INNOVATIONS IN LAND DATA GOVERNANCE:
UNSTRUCTURED DATA, NOSQL, BLOCKCHAIN, AND BIG DATA ANALYTICS UNPACKED

ROHAN BENNETT[1], MARK PICKERING[1] and JASON SARGENT[1]
1. Swinburne Business School, Swinburne University, Australia
rohanbennett@swin.edu.au

Paper prepared for presentation at the
"2018 WORLD BANK CONFERENCE ON LAND AND POVERTY"
The World Bank - Washington DC, March 19-23, 2018

Abstract

Unstructured data, non-relational databases, distributed databases (including blockchain technologies), and big data analytics potentially change the landscape for land data creation, management, and dissemination. The way land data is governed and the impact this can have on the delivery of land tenure is open to radical re-thinking. Drawing from international cases, this paper provides a state-of-the-art examination of cases, prototypes, and demonstrators where distributed, unstructured, non-relational, and analytical database tools are being explored and applied in the land sector. The aim is to deliver an insight into the opportunities, challenges, and lessons regarding the approaches, and how they are impacting on the broader delivery of land tenure security. The paper finds that whilst uptake of non-relational and distributed databases is occurring, it still remains largely at the level of proof-of-concept, demonstrator or pilot. Against other sectors, scaled uptake is occurring slower than anticipated meaning any assessment of the broader impacts on the land sector and broader society are somewhat premature. Meanwhile, emerging distributed analytical databases appears to be underexploited or at least underexplored by the land sector. Overall, the examined technologies are not only offering new operational approaches for the conventional land sector transaction, but also the creation of entirely new land related services, products, and actors. These embryonic developments are likely to greatly influence the land data governance landscape.

Key Words: Unstructured data, NoSQL, distributed databases, big data analytics, blockchain

1. INTRODUCTION

Unstructured data, non-relational databases, distributed databases, and big data analytics are relative newcomers to the domains of Information Systems and Computer Science, and the sub-branch increasingly known as Data Science (Dhar, 2013). In many application fields, including the domain of land governance, the concepts and related tools already impact upon methods of data creation, management, and dissemination: a suite of demonstrators, prototypes and pilots is increasingly evident.

These initial forays invite conjecture on the future of land data governance in terms of likely operational scenarios and the impacts any changes would have on the delivery of land tenure security more broadly. Like other technology-driven disruptions, the spectrum of hypothesized impacts for the sector includes everything from minor blips to existing operational procedures, to more radical suggestions of industry transformation (*c.f.* Anand, 2016).

This paper aims to deliver a balanced state-of-play insight into the opportunities, challenges, and lessons regarding these emerging land data governance approaches, and how they may impact on broader delivery of land tenure security. A contextual background on the major technical influences on land data governance over the proceeding decades is provided. Explanation of the current state-of-play follows, revealing that scaled utilization of unstructured data, distributed databases, and analytics tools is still emergent within the land sector. The paper unpacks these identified technologies, further exploring their nature and identifying early applications, demonstrations, and prototypes from the sector. On this, an exploratory approach is adopted, with published works, reports, grey literature, and author experiences informing the qualitative review. The resulting synthesis – supported by several analytical frameworks – seeks to outline the types and scale of uptake occurring, the impacts on land data governance, and the broader impact on land tenure security delivery.

2. LAND, DATA AND GOVERNANCE: A BRIEF BACKGROUND

2.1 UNPACKING THE NEXUS

Land governance is generally considered to concern: "the rules, processes and structures through which decisions are made about access to land and its use, the manner in which the decisions are implemented and enforced, and the way that competing interests in land are managed." (Palmer *et al*, 2009). Similarly, land data governance can be understood as the structures and processes for determining and enforcing how underlying land data is created, used, and managed. In this guise, it can be considered one technical subset of broader land governance.

Land governance and land data governance are studied due to the inherit complexities and challenges in delivering land tenure security. Land tenure security is sought after, studied, and measured in different ways by different disciplinary domains and industry sectors (Simbizi *et al*, 2013). Repeated failures to deliver it at scale, coupled with its recognized importance in supporting sustainable development, means it is an ongoing area of focus and investment for international agencies, donors, governments, private sector, and academia alike (Zevenbergen *et al*, 2013). Ingredients for land tenure security include, amongst others, stable political environments, funding mechanisms, mandates from leaders, societal backing, legal frameworks, administrative and technical capacities, supportive land information systems (LIS) – in the form of land registries or cadastres (Williamson *et al*, 2010), and at the heart of the nexus (at least for this work), the land data that populates those registries.

Those last elements, the land data and the information systems they feed, are the focus of much inquiry and development work over preceding decades. Land data supports land tenure security by enabling the creation and maintenance of accurate, authoritative, accessible, available and unambiguous information about people, land, and the rights that bind them (Williamson *et al* 2012; Bennett 2012). Three generations of LIS developments are evident (Bennett *et al*, 2017).

Each LIS generation builds on the previous, responding to the limitations and opportunities identified in past iterations[1].

2.2 DEFINING THE GENERATIONS

The 1st LIS generation, the conventional pre-2000s approaches, utilized standard data collection tools (e.g. total stations, aerial imagery), commercial GIS vendor platforms, and relational databases (*et al* Dale and McLaughlin, 1999; Williamson *et al*, 2010). The 2nd generation, the pro-poor approaches largely driven by the Global Land Tool Network (GLTN), sought to respond to the challenge of recording the 70% (~4 billion) unrecorded land interests worldwide (UN Habitat, 2008; 2012). These focused on creating tools that were more rapid and cheaper to apply, broadening the constellation of definable land interests, and utilizing alternative technologies (e.g. satellite imagery, mobile devices, open source technologies). The 3rd LIS generation is argued to be driven by responsibility, scalability, and industry innovation. These entrepreneurial approaches seek to transcend conventional institutions, technologies (e.g. UAVs, automatic feature extraction), and methods (de Vries *et al*, 2015; Bennett *et al*, 2017).

As mentioned, digital data lies at the heart of each LIS generation. Contemporary thinking makes several ‘societal pull’ pleas regarding the land data. First, it is ideally created in a collaborative and participatory fashion – with technical experts working alongside landholders and communities (Enemark *et al*, 2014). Second, data is flexible and inclusive in nature, representing the continuum of possible land tenure types that exist on the ground in any given context (Enemark *et al*, 2014). Third, the data is collected in a way that is affordable, reliable, attainable (at scale), and upgradable (Enemark *et al*, 2014). Scaled projects from Rwanda and Ethiopia, along with the developed Social Tenure Domain Model (STDM) software from GLTN, demonstrate that the above demands can be achieved by utilizing and aligning ‘technology push’ drivers with ‘societal pull’ drivers.

[1] Note: In the work at hand, the authors introduce ‘Generation 0’ – to encompass all those LIS where IT plays primarily only a limited supportive role to traditional paper based systems (See Section 7.4)

2.3 IDENTIFYING THE GAPS

However, whilst the more recent LIS generations demonstrate much progress towards supporting more inclusive approaches to delivering land tenure security in the form of novel data collection technologies and approaches, interestingly several key developments in data technologies appear underutilized, if not under explored. Meunier (2016) provides a useful framework for holistically understanding developments in data technologies including concepts of relational databases[2], non-relational databases, distributed databases, and analytical databases – and the intersections thereof - including NoSQL, NewSQL, distributed databases (including ledgers), and data analytics tools. The latter almost exclusively emerged from the mid-2000s onwards, with the scale of supporting technology infrastructure usually increasing in each case. We use an adapted version of the framework to guide subsequent analyses (Figure 1).

[2] Note, for this study, object-relational database management systems (ORDBMS) – as found in PostgreSQL – can be considered under the RDBMS category

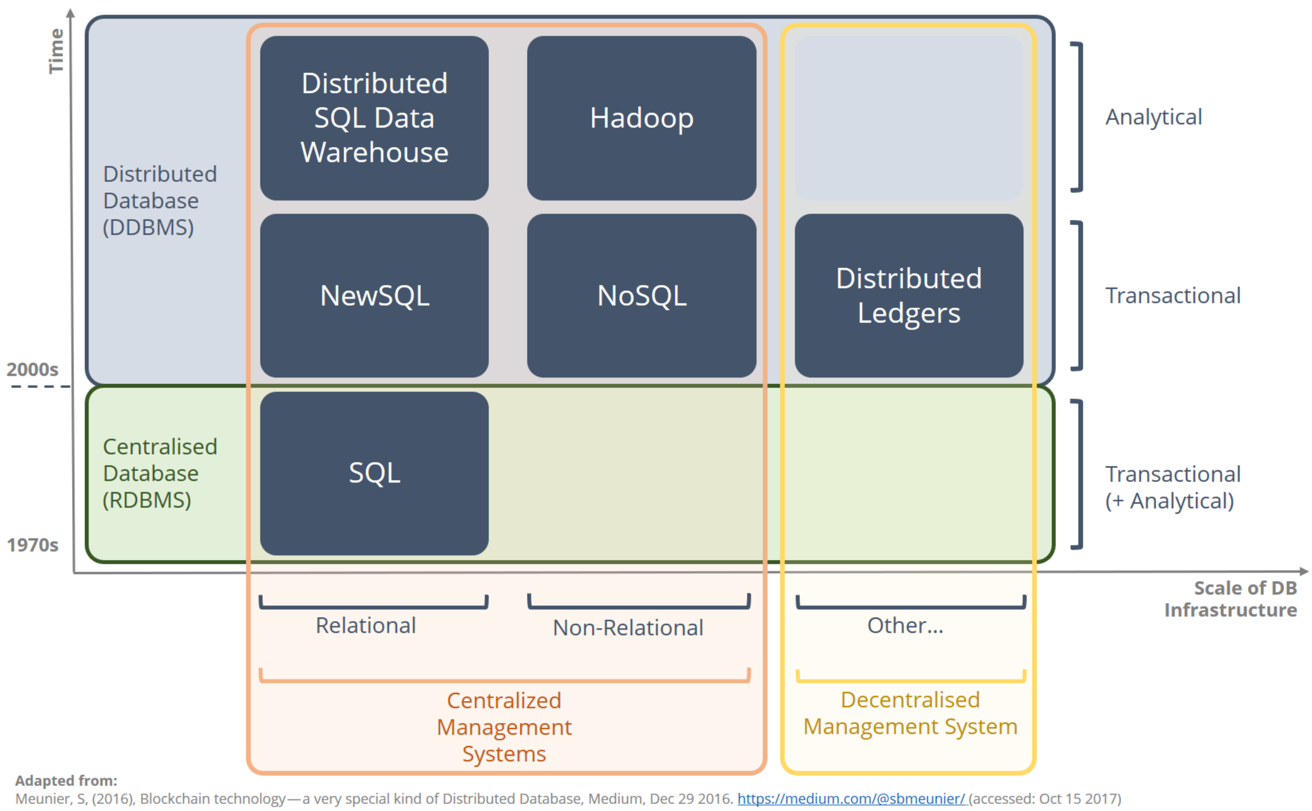

Figure 1. Overview of Contemporary Database Technologies (Adapted from Meunier, 2016)

Whilst it is an overstatement to suggest no example applications or implementations are evident, implementations and exploitation of the technologies has perhaps been slower than observed in other sectors. Hypotheses on the reasons for the lag are varied, but, ultimately untested. Is it truly the inherent conservatism within the sector? Perhaps an issue on financing? Or, is it a general lack of awareness of the technologies within the sector? Perhaps it's more to do with the significant challenges associated with implementing technical changes within such a complex sector? Or, do the nature of land data and land transactions not easily lend themselves to experiencing the benefits associated with the new technologies? Further still, perhaps there are more pressing legal, institutional, and social challenges that need addressing first? Getting answers to all these questions is beyond the scope of this paper, however, we make an initial attempt by exploring the tools, the state-of-play of their implementation in the land sector, challenges faced, and any broader impacts being experienced.

3. RISE OF UNSTRUCTURED DATA AND NON-RELATIONAL DATABASES

3.1 RETHINKING LAND DATA STRUCTURES

One contemporary ‘push’ technology already challenging conventional land data thinking is ‘unstructured data’. Contemporary database specialists refer to three major kinds of data: structured data, semi-structured data, and unstructured data (Molero *et al*, 2010). Structured data is that which is highly organized, usually by means of rows, columns, and related tables. Data types are set and input values are constrained. Its highly structured nature makes it easy to search and store – usually in relational database management systems (RDBMS). Unstructured data has no such rules for formatting and organization. It offers much more freedom in terms of data creation techniques (i.e. not just simple data entry forms), but, is more challenging to search and store coherently. Somewhere in between sits semi-structured data: the data is not formally organized in a way that RDBMS can handle, but, elements might be tagged to provide structure to the document, and make it more searchable.

The first generation of LIS were, generally speaking, built with the structured mindset (i.e. centralized relational databases) – notwithstanding the large amount of scanned deeds, titles, and/or plan documents linked to the systems. The structured approach fitted with pre-existing tabular paper-based systems used within registers from which the underlying data was transferred, and equally, at the time of LIS development, RDBMS were the mainstay, if only reliable database option available. The RDBMS arranged data about land into tables (i.e. a table for people, a table for rights, etc.) that generally mimicked, in a digital fashion, the existing paper-based registries of cadastral and mapping agencies. The SQL query language could be used to create, query, and update the tables. Implementations of RDBMs were usually centralized with all storage and maintenance occurring in a single location. Even with all the hype surrounding unstructured and semi-structured databases, these systems continue to represent 90% of the database market in terms of revenue (Meunier, 2016). Widely recognized vendors and offerings include MySQL, Oracle, IBM DB2, and PostgreSQL. By way of example, coupled with the open-source PostGIS, the latter represents the primary storage tool for Open Street Map (OSM)[3].

In this vein, RDBMS serve conventional land administration well: the RDBMS requirement for highly structured data aligns well with the highly controlled data inputs, strict management procedures (i.e. centralized land agencies), and transactional focus of conventional land administration systems. Many national land agencies continue to use the relational data model for organizing and managing their land rights data. The STDM software and versions of the Food and Agriculture Organization's (FAO) Solutions for Open Land Administration (SOLA) (i.e. Registry, Community Server) also make allowance for structured data (using PostgreSQL), although, semi-structured data types also play a role (Molero *et al*, 2010). Whilst these RDBMS-based solutions continue to be the mainstay for underlying land administration data management, the landscape is changing and systems utilizing non-relational, semi-structured and unstructured approaches are emerging.

3.2 LEVERAGING OFF THE SEMI-STRUCTURED TREND

Since the 2000s, like other sectors, semi-structured data organization techniques have permeated the domain of land administration – most prominently via GML (Oosterom *et al*, 2009) (derived from XML, Bray *et al*, 2008), LandXML (Karki *et al*, 2011; Aien *et al*, 2011; Karki *et al*, 2013; Oosterom *et al*, 2015), and RDF (and RDF/XML) (Beckett and McBride, 2004; and Çağdaş, V., & Stubkjær, 2015). On this, a suite of demonstrators, proof(s)-of-concept, and pilots utilizing semi-structured concepts and tools emerged – driven largely by two international trends: i) efforts aimed at the creation of international recognized standards of data models in the land administration domain (Lemmen *et al*, 2015; van Oosterom *et al*, 2006); and ii) developments in realizing 3D cadastres (Shojaei *et al*, 2015). For the former, the Land Administration Domain Model (LADM) or ISO 19152 provides the prime example. Whilst LADM is largely a conceptual data model, semi-structured data formats (i.e. GML) are/were at the forefront of

[3] More on data storage in OSM available at https://wiki.openstreetmap.org/wiki/Database. For data exchange, XML is used.

initiatives to realize a scalable and transferable implementation: the simplicity and flexibility of the semi-structured markup languages makes it highly suitable data as a data exchange format. More recently, efforts to make LADM interoperable with other international standards commenced, most prominently via ISO TC211 and the Open Geospatial Consortium's (OGC) Land Administration Domain Working Group[4] (Lemmen *et al*, 2017). The so-called LADM II will, amongst other aims, seeks to further link LADM to other developing spatial standards (e.g. BIM, CityGML, LandInfra). Initial technical forays are already evident (Atazedeh *et al*, 2017a; 2017b; Oldfield *et al*, 2017).

The popularity of semi-structured data relates to its simplicity in terms of data storage and exchange. Realizing these interoperability benefits does however require large amounts of transformation on existing datasets. On this, the proliferation and uptake of semi-structured standards spawned a new sub-industry focused on data re-design, data transformation, and increasingly important middleware (Crompvoets *et al*, 2009)[5]. The European Commission's INSPIRE directive – initially focused on the creation and harmonization of fundamental geospatial datasets across European countries – illustrates the significant workload requirements associated with transforming legacy datasets into formats that enable national and international interoperability[6].

3.3 EMPERIMENTING WITH THE GREAT UNSTRUCTURED

As developments relating to structured and semi-structured land data continue to roll out, the area of unstructured land data appears have the most potential for exploitation and growth. According to Harris (2016) and others[7], more data was created across 2016-2017 than the previous 5000 years of humanity. The majority of this data is unstructured – in the form of video,

[4] OGC Land Administration Domain Working Group: http://www.opengeospatial.org/projects/groups/landadmin
[5] See WeTransform and Hale Studio as an examples of organization/tool working in the data transformation domain: https://www.wetransform.to/news/2017/05/18/hale-studio-3-3-released/

[7] IBM on data volumes and creation: https://www.ibm.com/blogs/insights-on-business/consumer-products/2-5-quintillion-bytes-of-data-created-every-day-how-does-cpg-retail-manage-it/

audio, free form text, and/or IoT sensor created content. As per Molero *et al* (2010), whilst this data is indeed highly structured, in the context of data storage and searching tools, it is considered unstructured. Further adding to the plot are claims that only 0.5% of the data actually feeds into operational decision making. The aforementioned claims are open to question, however, even if only the order of magnitudes are considered reasonably accurate – an exploration appears warranted on how the data could be exploited to support land tenure security applications. Such an exploration gains traction in the context of claims that 70% or ~ 4 billion land interests remain unrecorded or unrecognized, by government or any other formal registration body (Zevenbergen *et al*, 2013).

Unstructured data already receives some consideration in existing LIS: many conventional LIS, if not most, have some provision for recording and storing various forms of unstructured data – consider the mass scanning/digitization of historical records (titles, deeds, plans) for example. Likewise, more recent developments including LADM and STDM data models provide for the inclusion of – or linkage to – unstructured data types (Lemmen *et al*, 2015). In these cases, the unstructured data supports the structured content – in order that a legitimate and/or legal land interest might be established.

Meanwhile, in the same period, two more radical experimental developments – namely the 'Talking Titler' methodology and the 'Open Source Cadastre and Registry' (OSCAR) projects – sought to incorporate unstructured (and semi-structured) at a more fundamental level. Both methods built from the premise that in many contexts, land tenure arrangements are extremely dynamic, if not conflict laden, and that any supportive data recordation tool must be inherently flexible and capable of dynamism with regards to data types and schemas. The 'Talking Titler' approach stemmed from the challenges in land reform projects in post-apartheid South Africa (Barry and Fourie, 2002) – with the concepts and related tools being iteratively developed since the early 2000s – as new technologies emerged. Latterly, it is being combined with Social Network Analysis (SNA), in seeking to visualize land tenure relationships, if not define and verify them (Barry *et al*, 2002; Barry and Khan 2000; Barry 2008; Barry *et al* 2013; Barry and Asiedu, 2017). With the same driver, the OSCAR project sought to utilize open source software and semantic web technologies – including XML and the Resource Description Framework

(RDF) – to enable the creation of an adaptive land administration system (Hall 2013; Hall and Hay, 2009).

In the context of land data governance, the two initiatives challenge convention and represent true innovation – although, transfer of either beyond R&D into a scaled production level environment remains elusive. On this, the nature of the data capture techniques (face-2-face interviews) coupled with the limited scale of data consumed and analyzed in the experiments[8] means the opportunities inherent in the exploitation of larger unstructured data repositories (e.g. social media repositories, stored outputs from smart sensors, and captured mobile device data), remain unexplored to a large extent. The challenge of determining whether emergent machine learning and data mining approaches to decipher meaning from multiple massive-scale data input sources, and assessing whether a form of land tenure verification (or personal or community land tenure audit trail) can be constructed in an automated fashion – from largely unstructured data sources (e.g. YouTube uploads, Facebook posts, Mobile GPS records, search strings, Online bill payments) – remains open. Likewise, the potential impacts, both positive and negative, and accompanying ethical issues also remain open to further exploration.

4. RAPID EVOLUTIONS AND EXPANSIONS OF DISTRIBUTED DATABASES

4.1 DISPERSING DATA AND STORAGE

Distributed database (DDBMS) offerings have matured, mutated, and morphed rapidly since the early 2000s (Manino, 2012). In more developed contexts, rapid improvements to internet connectivity and speeds have driven the innovations – as have the rise of business models based upon intelligence and analytics. Distributed databases differ from centralized RDBMs in that storage devices are dispersed across a network – and not directly attached to a central CPU

[8] Data was used from various village-scale research projects across Ghana, Nigeria, Somaliland and South Africa for Talking Titler, and sample data from Samoa in the case of the OSCAR implementation

(Meunier, 2016). These systems, in their various guises, create several opportunities for databases aimed at delivering land tenure security.

Distributed database offerings can be clustered into various flavours. Those of immediate interest to LIS – in terms of transaction orientations – are NoSQL databases (Straunch, 2019; Nayak *et al*, 2013) and so-called NewSQL databases (Aslett, 2011). Another broad group of distributed database offerings relate to those specifically designed for analytical purposes: those are dealt with subsequently – as are those with distributed management systems (i.e. blockchain/distributed ledger technology (DLTs).

4.2 EXPLORING APPLICATIONS OF NOSQL

NoSQL databases are non-relational: storage and retrieval is not based upon access to tables, data is stored in documents, and schemas are dynamic. The benefit is that frequent changes can be made to the database and the types of data added to the system, the systems tend to scale more easily, and they work well in distributed environments. However, the drawback is that installation and management is still maturing and that in some cases, characteristics of conventional database transactions – namely atomicity, consistency, isolation, and durability (ACID) – often key for land transactions, are compromised for scalability and speed.

At the level of R&D, Molero *et al* (2010) already mentions the potential of CouchDB, an open source NoSQL platform, in the context of establishing flexible database structures for land administration applications. Likewise, Hay (2013) explores NoSQL databases in the context of gracefully evolving database schemas and model elasticity – in the pursuit of an instrument centered land administration system. Reference is made to platforms such as BigTable and Cassandra and the data models of multidimensional maps of keys/pairs that underpin them. Hay (2013) cautions that unlike the NoSQL realm, in the context of land administration, data is typically not homogeneous in form and structure, suggesting they are perhaps not relevant. Meanwhile, building off Barry *et al* (2013), Egbulefu (2015) and Dabboor (2017) examine the connotations for NoSQL approaches in design work on an evolutionary and self-adapting land tenure information systems. In the case of Barry *et al* (2013), the NoSQL XML database known as eXist was utilized, although challenges were experienced with achieving true system self-adaptation. Dabboor (2017) demonstrates a potential application for adaptable land tenure information systems (LTIS) in the contexts of conflict and post-conflict scenarios. More recently, Visnjevac *et al* (2017) apply the concept to the domain of 3D cadastres. The increasing number of pilots and demonstrators suggest scaled uptake in the land administration domain is on the horizon.

At a scaled production level recognized platforms include MarkLogic, MongoDB, Apache Cassandra, Datastax, Redis, Riak, Google BigTable, CouchBase, and CouchDB. All are capable of handling large workloads at low latency with high throughput[9]. However, despite claims of grand uptake and impact (Hagen, 2014), examples of these tools applied to the land administration domain, at scale, remain scarce – at least in terms of public accounts. Esri, the market leader in commercial GIS software, announced integration with ArcGIS and MarkLogic's enterprise NoSQL DB as early as 2013[10], although, primarily only to support defense and intelligence communities, not land administration. Whilst not specific to land sector applications, McCarthy (2014) compares MongoDB (NoSQL) against PostgreSQL (RDBMS), in terms of storage and querying capabilities, using OSM data at global to local scales: MongoDB spatial capabilities are found to be lacking in comparison, however, it delivers high performance on large datasets. The final point is of particular interest in the context of the earlier point on large amounts of unstructured data that could potentially support land tenure security applications.

There are likely numerous reasons for the lack of uptake of true NoSQL databases in the land administration sector. McCarthy (2014) sheds some light here and perhaps the most influential reason is to do with the nature of the domain itself. NoSQL databases developed out of the need to cater for large transactional workloads, where low latency and high throughput are essential. These characteristics tend not to be core concerns of conventional land administration systems, and the underlying transactions, where transaction numbers (at least relating to ownership) are

[9] Google's NOSQL platform: https://cloud.google.com/bigtable/
[10] Esri partners with MarkLogic: http://www.esri.com/esri-news/releases/13-4qtr/esri-teams-with-marklogic-for-big-data-analysis-and-management

relatively small with respect to database capacity, and it remains accepted that transactions times are long in nature – being completed over days, weeks, and potentially months (at least for now).

This is not to say these generally accepted norms will hold into the future. Several trends suggest the relevance of NoSQL database technologies in the land administration sector is only likely to increase. First, a global trend in reduced times to transact and register can be observed according to World Bank studies[11]. The strengths of NoSQL databases, particularly as the platforms mature and spatial capabilities are enhanced, may be supportive in this regard. Second, the globalization of property markets and the requisite supporting administrative regimes also provides further relevance to NoSQL databases (McDougall *et al*, 2013): transaction workloads for LIS could be massively increased if administrative frameworks were to be aggregated at regional or global levels. Third, the unbundling of the land ownership rights into derivative land interests and complex commodities deemed marketable in nature (Wallace and Williamson, 2006; Bennett *et al*, 2008) – for example carbon rights, water rights, use rights, solar rights, biota rights, and so on – also potentially greatly increases the number of database transactions relating to land. Fourth, the underlying driver for Barry *et al*'s (2013) Talking Titler prototype remains: the need for a flexible and self-adaptive land transaction database – scaled for national use – is arguably only made possible via distributed NoSQL database and related technologies.

4.3 SKIRTING THE FRONTIER OF NEWSQL

Due to the limitations in earlier NoSQL databases to maintain ACID transactional characteristics an alternate modernized form of relational database management system emerged from the mid 2000s (Aslett, 2011). NewSQL databases are considered as RDBMS that match the scaled performance of NoSQL databases, whilst supporting the relational data model and use of SQL. NewSQL is essentially a label for a highly varied set of database offerings filling the gap between conventional RDBMS and NoSQL databases. Intended application areas for NewSQL

[11] World Banks' 'Doing Business' data on property registration: http://www.doingbusiness.org/data/exploretopics/registering-property

databases are those involving large numbers of short-lived, repetitive, transactions – each only including a small subset of data (Stonebraker et al, 2007). Response times are increased by removing complex joins between tables, and full table scans. More well-known NewSQL platforms, variously consisting of entirely new architectures and utilization of sharding[12], include Google Spanner, Clustrix, VoltDB, MemSQL, Pivotal's GemFire XD, NuoDB and Trafodion (Meunier, 2016). Whilst the majority of these are equipped to deal with spatial data types and related queries, there is little, or no published evidence of application in the domain of land administration.

5. HYPE, HOPE AND HAPPENINGS WITH DECENTRALIZED MANAGEMENT

5.1 DECENTRALIZING WITH BLOCKCHAIN

An even more recent development in distributed database offerings – also focused on transactions – are DLTs under which most blockchain discussions can be classified[13] (Meunier, 2016). Unlike NoSQL and NewSQL (and RDBMs), that maintain a logically centralized management system, DLTs systems use decentralized management. The setup is generally based upon a peer-to-peer network of node computers. Users lodge transactions into the network – the node computers validate the transaction and the user's status using algorithms – and once validated a new 'block' is created and added to the existing blockchain (Morrison and Sinha, 2016). This distributed setup, the driving force behind bitcoin and other cryptocurrencies, is now being piloted and implemented across many domains (Nofer *et al*, 2017; Scott *et al*, 2017) – including in the land sector – and will continue to be examined and applied, beyond the current hype period. The technology has the potential to change the nature of land administration processes, institutions, and how land tenure security is delivered (Miscione *et al*, 2017).

[12] Explanation on database sharding: http://www.agildata.com/database-sharding/
[13] DLTs and blockchain are considered synonymous hereon in.

WorldBank (2017) covers the potential applications for DLTs in the financial sectors generally, identifying land and ownership registries as opportunity spaces. Similarly, Pisa and Juden (2017) provide an overview of potential application in the development sector, incorporating a specific focus on securing property rights. Also from a global perspective, Velpuri (2017) and Velpuri *et al* (2017) couches land-related DLT opportunities in the broader context of unregistered lands and Fintech. Kshetri (2017) highlights the opportunity to combat in developing contexts. At a more refined analytical level, Anand *et al* (2016), Vos *et al* (2017), Lemieux, (2016), Veugar (2017), and Snäll (2017) explore DLT applications specifically for the land and real estate sectors, analyzing the implications, disruptions and consequences in a more circumspect fashion.

Meanwhile, Ølnes *et al* (2017) whilst appreciating the potential benefits of DLT land registry applications, suggest uptake must be needs driven – rather than technology driven – and that requisite institutional transformations are not yet well understood. Barbieri *et al* (2017) is more critical, suggesting DLT land registries may actually represent a retrograde step towards deeds style registration, and that distributed trust could eventually revert to unintended monopolization. Likewise, Arrunuda (2017) also outlines several unresolved difficulties relating to contract completion and enforcement[14]. Griggs *et al* (2017), using the Australian context, reveals that whilst DLTs may assist in preventing some common forms of fraud, the technology would still not mitigate others. In response to such arguments, Graglia (2017a) seek to dispel several myths surrounding DLT security, uptake, and control, whilst Graglia and Mellon (2017) go on to suggest that any successful implementation require several prerequisites to be in place – for example preliminary record digitization and established salutation for identity.

Various initiatives at country-level (if not regional) are evident – and in several cases well documented. The projects range from proofs-of-concept, demonstrators, pilots, and in some cases production level environments. Prominent examples include Sweden/Chromaway (Snäll, 2017; Hjelte, 2017), Georgia/Bitfury (Ugrekhelidze and Grigolia, 2017), Ukraine/Bitfury (Graglia, 2017); Dubai/ERES (Alsuwaidi, 2018); Brazil/Ubitquity (Anand *et al* 2016; Worldbank

[14] Other critical assessment are available in the European Property Law Journal's (Vol 6, No 3) special issue on blockchain applications in the property sector.

2017; Lemieux *et al*, 2018), Ghana/Bitland (Obderdorf, 2017; Kewell *et al* 2017), Global/Propy[15] – and the stalled Honduras/Factom case[16] (Annand *et al*; 2016; Drucker, 2016), amongst others.

Meanwhile, other cases are more aspirational, experimental, or hypothetical. In this regard, the analyses of Bal (2017), Spielman (2017), Kombe *et al*, (2017), van Bochhove *et al* (2016), and Savu *et al* (2017) focus respectively on potential DLT applications in India, Nashville (Tennessee, USA), Tanzania, the concept of 'bitsquares', and decentralizing property development quality assurance in the context of smart cities[17]. The abovementioned examples and case studies suggest an increasing number and flavours of documented DLT initiatives. In many cases it is simply too early to make claims about impacts: ongoing monitoring will reveal the relative costs and benefits. We now focus upon two less exposed cases to further enrich the discussion.

5.2 EXPLORING 'LAND LAY BY' IN KENYA

Rather than working with the Kenyan government to transform the country's land registries, Land Layby is attempting to gain a critical mass of citizens using their planned DLT system to register and transfer land ownership ultimately resulting in Ministry of Lands adoption (Land Layby, 2017). Land Layby is a Kenyan parent company with offices in Melbourne, London and the New York (Land Layby, 2018). Over recent years Land Layby's primary business model has been selling land through the use of land options. The company acquires large parcels of pre-developed land outside of major cities in Kenya and then sells options to customers to buy 1/8th of an acre blocks at an agreed price in the future, typically two years. The customer pays for the options in ten monthly instalments and settles prior to the option lapsing if they wish to proceed.

[15] See more: https://propy.com/
[16] Factom explanation of stall: https://www.factom.com/blog/a-humble-update-on-the-honduras-title-project
[17] It should be noted that the listed developments do not take into account initiatives occurring in non-English contexts (e.g French, Spanish, etc.)

In the meantime, Land Layby goes through the regulatory processes to subdivide the land. The offices in Melbourne, the UK and US enable the Kenyan Diaspora to purchase land back in Kenya. More recently, the company has commenced operations in Ghana.

Land Layby has announced plans to develop and introduce Land Layby DLT Land Registries in Kenya (Land Layby, 2018). The DLT registry is being developed in Etherium and is planned to include smart contracts to automate some components of land transactions. The company indicates that the DLT registry will mirror the records at the Ministry of Lands registry, although, how that will be achieved is somewhat unclear. The ledger is to be public but with multiple levels of permission to enable some aspects of the transaction, such as sales price, to be kept private. Intermediaries are to be incentivized to participate in the DLT registry and to authenticate transactions through the issuance of Harambee Tokens. The company is planning to deploy a minimum viable product of the registry in early 2018, perform an Initial Coin Offering (ICO) for the Harambee Tokens in mid-2018 and launch a Beta version with advanced features to the community in June 2019 (Land Layby, 2017).

Placing the above into context, a recent World Bank report noted problems with land governance and land information in Kenya (Kameri-Mbote, 2016). The report concluded that while there is a legal framework for the recognition and protection of land rights, these are not effectively implemented or enforced. Most land, particularly that in rural areas, is not registered, land information is largely manual reducing its effectiveness and “the need to re-organize, update, authenticate and digitise land information for ease of access and facilitation of the recognition and protection of rights” (p. 7). The Kenyan government has commenced the process of digitizing land records, launching an on-line search registry in Nairobi in January 2016 to be rolled out in 47 other registers (MyGov, 2016). In November 2017 the Lands CS indicated that his Ministry was seeking Sh17 billion from cabinet for a five year programme to complete digitizing the entire Kenyan land registration system with the objective of reducing corruption and increasing efficiency (Muhindi,2017).

Land Layby have indicated that they expect their planned registry to go beyond improving land information to substantially changing how land transactions are performed (Land Layby, 2018).

The company claims that using a DLT approach will increase the speed of transactions by having an agreed source of land information and reducing the need for intermediaries; will increase investment value by reducing transaction costs through the use of smart contracts; the immutable nature of information in the DLT will curb rampant fraud in the Kenyan land sector; including transaction data and history in the DLT will increase transparency (Land Layby, 2018).

The Land Layby Land Registries are in development and yet to be rolled out. The relative success or failure will emerge over time. The authors have identified a number of potential technical, political and financial challenges that will need to be overcome. Potential technical challenges include developing smart contracts which incorporate legal and regulatory land transaction requirements and authorities and authenticating whether terms have been met; populating the DLT with authenticated land ownership data where much existing data is manual or contested; dealing with conflicts between the DLT registry and the Ministry registry if the DLT registry is to be immutable but not all transactions are processed on the DLT registry. Potential political challenges include attempting the project without the support of the Kenyan government with land governance highly political and control of the land registry a 'critical factor in corrupt land dealings' (Basset, 2017) and the government currently focused on digitizing the current registers (Muhindi, 2017). Authenticating current land ownership and transactions is likely to require support from intermediaries, the people that Land Layby are seeking to remove from the process. Potential funding and business model challenges include the potential costs of development, data population of the DLT and authentication of transactions, which are likely to require substantial investment. The ICO of the Harambee Tokens is planned to raise capital for development and population of the DLT registry. The tokens are also planned to be used to incentivise participation in the platform and transaction authorisation. Funding of the project may be problematic if the ICO is unsuccessful either due to a general drop of investor interest in ICOs or if the Land Layby Business model is not seen as attractive. The economics of the proposed registry are underpinned by participants seeing value in the tokens and performing authentication in order to earn additional tokens.

Land Layby's attempts to establish an alternative DLT land registry will be interesting for scholars as another emerging case to explore the use of DLT in land registries, but also to examine the dynamics of a private provider endeavoring to set up a land registry without government support.

5.3 PROFILING 'FRACTIONAL PROPERTY INVESTMENT' IN AUSTRALIA

Fractional Property Investment represents another land related area where DLT may find application. The concept seeks to break conventionally registered land parcels into conceptual shares – potentially numbering in the hundreds or thousands. Investors then buy and sell those shares in a managed marketplace. The shareholders receive a portion of the regular rent payments in the event that the property is leased and are entitled to a share of the sale value if the property is sold[18]. The approach is intended to reduce the high entry barriers in land markets where start-up costs are prohibitively expensive. Some schemes seek to support both passive market investment, others enable investors to lead syndicates and ultimately live in the purchased house.

The concept sits at the intersection of the land sector and fintech: solution providers build upon established land administration infrastructures (although, may work quite separately from them) and utilize emerging technologies (in this case online crowdsourcing tools and *potentially* DLTs) to conceptualize, create and sell derivative land related value propositions. In this regard, fractional property shares merely represent another financial derivative underpinned by the formal land sector (Wallace and Williamson, 2006).

The model recently gained traction in Australia: property markets in the major capital cities are relatively expensive in global terms, following almost three decades of non-negative growth with only a slight hiatus following the 2008 global financial crisis. Solution providers emerge to service portions of the market that are unable to gain a foothold, primarily so-called millennials:

[18] More on fractional property investment: https://www.canstar.com.au/fractional-property-investment/

Brickx[19], DomaCom[20], and CoVesta[21] represent the prominent examples. BricVest[22] provides a similar example – based out of the United Kingdom. All are relatively new players with number of properties falling under the schemes still being relatively small (e.g. Brickx holds 14 properties at the time or writing, whilst DomaCom held 43 as of June 2017[23]).

Brickx is a fully online platform, with investment opportunities available for under $100. It divides each of its properties into 10,000 'bricks'[24]. Users create a digital wallet and must have a minimum deposit of $75 in the wallet upon start-up. As of February 2018, it had over 9000 members with more than 17,000 transactions being completed in the previous 90 day. The median sale time for each 'brick' is around 16 hours – and brick buyers are able to enter and leave the market at any time. The Brickx business model is based around the payment fees for each 'brick' transaction (1.75%), property management fees, and an annual audit fee. Legally, Brickx uses a unit trust arrangement for each property – the approach anchoring each property to the formal land administration system. CoVesta uses a similar model to Brickx, albeit offering the investor more insights and control over which properties are purchased. Typically after five years, a 'vote to sale' is held amongst all shareholders. Share sizes are also different. DomaCom uses different approach being regulated as a managed investment scheme (MIS). It provides flexibility in that investors can lead syndicates and control the selection of properties.

As outlined above, the motivation and benefits for Fractional Property Investment appear readily apparent, however, the approach raises several questions about rates of return on investment, and whether the approach really does anything to increase levels of property ownership and underlying tenure security. These issues are important in the context of developing countries where the model could be potentially applied in low-income or poverty stricken areas to support gaining footholds in the formal property market: the private sector, civil society, or even

[19] More on Brickx: https://www.brickx.com/how-it-works
[20] More on DomaCom: https://domacom.com.au/about-domacom/
[21] More on CoVesta: https://www.covesta.com.au/
[22] More on Bricvest: https://brickvest.com/en/
[23] Annual Report for DomaCom: https://domacom.com.au/investor-relations/financial-reports/
[24] Brickx' name for a fractional property share

government corporations could fill the gap between costly government land administration systems and the low-income landless.

The Australian technical platforms enabling Fractional Property Investment are not necessarily utilizing DLTs – although, Graglia and Mellon (2018) reveal the ongoing development of a Brickx-style DLT development known as pangio.io[25]. Looking ahead, according to head of blockchain at one of Australia's most prominent banks[26], fractionalization of large 'lumpy' assets into smaller tradeable shares is a key growth area. If true, there are significant implications for the land sector: the unbundling of land and resource rights into smaller commodities – a decades long trend often impeded by administrative and technical complexity (Bennett *et al*, 2008) – may find renewed traction and impetus. Drawing on an example from the Australian commercial lease sector[27], Mike Craglia[28] of the New America Foundation hypothesizes that it is these less administratively complex and less high valued land-related rights where DLTs may find scaled application initially. In particular, the ability to code or 'build-in' temporal, spatial, and other constraints via smart contracts, and the colored coin concept are particularly attractive (*c.f.* Anand *et al*, 2016). As always, this all relies upon the underlying base resource (i.e. the land) carrying both high enough value and high administrative and legal certainty.

6. THE SLOW BURN OF BIG DATA ANALYTICS FOR LAND TENURE

6.1 LOOKING FOR HADOOP

The other key development in the distributed database space is so-called 'Big Data'. These are distributed databases that have a specific focus on data analysis and analytics – as opposed to

[25] http://www.pangea.io/home
[26] More from Sophie Gilder: https://www.finder.com.au/commbank-possibilities-blockchain
[27] More on the Australian commercial lease example: https://bluenotes.anz.com/content/dam/bluenotes/documents/whitepaper%20_bank_guarantees_dlt_poc.pdf
[28] Personal communication on 18 January 2018 via Skype

transactions. These are distributed versions of earlier data warehousing technologies. The offerings can be relational (i.e. distributed SQL data warehouses e.g. Oracle Spatial) or non-relational (i.e. Hadoop, Open Source) (Hay, 2016). Either way, the focus is on delivering fast response times for queries on extremely large data sets – something centralized relational databases have tended to struggle with. The challenge applies to spatial applications, where large-scale spatial query processing is increasingly required. Already experimentation seeks to extend existing distributed/non-relational analytical databases (e.g. the Hadoop ecology) to support improved spatial querying: Huang *et al*'s (2017) development and testing GeoSpark SQL providing a pertinent example.

The domain of land administration systems has been typically slow to utilize data warehousing and big data analytics: the historical focus has been on delivering transactional efficiency, with less interest in extracting business or social insights from the data. That said, a body of discourse suggests increasing interest in big data analytics across the land sector.

Kwanya's (2014) analysis suggests high levels of viability and fitness with regards to utilizing big data technologies within Kenya's land sector, at least based on the perceptions of land sector actors, however the specifics of the so-called 'big data' implementation are not really touched upon at all. Meanwhile, Zoomers *et al* (2016) explores the potential role of big data in the global land rush discourse in rural areas: big data can enable new networked and visual analyses of the global land grab. However, these opportunities are balanced against the need to negate against purely quantitative analysis and poor cartographic expression. Similar to rural areas, urban areas also receive big (land) data attention with regards to analytics and visualization in the contexts of smart cities and 3D cadastres (*et al* Sabri *et al*, 2015; Jazayeri *et al*, 2014; Pouliot *et al*, 2016).

Possibly the greatest untapped opportunity in the big data space for land administration is in land tenure security provision – particularly in developing contexts (Barry, 2013). As discussed under the pretexts of unstructured data and NoSQL, by bringing together both structured and unstructured data – from a myriad of official and less official sources – there lies the potential to apply descriptive machine learning techniques in order to reverse-engineer 'land tenures' through the identification of patterns in social and spatial data. As mentioned earlier, whilst a slum dweller in a given context may not have access to a 'legal title' over a land holding, records from social media, phone records, utility payments, can be combined using distributed data warehousing tools and thus demonstrate a dweller has a 'legitimate' holding – based on history of occupation and servicing of the land. Barry and Asiedu (2017) already demonstrate the potential for the approach in work on further work combining the Talking Titler methodology, data mining and social network analysis. The highly innovative cross-disciplinary work invites a radical rethinking of how land rights established and land tenure is security. It also represent a sought after example of balanced technology-push and societal-pull forces. However, like most big data innovations for the land sector, the concept remains to be scaled to production level. Additionally, the examples provided above only deal with descriptive machine learning: the potential for predictive or prescriptive (i.e. recommendations) analytics tools in the land sector remains untapped[29].

7. SYNTHESIS ON IMPLEMENTATIONS, IMPACTS, AND INSIGHTS

Condensing the abovementioned technology developments, application areas, and case examples, a synthesis of key findings is presented. A specific focus is placed on: gauging trends in overall uptake; the apparent gap regarding analytics; expectation management within the land sector; whether the new technologies are impacting on operations, strategy, or both; and the challenge of assessing impacts generally.

In terms of uptake of the emerging data governance technologies on the land sector, a general trend towards uptake can be observed. Like other sectors, whilst centralized/relational databases remain significant in the land sector and 1st Generation LIS, there is evidence of experimentation, demonstrators, and some scaled uptake of combinations of distributed and non-relational databases in 2nd and 3rd Generation LIS. Most examples sit at the level of demonstrator or pilot, but the count and diversity in applications is quickly increasing. Meanwhile, the application of

[29] For more on descriptive, predictive and prescriptive machine learning see: https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/an-executives-guide-to-ai

distributed analytical databases in the land sector appears minimal. Very few, if any, examples of distributed SQL data warehouses or application of Hadoop are apparent. The same applies to applications of NewSQL databases where examples of land-related applications remain scarce. As mentioned in the text, there is likely multiple reasons for the minimal uptake, however, untapped opportunities as espoused in Barry (2013), Zoomers et al (2016), and Cunningham et al (2018) are worth of further examination.

Despite the experimentation and demonstrations, scaled uptake appears to lag behind expectation, particularly in terms of the scale of 2nd Generation LIS implementation. In many regards, this lag is no different to other sectors. These scaled implementations are likely to rely on the 'abilities and cooperation between well trained notaries, surveyors, registrars, and information managers' as much as any technological issues (*c.f.* Henssen, 2010): HR capacity only appear to be gaining in importance when it comes to technology diffusion (Hunt and Allas, 2018).

Uptake and the scale of it aside, are the new data technologies actually changing the nature of land administration? We answer this question using Bennett et al (2005) and Applegate et al's (2003) theoretical frame. In brief, the new data technologies are influencing conventional land administration agencies (e.g. DLT examples in Georgia and Sweden) but, also enabling new products and services (e.g. Land Lay By). However, these innovations need to be couched in the broader context where the majority of conventional LIS are still in formation, piecemeal, incomplete, and sitting alongside equally incomplete or corrupted traditional paper-based systems. Building on Bennett *et al's* (2017) model, we refer to these systems as Generation 0 and these may still account for as much as 70% of existing LIS globally. Perhaps the biggest question to ask is should the 70% seek to move systematically into Generation 1 systems or leapfrog into later generations. If the latter, what are the most suitable development pathways?

The above assessment are largely preliminary, and could be built upon by further comparing the capabilities of the new technologies against commonly agreed upon land sector frameworks and requirements – for example: the four basic principles of land registration as espoused by Henssen (2010) including 'booking', 'consent', 'publicity', and 'speciality' – and the add-on Torrens principles of 'mirror' 'curtain' and 'insurance'. Already, Vos *et al* (2017) provides an initial examination in the context of DLTs, finding DLT may support some elements, undermine others, dependent on the specific implementation. Likewise, Larsson's (1991) 4 levels of evidence for land transactions – evolving from witnesses, through deeds, to deeds registration, and finally title provision – provide another frame. Zevenbergen (2017) undertakes an initial review concluding that emerging technologies are digitally replicating the forms of evidence, and that on the surface at least, DLT would appear a reversion to Deeds registration. Other frameworks worthy of comparison might include, in no particular order, Dale's (2000) property market pillars; Kaufman and Steudler's (1998) Cadastre 2014 statements; Enemark *et al*'s (2014) Fit for Purpose Land Administration principles and elements; the 4 conventional tenure types (*c.f.* Dalrymple, 2004); UN Habitat's (2008) Continuum of Land Rights concept; Dale and McLaughlin's (1999) 3 cadastral types (fiscal, juridical, multipurpose); Bennett *et al*'s (2008) RRRs toolbox, Bennett *et al*'s (2012) categories of land data application; and Bennett *et al*'s (2010) 'cadastral future's statements.

The above suggestions again bring to the surface the challenges of assessing land related interventions – including those related to land data technologies. Impact assessment is a fraught exercise. At a fundamental level, any implementation should seek to demonstrate enhancements related to the usual aims of land administration (e.g. tenure security) (Henssen, 2010) and also make assessments about system security, clarity, simplicity, timeliness, fairness, accessibility, cost and sustainability (Henssen, 2010). Going further, land data is part of a wider LIS including business logic, presentation layers, all couched within an operational setting consisting of people, processes, and other technical and non-technical systems: a more robust assessment would need to incorporate the flow-on effects to these sub-systems. Whilst there already exists a significant and growing body of activity in this domain (e.g. LGAF, UN SDGs[30] and PRIndex) the area of Social Impact Assessment (SIA) may provide some relevance to the land sector and require further exploration. The emergent frameworks are holistic, multidimensional, responsible and

[30] More on UN-SDGs: https://sustainabledevelopment.un.org/sdgs

pragmatic (*c.f.* Aledo-Tur and Domínguez-Gómez, 2017; Mahmoudi *et al*, 2013; McLoughlin *et al*, 2009) are potentially worthy of further exploration and integration.

8. CONCLUSIONS

This paper revealed that unstructured data, non-relational (NoSQL) databases, distributed databases, and big data analytics offer the potential to change land data governance in terms of creation, management, and dissemination. It is argued that the way land data is governed and the impact this may have on the delivery of land tenure is open to re-thinking. A state-of-the-art examination was undertaken of cases, prototypes, and demonstrators where distributed, unstructured, and analytical database tools were being experimented in the global land sector. An insight into the opportunities, challenges, and lessons regarding the approaches, and whether they are impacting on the broader delivery of land tenure security was provided. It was found that whilst uptake of non-relational and distributed databases is occurring, it still remains largely at the level of proof-of-concept, demonstrator or pilot. Against other sectors, scaled uptake is occurring slower than anticipated, but, any assessments of the broader impacts on the land sector and broader society are likely to be premature. Meanwhile, emerging distributed analytical databases appear to be underexploited or at least underexplored by the land sector. Overall, the examined technologies are not only offering new operational approaches for the conventional land sector, but also the creation of entirely new land related services, products, and actors. It is recommended in all cases, more thorough examinations of the legal, ethical, and institutional implications is undertaken.

ACKNOWLEDGEMENTS

We the authors would like to acknowledge the support and any contributions from our colleagues at the Swinburne Business School, Faculty of Business and Law, Swinburne University of Technology. The authors also wish to acknowledge that at the time of upload, an extended version of this paper may be under review within an academic journal

REFERENCES

Aien, A., Rajabifard, A., Kalantari, M., Williamson, I., & Shojaei, D. (2011). 3D cadastre in Victoria Australia. GIM International, 25(8), 16-21.

Aledo-Tur, A., & Domínguez-Gómez, J. A. (2017). Social Impact Assessment (SIA) from a multidimensional paradigmatic perspective: challenges and opportunities. Journal of environmental management, 195, 56-61.

Allas, T., and Hunt, V., (2018), Accelerating the diffusion of technology-enabled business practices, Mckinsey and Company, McKinsey Quarterly (February 2018), https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/accelerating-the-diffusion-of-technology-enabled-business-practices (accessed February 2018)

Alsuwaidi, K., (2018), Dubai Real Estate Block Chain (Extended Abstract), In Proceedings of the 2018 World Bank Conference on Land and Poverty, (March 19-23), The World Bank, Washington, DC, USA.

Anand, A., McKibbin, M., & Pichel, F. (2016). Colored coins: Bitcoin, blockchain, and land administration. In Annual World Bank Conference on Land and Poverty.

Arrunada, B., (2017), Blockchain's Struggle to Deliver Impersonal Exchange, In Proceedings of the 2017 World Bank Conference on Land and Poverty: Responsible Land Governance–Towards an Evidence-Based Approach, The World Bank, Washington, DC, USA.

Aslett, Matthew (2011). "How Will The Database Incumbents Respond To NoSQL And NewSQL?", 451 Group (published 2011-04-04). (accessed 08/02/2018)

Atazadeh, B., Kalantari, M., Rajabifard, A., Ho, S., & Champion, T. (2017a). Extending a BIM-based data model to support 3D digital management of complex ownership spaces. International Journal of Geographical Information Science, 31(3), 499-522.

Atazadeh, B., Kalantari, M., Rajabifard, A., & Ho, S. (2017b). Modelling building ownership boundaries within BIM environment: A case study in Victoria, Australia. Computers, Environment and Urban Systems, 61, 24-38.

Applegate, L. M., Austin, R. D., & McFarlan, F. W. (2006). Corporate information strategy and management. McGraw-Hill/Irwin Custom Publishing.

Bal, M., (2017), Securing Property Rights in India through Distributed Ledger Technology, ORF Occasional Research Paper, New Dehli, India.

Barbieri, M., *et al* (2017), Blockchain – Can This New Technology Really Revolutionize The Land Registry System? In Proceedings of the 2017 World Bank Conference on Land and Poverty: Responsible Land Governance–Towards an Evidence-Based Approach, The World Bank, Washington, DC, USA.

Barry, M. (2013). Land tenure administration and GIS: trends and challenges. South African Surveying and Geomatics Indaba. Johannesburg, South Africa.

Barry, M., & Asiedu, K. (2017). Visualising changing tenure relationships: the talking titler methodology, data mining and social network analysis. Survey Review, 49(352), 66-76.

Barry, M., & Fourie, C. (2002). Analysing cadastral systems in uncertain situations: a conceptual framework based on soft systems theory. International Journal of Geographical Information Science, 16(1), 23-40.

Barry, M., Molero, R., & Muhsen, A. (2013). Talking titler: Evolutionary and self-adaptive land tenure information system Development. South African Journal of Geomatics, 2(1), 1-12.

Basset E.M. (2017). The challenges of Reforming Land Governance in Kenya Under the 2010 Constitution. Journal of Modern African Studies. 55(4). Pp. 537-566.

Bennett, R. M. (2012). The cadastral divide: a view from the bridge. Geospatial world, 2(11), 42-45.

Bennett, R., Wallace, J., & Williamson, I. (2005). Integrated land administration in Australia: the need to align ICT strategies and operations. SSSI Conference, Melbourne, Australia.

Bennett, R. M., Wallace, J. & Williamson, I. P. (2008). A Toolbox for mapping and managing new interests over land. Survey Review, 40(307), 43-53.

Bennett, R., Rajabifard, A., Kalantari, M., Wallace, J., & Williamson, I. (2010). Cadastral futures: building a new vision for the nature and role of cadastres. In FIG Congress (pp. 11-16).

Bennett, R. M., Gerke, M., Ho, S., & Chipofya, M. (2017). Building Third Generation Land Tools: Its4land, Smart Sketchmaps, UAVs, Automatic Feature Extraction, and the...

Bray, T., Paoli, J., Sperberg-McQueen, C. M., Maler, E., & Yergeau, F. (2008). Extensible markup language (XML) 1.0.

Çağdaş, V., & Stubkjær, E. (2015). Core immovable property vocabulary for European linked land administration. Survey Review, 47(340), 49-60.

Crompvoets, J., Rajabifard, A., van Loenen, B., & Fernandez, T. D. (2009). Future directions for SDI assessment. In GSDI 11 World Conference , Rotterdam, The Netherlands .

Cunningham, S., David, J., Dohrmann, T., (2018), The trillion-dollar prize: Plugging government revenue leaks with advanced analytics, Mckinsey&Company, http://www.mckinsey.com/industries/public-sector/our-insights/the-trillion-dollar-prize-plugging-government-revenue-leaks-with-advanced-analytics?cid=eml-web (accessed 15 February 2018)

Dabboor, A, (2017), An Adaptive Land Tenure Information System Database Design for Conflict and Post Conflict Situations, MSc Thesis, University of Calgary, Canada.

Dale, P. (2000). The importance of land administration in the development of land markets: a global perspective. Proceedings of UDMS 2000, 31-41.

Dale, P., & McLaughlin, J. (2000). Land administration. OUP Catalogue. UK.

de Vries, W. T., Bennett, R. M., & Zevenbergen, J. A. (2015). Toward responsible land administration. In Advances in responsible land administration. CRC Press.

Dhar, V. (2013). Data science and prediction. Communications of the ACM, 56(12), 64-73.

Drucker, P. (2016). Blockchain applications in the public sector. London, United Kingdom: Deloitte

Egbulefu, I., (2015), Designing a Technical Solution for Managing Land Tenure Data in a Customary Setting, MSc Thesis, University of Calgary, Canada.

Enemark, S., Clifford Bell, K., Lemmen, C., & McLaren, R. (2014). Fit-for-purpose land administration. FIG Denmark.

Craglia, M, (2017a), 5 Myths about Blockchain: Moving the conversation past current confusion so we can bring the best technology to land and property registries (April 14), New America Foundation, bit.ly/FPR-5-BC-Myths (accessed Feb 13 2018)

Graglia, M., (2017b), Will Blockchain Work in Ukraine? (June 15) New America Foundation, bit.ly/FPRBCUA1 (accessed Feb 13 2018)

Graglia, M., and Mellon, C., (2017), Pre-requisities for Incorporating Blockchain into a registry (July 31), New America Foundation, bit.ly/FPRprereq (accessed Feb 13 2018)

Graglia, M., and Mellon, C., (2018), Blockchain, Real Estate, and Land Governance, In Proceedings of the Annual World Bank Conference on Land and Poverty, March 19-23, Washington DC, United States.

Griggs, L. D., Thomas, R., Low, R., & Scheibner, J. (2017). Blockchains, trust and land administration: The return of historical provenance. *Property Law Review*, *6*, 179-194. (abstract only)

Harris, R., (2016), More data will be created in 2017 than the previous 5,000 years of humanity, App Developer Magaine, 23/12/2016 https://appdevelopermagazine.com/4773/2016/12/23/more-data-will-be-created-in-2017-than-the-previous-5,000-years-of-humanity-/ (accessed 05/02/2018)

Henssen, J. (2010). Land registration and cadastre systems: principles and related issues. Lecture Notes. Technische Universität München, Germany

Hagan, S., (2014), Land Administration and Management: Big Data, Fast Data, Semantics, Graph Databases, Security, Collaboration, Open Source, Shareable Information Platform, FIG Working Week.

Hay, G., (2013), Architecture for instrument-centred land administration applications, PhD Thesis, University of Otago, Dunedin, New Zealand.

Hay, G. (2016). Cloud, Mobile and Big Data Technologies in Responsible Land Administration. In 17th Annual World Bank Conference on Land and Poverty, Washington DC.

Hay, G., & Hall, B. (2009). Architecture for an Open Source Semantic Spatio-Temporal Land Administration Application. In FIG-Working Week.

Hjelte, H., (2017), Setting Up a Real-Estate Pilot, In Proceedings of the 2017 World Bank Conference on Land and Poverty: Responsible Land Governance–Towards an Evidence-Based Approach, The World Bank, Washington, DC, USA.

Huang, Z., Chen, Y., Wan, L., & Peng, X. (2017). GeoSpark SQL: An Effective Framework Enabling Spatial Queries on Spark. ISPRS International Journal of Geo-Information, 6(9), 285.

Jazayeri, I., Rajabifard, A., & Kalantari, M. (2014). A geometric and semantic evaluation of 3D data sourcing methods for land and property information. Land Use Policy, 36, 219-230.

Kameri-Mbote, P. (2016). Kenya Land Governance Assessment Report, World Bank. 27 June.

Karki, S., Thompson, R., McDOUGALL, K., Cumerford, N., & Van Oosterom, P. (2011, November). ISO land administration domain model and LandXML in the development of digital survey plan lodgement for 3D cadastre in Australia. In 2nd International Workshop on 3D Cadastres (pp. 65-84).

Karki, S., Thompson, R., & McDougall, K. (2013). Development of validation rules to support digital lodgement of 3D cadastral plans. Computers, Environment and Urban Systems, 40, 34-45.

Kaufmann, J., & Steudler, D. (1998, July). A vision for a future cadastral system. In Working group (Vol. 1). FIG. Switzerland.

Kewell, B., Adams, R., & Parry, G. (2017). Blockchain for good?. Strategic Change, 26(5), 429-437.

Kshetri, N., (2017), "Will blockchain emerge as a tool to break the poverty chain in the Global South?." Third World Quarterly 38, no. 8 (2017): 1710-1732. Cited In: Ølnes *et al* (2017).

Kombe, C., Manyilizu, M., and Mvuma, A., (2017), Design of Land Administration and Title Registration Model Based on Blockchain Technology, Journal of Information Engineering and Applications, Vol.7, No.1, 2017

Kwanya, T. (2014, September). Big data in land records management in Kenya: a fit and viability analysis. In *International Conference on Knowledge Management in Organizations* (pp. 15-24). Springer, Cham.

Land Layby (2017). Land Layby Kenya Limited. 2017-2018 Harambee Token. Information Memorandum. Land Layby Kenya Limited.

Land Layby, (2018). Land Layby websites in Kenya and Australia. https://www.landbankingkenya.co.ke/main/what-we-do/ http://landlayby.com.au/BlockChain.aspx Sighted 9 February 2018.

Lemieux, V.L., (2016). Trusting records: is Blockchain technology the answer?. Records Management Journal, 26(2), 110-139.

Lemieux, V.L., Flores, D., Lacombe, C., (2018), Title and code: Real Estate Transaction Recording in the Blockchain in Brazil (RCPLAC-01)-Case Study 1, DOI 10.13140/RG.2.2.10569.85606

Lemmen, C., Vos, J., & Beentjes, B. (2017). Ongoing Development of Land Administration Standards. European property law journal, 6(3), 478-502.

Lemmen, C., Van Oosterom, P., & Bennett, R. (2015). The land administration domain model. Land Use Policy, 49, 535-545.

Lemmen, C. H. J., van Oosterom, P. J. M., Kalantari, M. Unger, E. M., Teo, C. H., & de Zeeuw, K. E. E. S. (2017, March). Further Standardization in Land Administration. In Proceedings of the 2017 World Bank Conference on Land and Poverty: Responsible Land Governance–Towards an Evidence-Based Approach, The World Bank, Washington, DC, USA.

Mahmoudi, H., Renn, O., Vanclay, F., Hoffmann, V., & Karami, E. (2013). A framework for combining social impact assessment and risk assessment. Environmental Impact Assessment Review, 43, 1-8.

Mannino, M. V. (2004). Database design, application development, and administration. New York: McGraw-Hill.

McDougall, K., Bennett, R., & Molen, P. V. D. (2013). The global cadastre. GIM International, 27(7), 30-34.

McCarthy, C. (2014). Does NoSQL have a place in GIS?-An open-source spatial database performance comparison with proven RDBMS, MSc Thesis. University of Edinburgh. UK.

McLoughlin, J., Kaminski, J., Sodagar, B., Khan, S., Harris, R., Arnaudo, G., & Mc Brearty, S. (2009). A strategic approach to social impact measurement of social enterprises: The SIMPLE methodology. Social Enterprise Journal, 5(2), 154-178.

Meunier, S, (2016), Blockchain technology—a very special kind of Distributed Database, Medium, Dec 29 2016. https://medium.com/@sbmeunier/blockchain-technology-a-very-special-kind-of-distributed-database-e63d00781118 (Oct 15 2017)

Miscione, G., Ziolkowski, R., Zavolokina, L., & Schwabe, G. (2017). Tribal Governance: The Business of Blockchain Authentication. Prepared for the Hawaii International Conference on System Sciences (HICSS) 2018

Molero, R., Barry, M., Hunter, A., & Shunnar, T. (2010). Flexible Database Structures for Land Records.

Morrison, A, and Sinha, S., (2016), A primer on blockchain, PWC Next in Tech, http://usblogs.pwc.com/emerging-technology/a-primer-on-blockchain-infographic/ (Oct 15 2017)

Muhindi, S. (2017). Land Ministry Will Spend Sc17 billion to Digitise land Registration Systems. The Star. Nov 17. https://www.the-star.co.ke/news/2017/11/17/lands-ministry-will-spend-sh17-billion-to-digitise-land-registration_c1670885. Sighted 10 February 2018.

MyGov, (2016). Good News for Kenyans as Land Search Rent Payment Goes Online. http://www.mygov.go.ke/good-news-for-kenyans-as-land-search-rent-payment-goes-online/ . 22 January 2016. Sighted 10 February 2018

Nayak A., Poriya A., and Poojary D., 2013, "Type of NOSQL Databases and its Comparison with Relational Databases", International Journal of Applied Information Systems (IJAIS), New York, USA, ISSN : 2249-0868.

Nofer, M., Gomber, P., Hinz, O., & Schiereck, D. (2017). Blockchain. Business & Information Systems Engineering, 59(3), 183-187.

Oberdorf, V., (2017), Building Blocks for land administration: The potential impact of Blockchain based land administration platforms in Ghana, MSc Thesis, Utrecht.

Ølnes, S., Ubacht, J., & Janssen, M. (2017). Blockchain in government: Benefits and implications of distributed ledger technology for information sharing. Government Information Quarterly. Vol 34. Issue 3 (September 2017).

Pisa, M., and Juden. M., 2017. "Blockchain and Economic Development: Hype vs. Reality." CGD Policy Paper. Washington, DC: Center for Global Development. https://www.cgdev.org/publication/blockchain-and-economic-development-hype-vs-reality

Pouliot, J., Wang, C., Hubert, F., Ellul, C., & Rajabifard, A. R. (2016, October). 3D Cadastre visualization and dissemination: Most recent progresses and future directions. In Proceedings of the 5th International FIG Workshop on 3D Cadastres (pp. 337-359). FIG.

Sabri, S., Rajabifard, A., Ho, S., Namazi-Rad, M. R., & Pettit, C. (2015). Alternative planning and land administration for future smart cities [leading edge]. IEEE Technology and Society Magazine, 34(4), 33-73.

Savu, I., Carutasu, G., Popa, C. L., & Cotet, C. E. (2017). Quality Assurance Framework for New Property Development: A Decentralized Blockchain Solution for the Smart Cities of the Future. Res. & Sci. Today, 13, 197.

Scott, B., Loonam, J., & Kumar, V. (2017). Exploring the rise of blockchain technology: Towards distributed collaborative organizations. Strategic Change, 26(5), 423-428.

Shojaei, D., Rajabifard, A., Kalantari, M., Bishop, I. D., & Aien, A. (2015). Design and development of a web-based 3D cadastral visualisation prototype. International Journal of Digital Earth, 8(7), 538-557.

Simbizi, M. C. D., Bennett, R. M., & Zevenbergen, J. (2014). Land tenure security: Revisiting and refining the concept for Sub-Saharan Africa's rural poor. *Land use policy*, *36*, 231-238.

Snäll, M., (2017), Blockchain and Land Register - A New "Trust Machine"? Annual Work Bank Conference on Land and Poverty, March 20-24, Washington D.C., United States.

Stonebraker, M., Madden, S., Abadi, D. J., Harizopoulos, S., Hachem, N., & Helland, P. (2007, September). The end of an architectural era:(it's time for a complete rewrite). In Proceedings of the 33rd international conference on Very large data bases (pp. 1150-1160). VLDB Endowment.

Spielman, A. (2016). Blockchain: digitally rebuilding the real estate industry (Doctoral dissertation, Massachusetts Institute of Technology).
UN-HABITAT, G. L. T. N. (2012). Designing a Land Records Systems for the Poor. Nairobi.

Strauch C., 2009, "NoSQL Databases", Stuttgart Media University, Germany.

Ugrekhelidze, P., and Grigoli, E., (2017), Using Blockchain in Georgia, In Proceedings of the 2017 World Bank Conference on Land and Poverty: Responsible Land Governance–Towards an Evidence-Based Approach, The World Bank, Washington, DC, USA.

UN Habitat / GLTN, (2008), Secure land rights for all, UN Habitat, Nairobi, Kenya

UN-Habitat / GLTN, (2012). Handling Land–Innovative Tools for Land Governance and Secure Tenure. Nairobi. Kenya

van Bochove, D., Bruin, L., & Lemmen, C. (2016). From bitcoins to bitsquares. GIM International, 30(9), 36-37.

van Oosterom, P., Lemmen, C., Ingvarsson, T., van der Molen, P., Ploeger, H., Quak, W., ... & Zevenbergen, J. (2006). The core cadastral domain model. Computers, Environment and Urban Systems, 30(5), 627-660.

van Oosterom, P., Groothedde, A., Lemmen, C., van der Molen, P., & Uitermark, H. (2009). Land Administration as a Cornerstone in the Global Spatial Information Infrastructure. IJSDIR, 4, 298-331.
Palmer, D., Fricska, S., & Wehrmann, B. (2009). Towards improved land governance. UN-HABITAT.

van Oosterom, P., & Lemmen, C. (2015). The land administration domain model (LADM): motivation, standardisation, application and further development. Land use policy, 49, 527-534.

Velpuri, M. (2017). Joining the blockchain gang. Land Journal, 8.

Velpuri, M., Pidugu, A., Velpuri, J., Bhamidipati, S., Sharma M., Maringanti, C., (2017), Enabling Formalising Of Informal Markets Through Block chain For Unregistered Real estate, In Proceedings of the 2017 World Bank Conference on Land and Poverty: Responsible Land Governance–Towards an Evidence-Based Approach, The World Bank, Washington, DC, USA.

Veuger, J. (2017). Attention to Disruption and Blockchain Creates a Viable Real Estate Economy. Journal of US-China Public Administration, 14(5), 263-285.

Višnjevac, N., Mihajlović, R., Šoškić, M., Cvijetinović, Z., Bajat, B., (2017) Using NoSQL Databases in the 3D Cadastre Domain, Journal of Geodetic Engineering (Slovenia),
DOI: 10.15292//geodetski-vestnik.2017.03.412-426

Vos, J., Lemmen, C., Beentjes, B., (2017), Blockchain-Based Land Administration, feasible, illusory or a panacea? In Proceedings of the 2017 World Bank Conference on Land and Poverty: Responsible Land Governance–Towards an Evidence-Based Approach, The World Bank, Washington, DC, USA.

Wallace, J., & Williamson, I. (2006). Building land markets. Land use policy, 23(2), 123-135.

Williamson, I., Enemark, S., Wallace, J., & Rajabifard, A. (2010). *Land administration for sustainable development* (p. 487). Redlands, CA: ESRI Press Academic.

Williamson, I., Rajabifard, A., Kalantari, M., & Wallace, J. (2012, November). AAA land information: accurate, assured and authoritative. In *Proceedings of 8th FIG Regional Conference, Montevideo, Uruguay* (pp. 26-29).

Zevenbergen, J., Augustinus, C., Antonio, D., & Bennett, R. (2013). Pro-poor land administration: Principles for recording the land rights of the underrepresented. *Land use policy, 31*, 595-604.

Zevenbergen, J.A. (2017) Promising tech advances for tenure security? : powerpoint. Presented at: 2017 Conference on land policy in Africa, 14-17 November 2017, Addis Ababa, Ethiopia. 33 slides

Zoomers, A., Gekker, A., & Schäfer, M. T. (2016). Between two hypes: Will “big data” help unravel blind spots in understanding the “global land rush?”. *Geoforum*, *69*, 147-159.