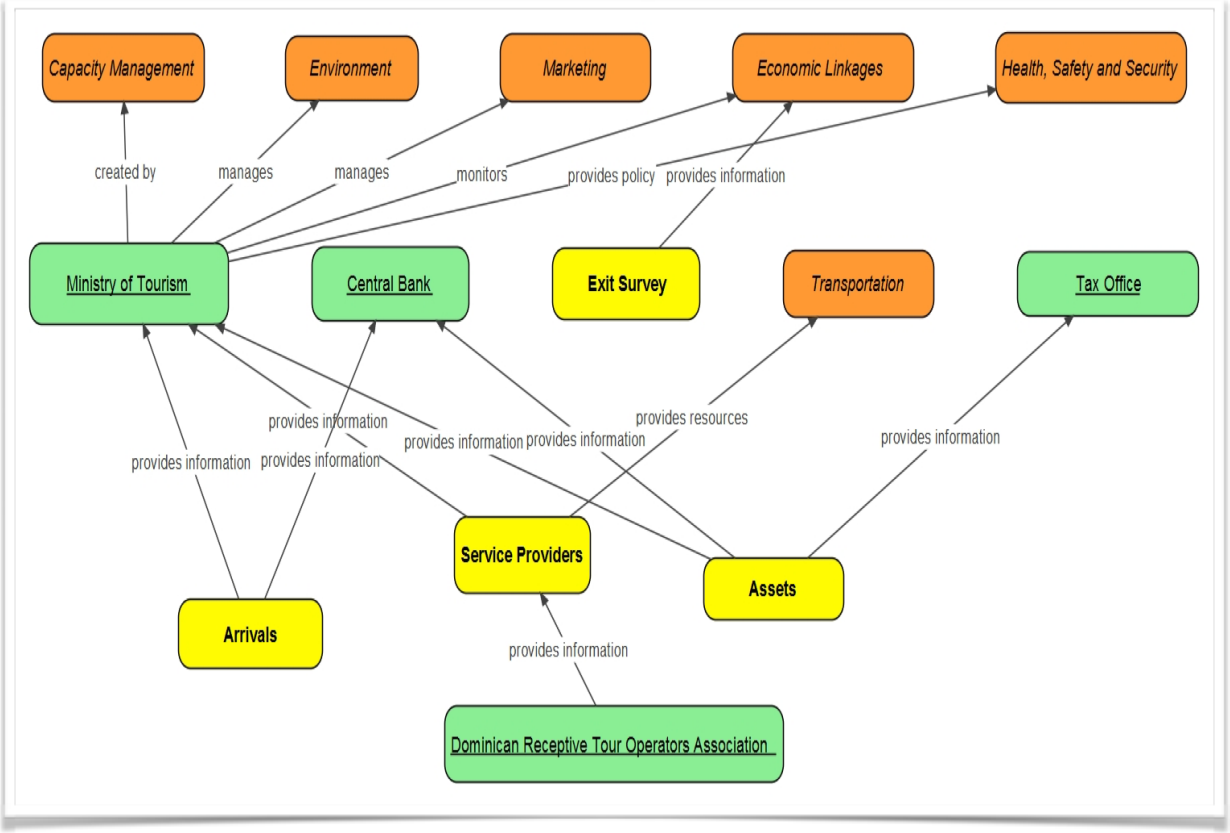


Figure 5: The Dominican Republic Tourism Stakeholder Map

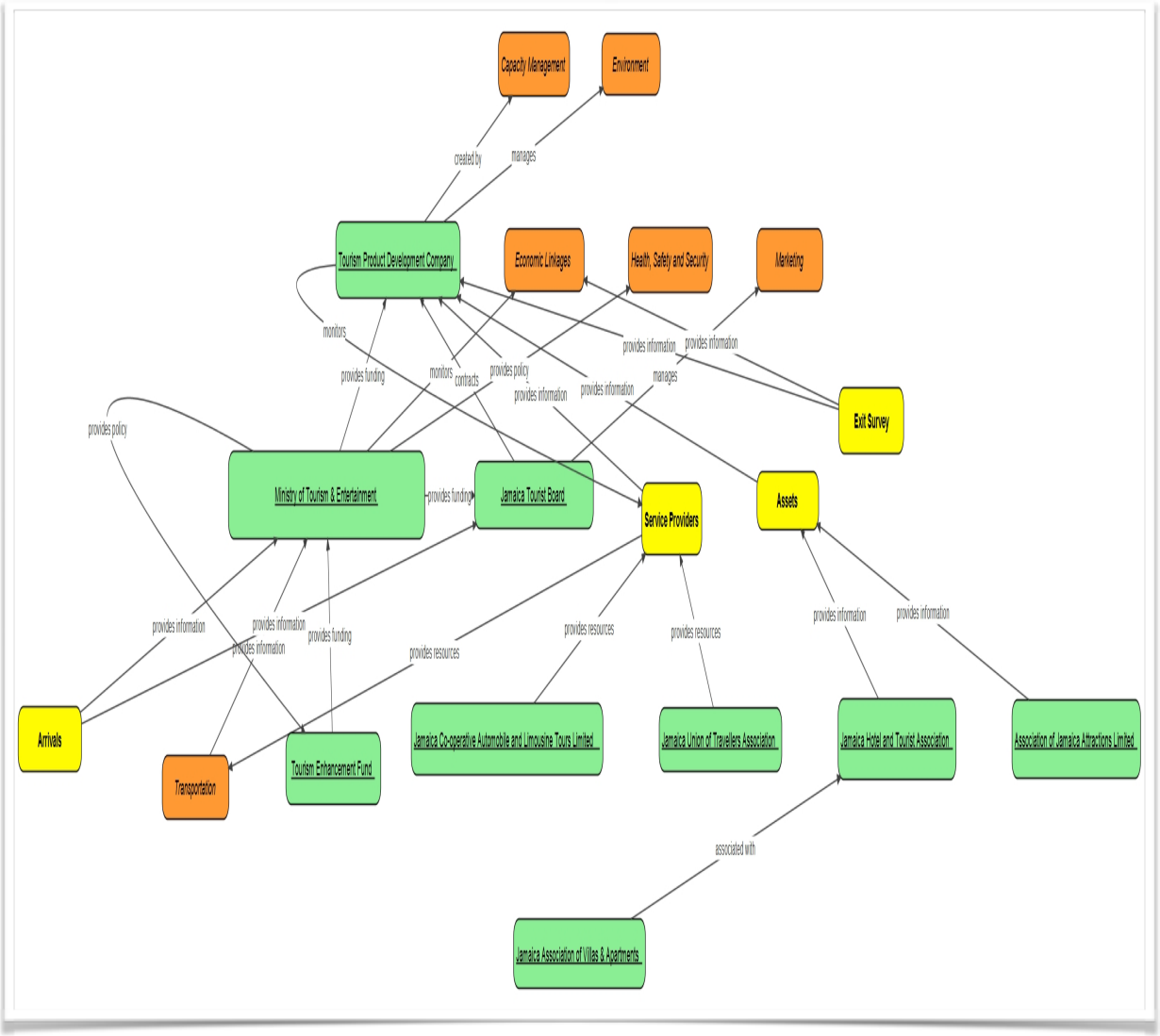


Figure 6: Jamaica Tourism Sector Stakeholder Map

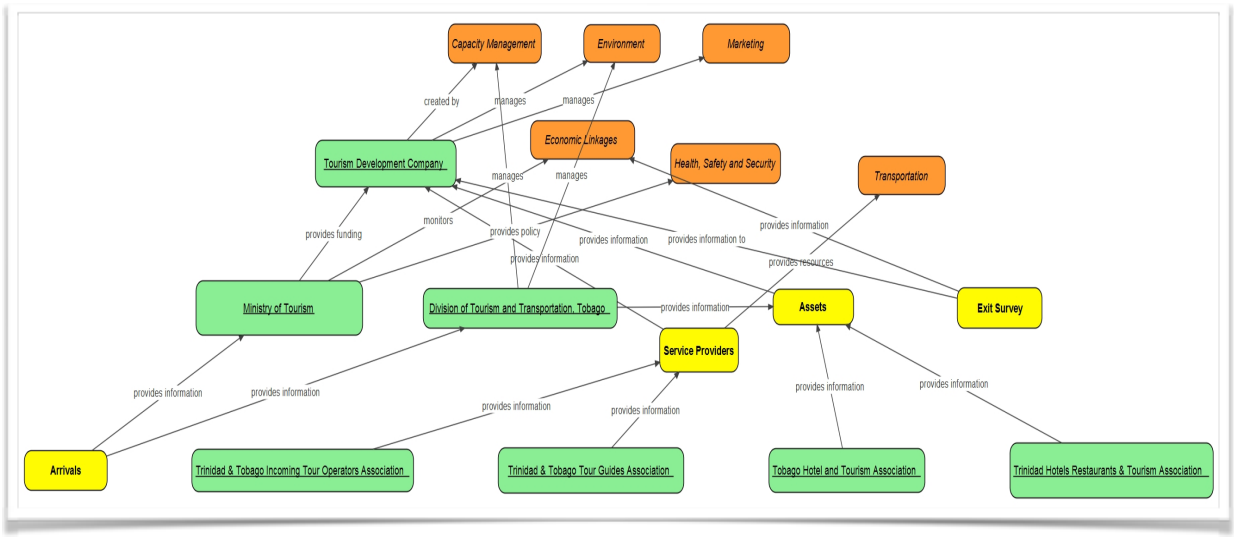
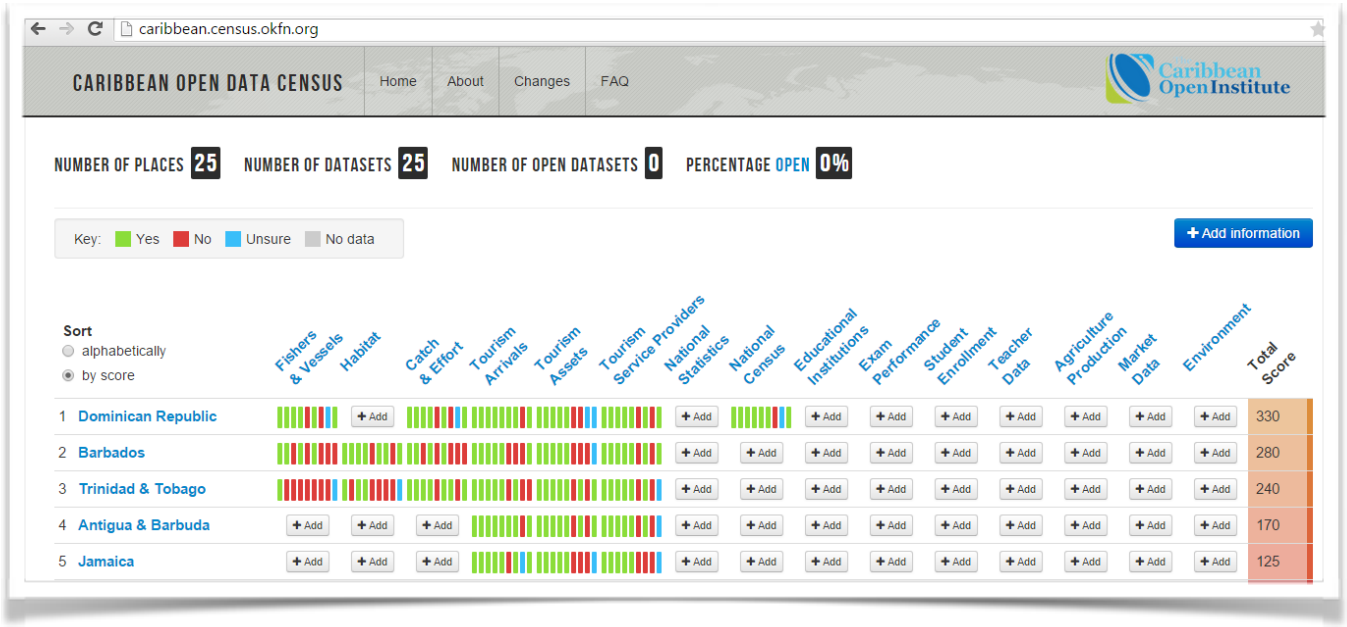


Figure 7: Trinidad and Tobago Tourism Sector Stakeholder Map

Appendix II Caribbean Open Data Census (five countries)



Source: Caribbean.census.okfn.org (retrieved 02/05/2015)

Appendix III Caribbean Tourism Data Ecosystem (tie strengths shown by line thickness)