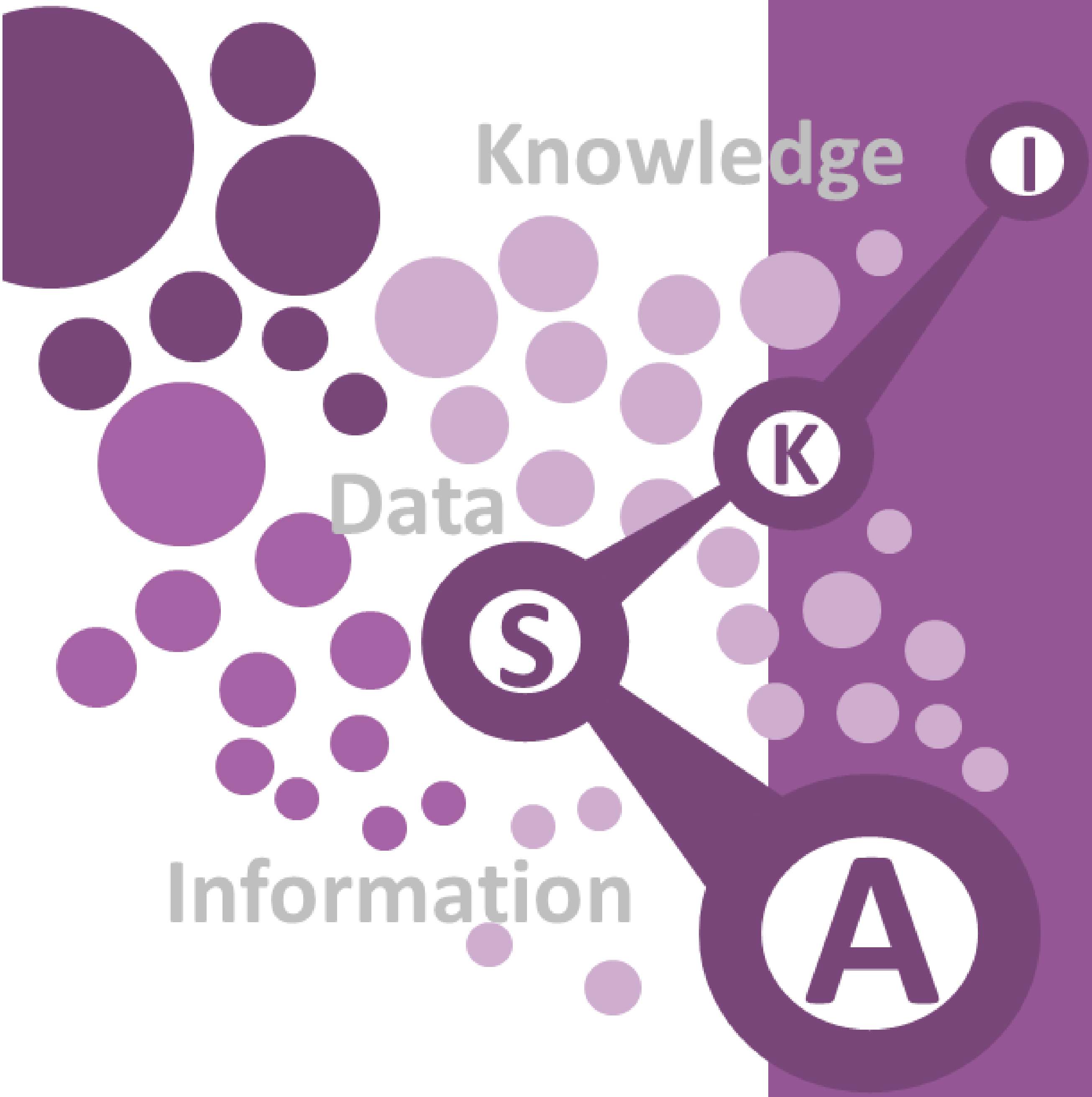
Spatial Knowledge Infrastructure:

Research Agenda and Cadastral Data Case Study



Australia and New Zealand
Cooperative Research Centre
for Spatial Information

Production Notes:

All currency included in this white paper is Australian dollars unless stated otherwise.

Authors:

Dr Lesley Arnold, Curtin University

Kylie Armstrong, Program Director Next Generation Spatial Infrastructures (CRCSI)

Reviewers:

Dr Russell Priebbenow, Director, Cadastral and Geodetic Services, DNRM, QLD

Dr David McMeekin, Curtin University

Emeritus Prof Geoff West, Curtin University

Professor Matthew Duckham, RMIT University

Serryn Eagleson –Associate Research Director, CRCSI

Darren Mottolini –Business Development and Research Manager, CRCSI

Maurits Van Der Vlugt – Director Mercury Solutions

Peter Woodgate – Former CEO, CRCSI

Australian Government
Department of Industry, Innovation and Science

Business
Cooperative Research Centres Programme

crc•si

Contents

Executive Summary

The CRCSI[1] has proposed a next generation Spatial Knowledge Infrastructure (SKI) model that will allow traditional Spatial Data Infrastructures to transform to support on demand knowledge creation and seamless support for federated supply chains. The modernisation of land and property information as described in Cadastre 2034 [1] has introduced exciting new opportunities to enable more dynamic and information intensive decisions to support our emerging digital economy and the exponential rise of spatially aware and equipped citizens. Looking through the modern Cadastral lens, this report describes how the SKI model can be used to create a paradigm shift in the delivery of whole of landscape federated supply chain by leveraging a network of data, analytics, expertise and policies in a way that allows people and machines, whether individually or in collaboration, to integrate in real-time spatial knowledge to improve decision-making and problem solving.

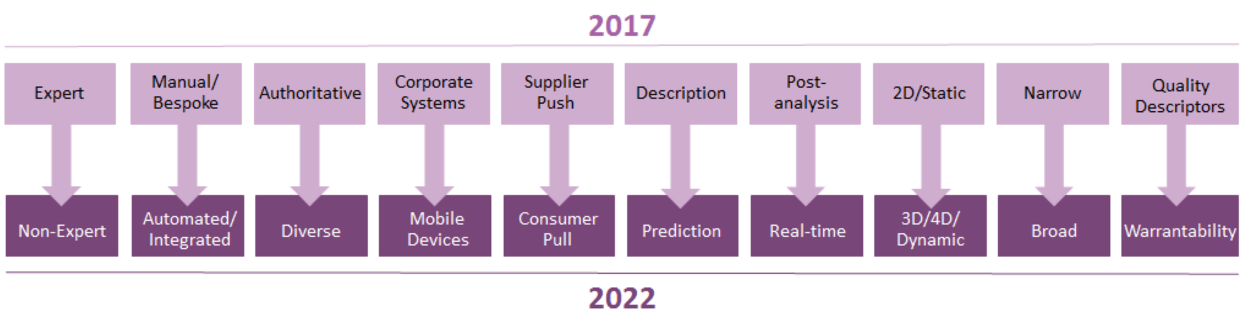
The SKI white paper, published in March 2017 [2], outlined the value proposition that justifies the move to develop and embrace the next generation SKI concepts to capture the power of emerging technologies and changing demands of users. It set out the capabilities and types of innovation and new practices that industry will transition to over the next five years.

In doing so, government, business and community sectors will be able to analyse information through a global network of data to exploit the unique properties that new knowledge brings – be it a competitive advantage, delivery of new products, faster services or simply the ability to make sound decisions from having access to new insights.

Leveraging global trends that are transforming information management in all industries, the CRCSI has combined the capabilities of these trends with emerging spatial research to summarize the changes required to move to a knowledge-based environment for customised and real-time decision-making (as shown in the diagram below). The characteristics of the digital information world, as speculated in the SKI paper 2022 [2] have been summarised in the diagram below.

This paper provides a working example of how the SKI concepts can transform existing practices and allow them to drive new knowledge-based activities. The cadastral supply chain is used as a prime example as it is fundamental to the orderly functioning of society, underpinning the all-important property and development sector, and is consistent with the clear future vision captured in Cadastre 2034 [1].

This case study compares the current supply chain activities with the future vision of a seamless, federated national cadastre that extends the digital cadastre to include linkages to all Rights, Responsibilities and Restrictions (RRR) within an accurate 3D/4D digital environment, and charts the changes in the context of an SKI.



[1] CRCSI will transition to FrontierSI 1 July 2018.

1. Introduction

The Cooperative Research Centre for Spatial Information (CRCSI) conceptualised a Spatial Knowledge Infrastructure (SKI) model that moves the agenda from more traditional Spatial Data Infrastructure (SDI) concepts, to automatically creating, sharing, curating, delivering and using knowledge (and not just data and information) in support of the digital economy and the rise of spatially aware and equipped citizens [1].

Just how the SKI concepts can be delivered and why an SKI is necessary, are explored in this second research paper that sets out the key differences between supply chains in current SDI models and those envisaged in the future SKI. The transition to a modern digital cadastre as described in Cadastre 2034 [1] is used to examine the need for change. This paper explains the necessary changes along with the research and development required to streamline data supply, improve information value and increase knowledge utility.

1.1 Looking to the Future

We have entered a new era of the 'knowledge society' where raw data is commonly processed to create information that can be analysed and the results used to make decisions and take effective action [3].

The richness of second generation information technologies (Web 2.0), has meant that the creation and sharing of information takes place at a much faster rate and on a far more participative scale than ever before [4].

Devices capable of delivering information (and not just data) now proliferate the market [5]. Today, personal assistants are available on most mobile devices [6], meaning people can connect 'anywhere' to a network of information simply by posing a question using natural language.

However, the capacity for individuals (and society) to acquire knowledge 'at will' from an integrated global resource network is not quite there yet. Personal assistants have not progressed beyond basic tasks – answering questions about the weather, sport scores and calendar appointments [6], and search and query engines currently suffer from unstructured data and the lack of spatial analytics for complex query computations [7].

Information alone is not enough to create new knowledge. Our questions are often multifaceted, and answers require spatial analysis using more than one data source. For example….Where do we locate the new hospital?.... Which areas should be declared fire risk zones?.... Where should I evacuate to in the event of a flood ?.... What are the main concerns of my constituents?... Should we insure this property?...and so on.

Finding answers to these questions is complex. SDIs today lack the critical analytical capabilities required to automatically infer new knowledge from information. Human intervention is frequently required to curate and interpret information.

In addition, the value of answers, as new knowledge, is dependent on how much a person *believes* the information to be true and has the conviction to act on the knowledge they acquire [8]. The ability to communicate the reliability of information requires data provenance (metadata plus lineage) that in today's SDIs is often unknown, and if known, is not in context of a user's questions.

A new approach to knowledge creation is required – one where knowledge can be easily extracted from data, particularly in times of crisis where immediacy and reliability are paramount. These new methods need to enable real-time, consistent and richer information discovery to meet society's quest for *knowledge on-demand*.

'New knowledge is the most valuable commodity on earth. The more truth we have to work with, the richer we become."

- Kurt Vonnegut

1.2 Semantic Web Potential

Sourcing information on demand and in a way that it is fit for use, is a key motivator behind the next generation Spatial Knowledge Infrastructure (SKI).

The proposed SKI is not a centralised service nor a distinct entity. The SKI concept embodies the behaviour of resources available via the Web of Data. The Web of Data, also referred to as the Semantic Web or Web 3.0, has potential to move information discovery and knowledge creation forward [9].

The transitionary technologies and methods, required to move from an SDI to an SKI, coincide with a shift from Web 2.0 functionality to enhanced Semantic Web 3.0 knowledge capabilities, and potentially beyond to the anticipated Web 4.0 that will enable ultra-intelligent social connectivity (Figure 1).

> **The Semantic Web will reshape how we view current SDIs – moving from a closed framework for data sharing that is centred on *data distribution portals*, to a virtual connected open Web framework that enables *knowledge* creation on-demand.**

New query capabilities to find and process information will stem from a combination of problem-directed spatial analytics and Artificial Intelligence (AI), such as natural language query processing, semantic queries, spatial filtering and machine learning algorithms. Knowledge representation will be enabled through domain ontologies and rules that capture the relationships between entities and concepts, and knowledge discovery will be achieved through the automatic orchestration of geoprocesses that process queries by integrating and linking data on-the-fly. The trustworthiness of answers will be enabled through data provenance traceability, and algorithms that rank, rate and communicate reliability to end users.

While many of these techniques have been applied to query systems in the past, the Semantic Web and Linked Data paradigm offers a whole new environment for innovation potential where real-time interconnectivity with Internet of Things (IoT) devices is possible. Linked Data, combined with intelligent agents, is currently being used to integrate and draw meaning from information independent of information communication technology [10]. This is leading to more intelligent search capabilities and the potential for advanced on-demand query processing underpinned by spatial analytics.

The uptake of Sematic Web technologies to-date has been slower than expected; far slower than that experienced with the transition to Web 2.0. The success of second generation Web technologies was driven by society's acceptance and zeal for uploading content en masse. As a consequence, Web 2.0 techniques were adopted quickly.

In contrast, the transition to the Semantic Web has not achieved immediate benefits for society. Knowledge on-demand capabilities requires data producers to manage and publish their spatial data differently. Many producers have been slow to adopt Semantic Web formats, preferring to focus on organisational productivity improvements such as federation of information, automated workflows, and the transition to cloud based platforms. These driver are far more appealing than the content-sharing user-driven benefits resulting from the transition to Web 2.0. Consequently, change is occurring at a much slower rate.

To embrace Semantic Web concepts, and thus the SKI, spatial information needs to be available as Linked Data [11]. Linked Data is the key to integrating heterogeneous data and making it machine-readable and thus, more discoverable and available for on-the-fly data querying.

Persistence is a key element to ensuring users are able to connect to information on a continuous basis. In the future, assigning (also referred to as URI Minting) Uniform Resource Identifiers (URIs) to data, investing in resources to link data to other data, and repairing broken links should they arise [11] will become the new norm. In this future, organisations will benefit from the ability to integrate data from a variety of other information resources previously not possible; opening up the opportunity for future *knowledge on-demand* products and services.

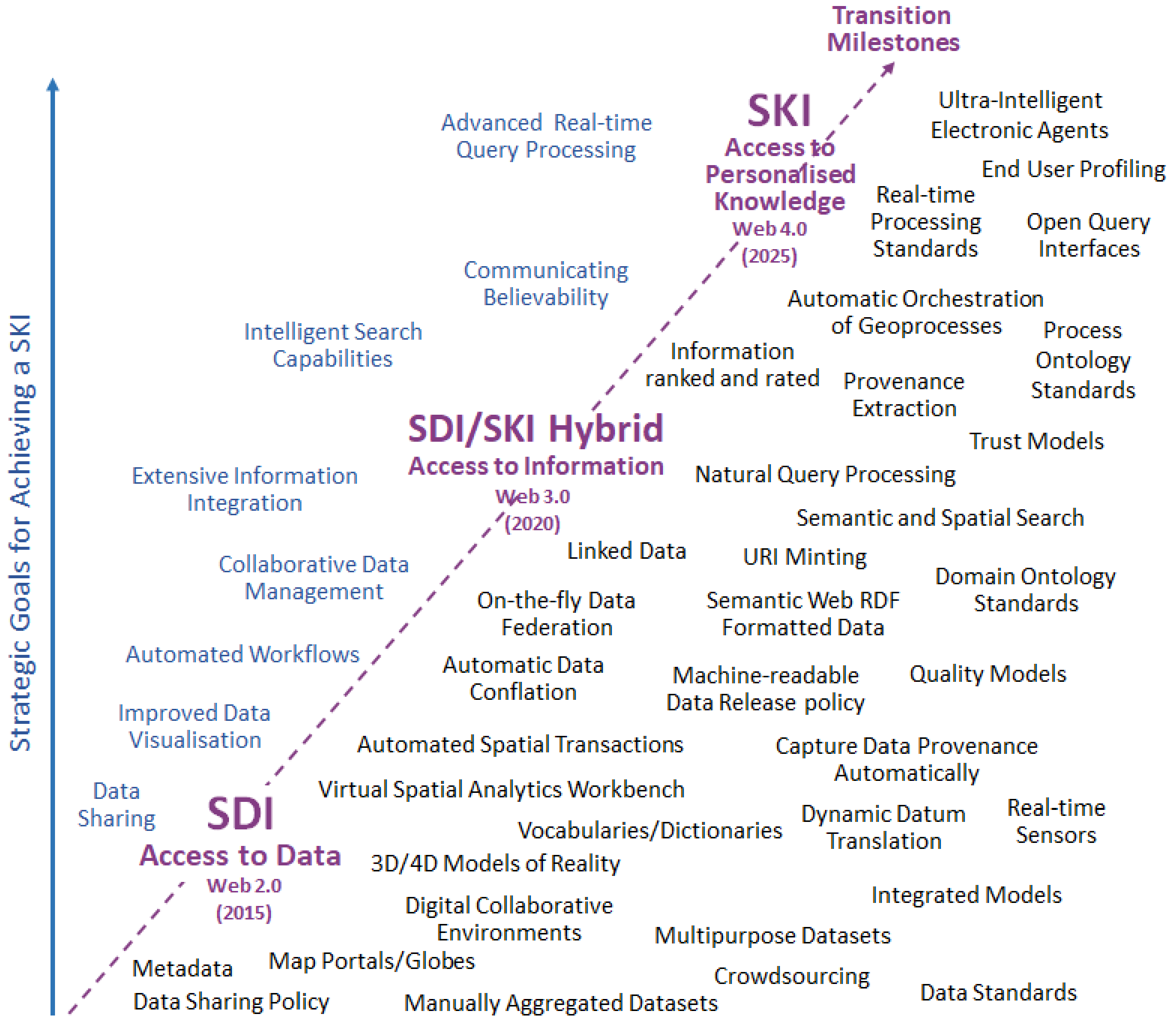


Figure 1: Transitioning to a Spatial Knowledge Infrastructure. Adapted after [9]

1.3 Transitionary Methods

Figure 1 depicts the change in technologies that will enable the transition towards improved knowledge creation; where technology, policies, supply chain collaboration and community partnerships will shape and tackle;

1. *Improvements to Data Supply* – to overcome the multifaceted limitations of current spatial data supply chains;
2. *Improvements to Information Value* – to automatically create new interpreted information from basic data using sophisticated spatial analytics and enhanced visualisations; and
3. *Increased Knowledge Utility* – to advance data querying and knowledge inferencing so that end-user questions can be responded to in real-time, reliably and in a way that supports their specific needs.

The application of these technologies along the spatial data supply chain is discussed in more detail in the Sections 7 to 9.

2. Digital Cadastre - SKI Context

The ability to draw knowledge from integrated *land and property data* is encapsulated in the vision for the future federated cadastral system that envisages being able to readily and confidently identify the location and extent of all rights, restrictions and responsibilities (RRR) pertaining to land and real property [2].

In this case study, the digital cadastre is examined in the context of what research is required to transform cadastral systems from an SDI delivery model to an SKI. The digital cadastre forms the basis for the majority of location-based applications across Australia [12], it is managed by government at a state and territory level and nationally delivered through PSMA Australia and organised within the Foundation Spatial Data Framework (FSDF).

The cadastral system, in combination with the land registration system, is used to define and reinforce property rights across Australia and underpins land and property assets worth over $5.2 trillion, nationally [2].

The growth of the knowledge economy is largely dependent on the quantity, quality, accessibility and usability of information and the knowledge that can be derived from it.

The better information landowners, investors and land managers have, the better use they can make of their land asset [2].

Being able to unlock the capacity of this information and improve its versatility and usability for evidenced-based decision-making, has the potential to benefit the economy in the short term through better knowledge of earning capacity on agricultural and grazing land etc., and in the longer term through the sustainable management of resources, and reduced financing risk on mortgage lending and other initiatives [2].

Today, businesses and individuals regularly profit from knowledge about land and this is fuelling the need for knowledge production technologies.

This 'Knowledge Commodity' future is likely to be one of the next business and social disrupters, particularly from the perspective of how cadastral information is leveraged with social, business, environmental and economic technologies, such as Twitter, real estate values, air quality monitors and financial systems, respectively.

> **To create this new knowledge future, the digital cadastre and other government land and property information will need to be able to be combined, contrasted and analysed with other datasets.**

Community data, social media, online encyclopaedias and other data on the Web, provide a rich source of diverse knowledge that is insightful (e.g. community opinions on the most liveable suburb) and immediate (e.g. fire report twitter feeds). When combined with this Web of open data, semantically-enabled government data has the potential to unlock new knowledge (Figure 2).

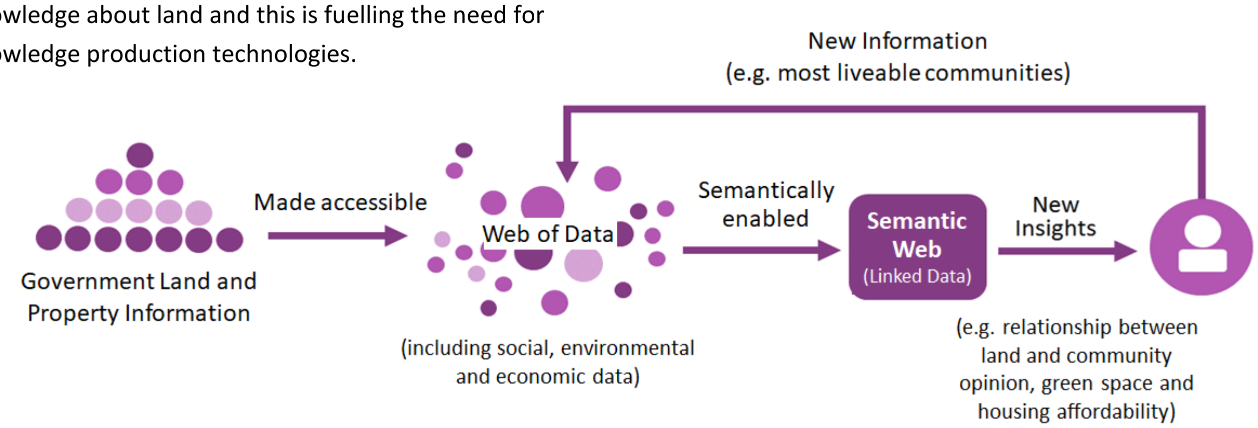


Figure 2: Semantically-enabled cadastral information contributes to the discovery of new knowledge

Currently, the integration of the digital cadastre with other data is constrained. Land and property information is locked behind government data portals, and while data is accessible to the broader public, it cannot be queried in a way that is meaningful to people who have a mix of vague, precise and implicit requirements. In addition, national access to information is even more limited due to lack of harmonisation and relevance to the land development process.

Traditional push production supply chains are no longer adequate. Current query-based systems typically cater for anticipated customer queries and data needs (usually developed through market surveys). However, application interfaces are typically hardcoded to respond to these expected queries, and they do not allow for customised data products or queries that are arbitrary in nature. Predicting demand frequency and data quality requirements is a challenge for producers. Spatial data acquisition is costly and time consuming, and vector data in particular is expensive to update. Being able to focus effort where it is needed and at the right time is critical, and a new approach to supply chain management is required.

The challenge for producers of cadastral information is to balance low demand and variability products and services with high-demand and variable end-user requirements. The balance between these competing perspectives is the point up to which demand is certain and can be forecast with relatively high accuracy and, where operational efficiencies are more feasible due to larger economies of scale, and lower setup costs and lead times (Figure 3). This point is referred to as the push/pull interface.

> **To thrive in the knowledge economy of the future, land and property information needs to be viewed in terms of its potential knowledge production capabilities. This means future development needs to focus on automating the push/pull interface in supply chains, to deliver new insights to consumers and decision-makers in a way not possible today**.

While pushing data out to end-users will remain a central role for producers of cadastral data, and delivering quality information will continue to be an important element of their supply chain value activities, producers of the cadastre will need to expose their data in Semantic Web format. Users will have the ability to pull data and create information products on the fly and take advantage of the new knowledge creation and Linked Data processes so that end-user queries can be responded to without having to specifically configure systems or modify data structures.

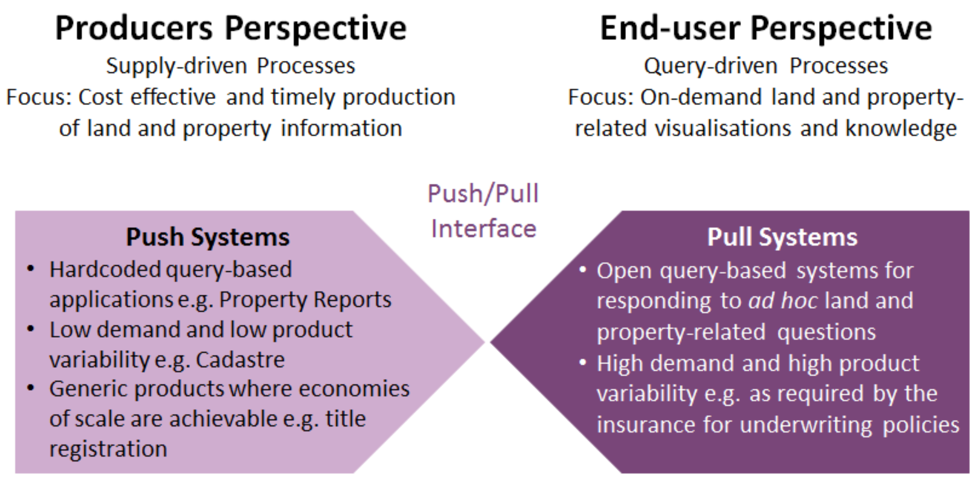


Figure 3: Characteristics of 'Push' production supply chain verses 'Pull' production 'open query-based systems' for land and property information (adapted from [13] where it is applied to manufacturing)

3. Case for Change

Two very different perspectives have emerged in knowledge management research: a commodity view which sees knowledge as something to be acquired, stored and converted, and a community perspective which emphasises knowing and the ability to act on what one knows [4].

This section considers both perspectives – the need to streamline and modernise cadastral data production as well as enabling end users to discover knowledge in real-time. Currently, cadastral data supply chains are geared towards data provision and not the creation of *knowledge* for planning and decision-making. A new approach is required – one that integrates and automates information and knowledge flows from both a producer and end-user viewpoints, respectively.

3.1 Producer's Perspective

From a data producer/supplier perspective, there is a critical need for cadastral data production to be more efficient. The demographic structure, skill sets and size of land agencies will change over the next five years, creating additional workforce pressures. New methods are required to automate and systematise workflows, and partnerships and collaborative digital environments are required to sustain the foundations of service delivery in the longer term.

Cadastral databases in many jurisdictions are currently based on out-dated technologies that lack automated processes for incorporating digital cadastral survey data. This has resulted in inefficiencies and higher than necessary production costs due to a reliance on manual processes. Existing workflows for updating the digital cadastre require that survey precision data is distorted to fit imprecise digitised land boundaries. This process constantly degrades the digital cadastre. As a consequence, applications on mobile devices, which are constantly improving in accuracy, are becoming out-of-sync with digital land boundaries.

The digital cadastre is currently an underutilised resource because of inherent data inaccuracies. Spatial inaccuracies make it difficult to integrate the digital cadastre with other interests on land, as coincident boundaries are not aligned. This impacts on the accuracy of data analysis results and the trustworthiness of the data for decision-making.

The lack of automation, from survey data acquisition through to its integration in the digital cadastre, is limiting the capacity to meet emerging business needs, such as Smart Cities that require integrated and real-time information. In addition, the precision agriculture and mobile mapping sectors require sophisticated management of the dynamic datum, and this calls for real-time transformations from field data to digital cadastre, and from digital cadastre to end user software applications, such as location-based services.

> **It is also becoming increasingly difficult to manage the continued densification and complexity of land developments (above and below ground) as data visualisations and legal instruments are currently limited to 2.5 dimensions [2].**

Cadastral data supply chains are not streamlined and duplicated datasets and data inconsistencies can occur leading to misinterpretation. Multiple copies of the digital cadastre have arisen within each jurisdiction and point of truth is not enforced. The telecommunications, utilities and local government sectors often have their own copy of the digital cadastre for internal business needs, such as managing concept plans (a.k.a. Precal), and perform upgrades to their cadastre in isolation of, and prior to, the availability of the authoritative source. Sharing upgraded data and survey plan updates is difficult as datasets become inconsistent over time and information has to be manually integrated.

Extracting knowledge from cadastral information is a key driver for reform. However, current cadastral data models are restrictive and, given the rapidly growing internet of things, a more integrated land and property model is required to augment knowledge-based systems.

In addition, new sourcing capabilities are needed to systematically integrate other land-related features to achieve a more comprehensive property model. Examples include the precise location of trees, buildings and fence lines captured by surveying firms as part of the land development process; and local knowledge such as the location of sheds, water points and private roads, which can be provided through volunteer mapping programs and crowdsourcing. These additional features can make a significant contribution to the land and property knowledge-base for emergency services, urban planning, building insurance and infrastructure risk assessment etc.

3.2 User's Perspective

There are decision-support and knowledge-based systems that use basic spatial analytics as the software core to answer land and property queries. Western Australia's Interest Enquiry [14] system, for example, enables the consumer to purchase a property report that details interests incumbent on land using a point in polygon query intersecting with several layers of land and property-related information.

Such systems are based on knowing and understanding consumer needs, and query interfaces are predefined and hardcoded. However, these interfaces are inflexible as the number of query possibilities is restricted to those that can be pre-empted.

In reality, end-user queries are unpredictable and cover a huge array of knowledge domains. The challenge is to move beyond static query systems to dynamic 'machine-learning' models that have the capacity to compute a diverse range of consumer queries.

This calls for more open query interfaces underpinned by sophisticated natural query language processing; on-the-fly spatial analytics using new machine learning techniques; Artificial Intelligence (AI); and an array of visualisation methods.

The SKI of the future is envisaged as responding to these challenges by being able to accommodate non-spatial experts and their unpredictable information needs. This will be enabled through the development of knowledge-based services that are an additional capability not currently available with current SDIs.

3.3 Moving forward

The transitionary steps for the SDI –SKI are explored in [2]. In terms of transitionary characteristics of the digital cadastre, Figure 4 illustrates where we are today and where we expect to transition to over the coming years. More detail is illustrated Figure 6.

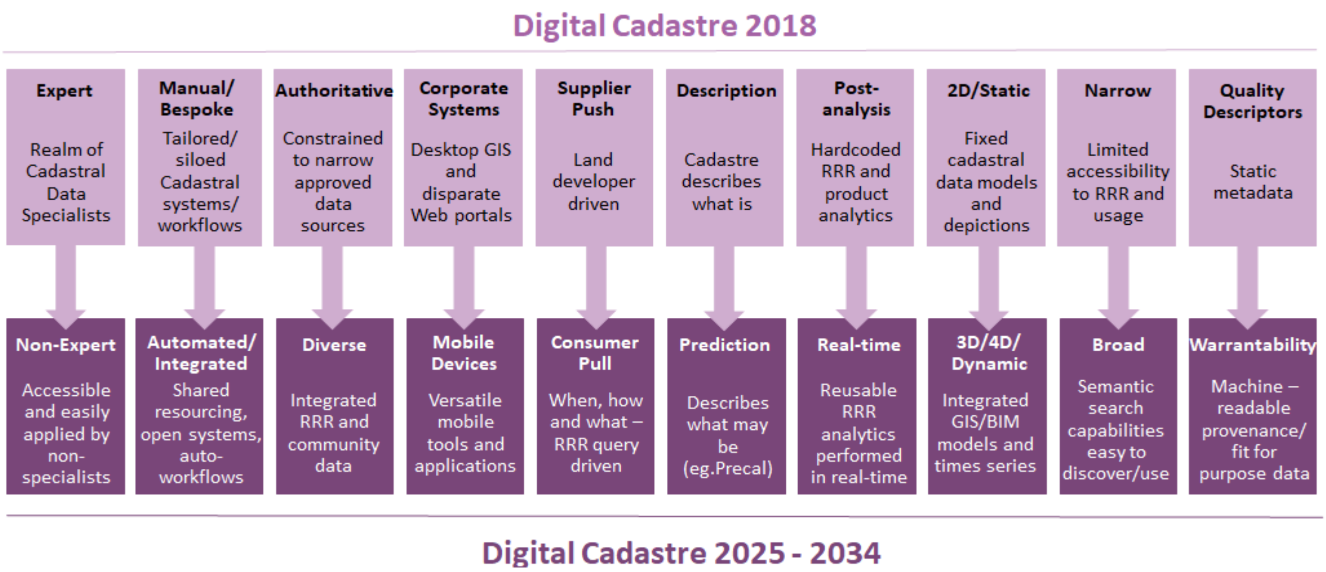


Figure 4: Transitionary characteristics of the digital cadastre anticipated over the coming years.

4. Knowledge-based Services

There are a multitude of rights, restrictions and responsibilities (RRR) pertaining to land and real property. Understanding what these interests are is a dilemma for land owners and developers, as information is spread across multiple systems of government and in various formats (both aspatial and spatial), making data integration and querying problematic.

Semantic Web tools and formats make it possible to find and integrate content (in a variety of forms) beyond the current boundaries of applications and systems [15]. This content can be interpreted more meaningfully when combined with *domain ontologies* [16].

Domain ontologies are a way to diagrammatically represent the relationships between concepts (things) and how they relate to other concepts in a given domain of knowledge. For example, an RRR Ontology represents land and real property concepts (or elements) and their relationship with other government concepts, such as policies that describe what can be done on land (rights), what cannot be done (restrictions) and what must be done (responsibilities). The Web Ontology (Markup) Language OWL is used to publish and share these relationships [17].

Domain ontologies are currently being studied under the CRCSI research program to enable on-the-fly Linked Data querying for decision makers. Figure 5 illustrates conceptually how a RRR query is executed using the Semantic Web (Web of Data). The example poses a query to discover if a piece of land is suitable as a market garden. In this example, 'land' suggests a spatial query, although the results may be provided in tabular or graph form depending on the type of query.

There is also a need to communicate the trustworthiness and relevance of the information to the user. This is where data provenance becomes fundamentally important. For derived knowledge to be of value, the end-user has to believe the information to be true or at least understand its reliability [18]. Communication tools, be they rating or ranking systems, need to convey the credibility and applicability of the query response. Methods will need to take into account the underpinning quality and lineage of the data used to process the query.

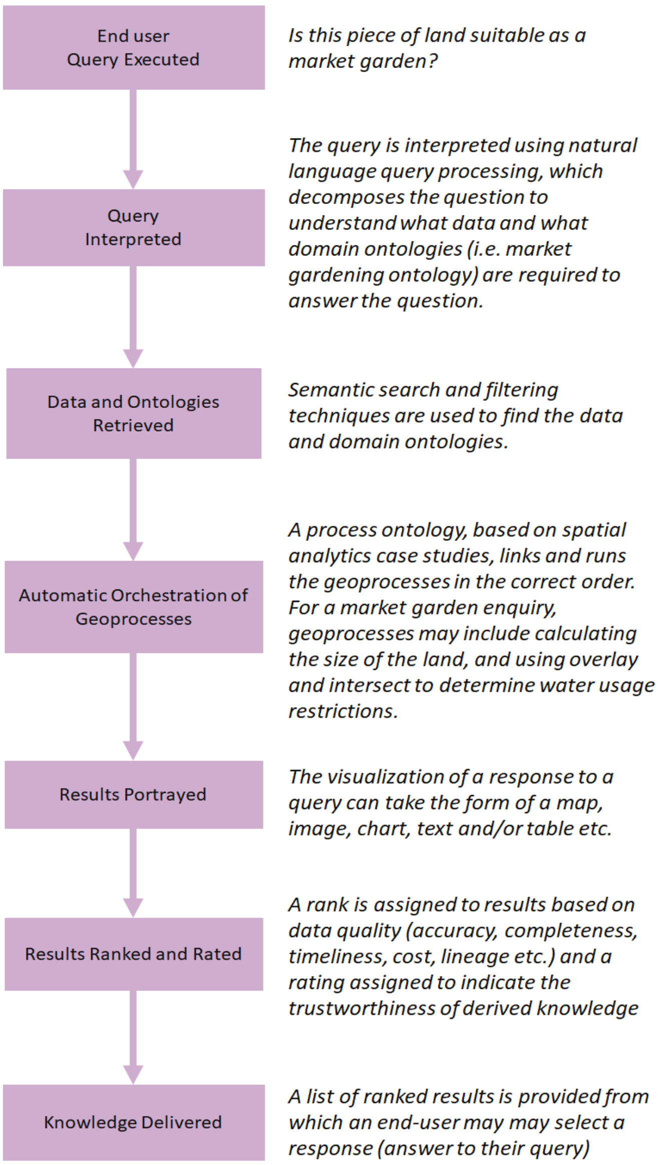


Figure 5: Conceptualised process flow for enquiring on rights restrictions and responsibilities

5. The Step Change

The four main transformative characteristics of the SDI-SKI step change relate to the policies, technologies and change in attitude towards; open 'pull-based' query systems, adoption of Linked Data formats, supply chain collaboration and cooperation, and increased trust in community data.

Firstly, it is envisaged that end-users will have the capability to query cadastral information in a way that is sensitive to their needs. This contrasts with the SDI approach where the digital cadastre is pushed out to end-users in a format and quality that is a best estimation of their needs, and where enquiry systems are predefined with a standard set of query possibilities and outputs. This new capability will be achieved through open query (search) interfaces and automated spatial analytics underpinned by Linked Data, natural language processing, and Semantic Web formatted data.

Secondly, a Global Resource Network (Web of Data) is envisaged from which data can be 'pulled' via a search or query tool to deliver application-specific capabilities. Duplicate versions will be managed using automatic conflation tools to create a single point of truth dataset that is representative of the user's needs. State and Territory datasets will be federated on-the-fly to create seamless nationwide coverage to enable regional queries to be performed. These capabilities contrast with current SDIs where multiple copies of similar cadastral datasets exist, and are logically inconsistent and create ambiguity for end-users.

Thirdly, collaboration and cooperation will be a significant characteristic of cadastral data supply chains. Information silos will gradually disappear as producers of the digital cadastre better understand their roles and responsibilities as value creation participants in a much broader supply chain. Upstream and downstream interactions will be enhanced through automated 'self-service' machine-readable spatial transactions, and the need to copy and maintain duplicate versions of the digital cadastre will no longer exist with the advent of more sophisticated property models.

And fourthly, community data will evolve as a powerful contributor to knowledge production. When combined with authoritative data, community data contributes contextual information that is far timelier, localised and specific. Provenance models for cadastral and crowdsourced data will enable information reliability to be automatically captured, sorted, tracked and managed within (or external to) authoritative data environments, and warrantability communicated to end-users. This contrasts with current SDIs, which rely on dataset level metadata and manual validation and, as a consequence, crowdsourced information is rarely used today in authoritative environments.

Transformative characteristics of the SDI-SKI step change are the polices, technologies and change in attitude towards open 'pull-based' query systems, adoption of Linked Data, supply chain collaboration and cooperation, and increased trust in community data.

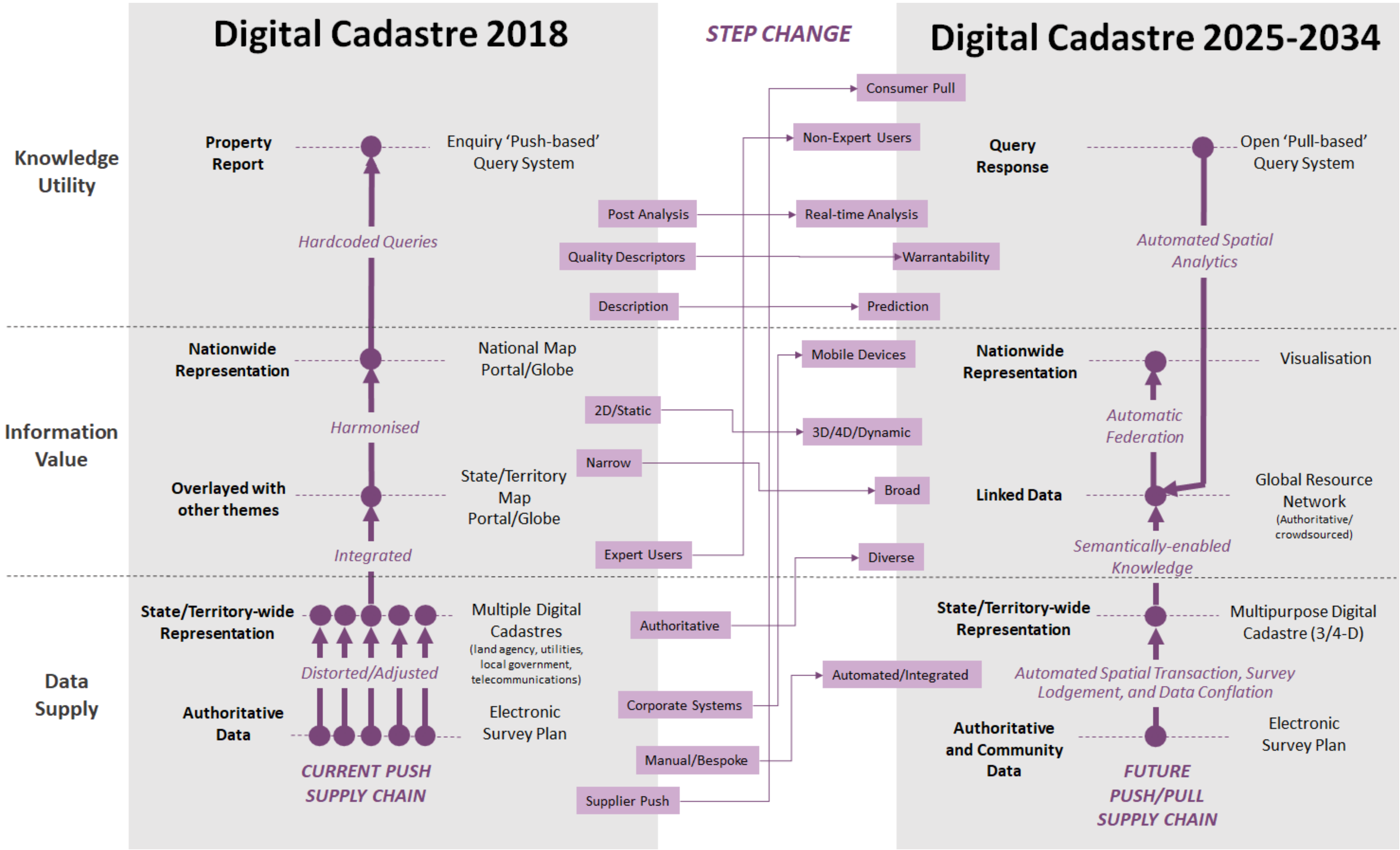


Figure 6: The step change required to move from current 'push' spatial data supply chains to the 'push/pull' supply chains that will characterise the next generation Spatial Knowledge Infrastructure

The ten key differences between the digital cadastre in the current SDI and the digital cadastre envisaged in the future SKI are illustrated in Figure 6.

The changes that can be expected over the next five years will move cadastral information from:

1. an environment where data collection and usage is predominantly the realm of cadastral data specialist using large corporate systems *to* more accessible **mobile** capabilities where **non-experts** and other domain experts are the dominant users;
2. manual and bespoke workflows and duplicate data management environments *to* more **integrated** environments that take advantage of **shared** resourcing regimes;
3. a narrow range of authoritative cadastral data sources *to* a **diverse** environment of land and property information where a variety of government data is available to be networked together, and **crowdsourced** data is routinely used as a reliable source of information in **combination** with the authoritative digital cadastre;
4. large corporate systems *to* more mobile tools and applications making cadastral information more **accessible**;
5. a supply dominated cadastral information publishing environment *to* one that also allows the end-user to **pull** information at will;
6. cadastral information that is used mainly as a method of description (what is) to one that is easily combined and linked with other data and used to **predict** what may be;
7. post-analytics that are purpose-specific *to* **open** and **reusable** analytics that are performed in **real-time** using land and property data;
8. 2D/Static cadastral data models and depictions ***to*** **3 and 4 dimensional** representations and **dynamic** visualisations that enhance land administration, planning and prediction;
9. inadequate data discoverability and narrow usage *to* **intelligent semantic** search capabilities and broader usage of the cadastre and related elements; and
10. cadastral data quality descriptors that are largely static metadata *to* more comprehensible **machine-readable** systems that communicate **warrantability** and fit for purpose to the user.

6. Making the Transition: The Research Agenda

The transition to a next generation Spatial Knowledge Infrastructure will require innovation and research in the areas of data supply, information value and knowledge utility in order to meet future demands and challenges.

This transition will not occur overnight. Nevertheless, while research is developing new methods, there are activities that producers of the digital cadastre can undertake now to lay down the foundations for more streamlined supply of data, and prepare for the Semantic Web future. These actions are listed in Table 1 and explained in Sections 7 to 9. Figure 7 illustrates where these actions and research activities occur along the 'push-pull' supply chain.

In terms of **improving data supply**, those actions worth considering now include:

- Automated data capture and validation to improve collection at point of data entry including automated collection of data provenance and Block Chain metadata.
- Collaborative supply chain management achieved through a strategic and integrated framework that manages supply chain complexity and coordinates activities.
- An integrated Land and Property Model suited to multiple cross-sector business needs that can be collectively maintained.

In terms of **improving information value**, activities that can be undertaken now include:

- Creating and sharing integrated information as Linked Data and developing machine-readable domain ontologies to share a common understanding of information concepts and relationships in the Cadastral Domain, such as Rights, Restrictions and Responsibilities.
- Fostering more diverse representations through 3-dimensional information modelling and dynamic visualisations for time-series analysis and movement analytics.
- Achieving information release readiness through spatial accuracy improvements, dynamic datum transformations, exposing linkable data and ontologies to manage data release rules automatically.

Meanwhile, research is focussing on the 'pull' supply chain to **improve knowledge utility.** Activities underway include:

- Query Execution through open interfaces to answer arbitrary end-user questions, and the ability to better understand the needs of the person posing the query through end-user profiling algorithms.
- Query processing to automatically decipher and process queries in the context of the end user using natural language processing to decompose semantic queries, spatial filtering to locate the right data, and process ontologies to orchestrate geoprocessing workflows, so that answers can be obtained from multiple sources as required.
- The ability to communicate the reliability and suitability of query responses using ranking and rating systems.

In moving to a SKI there are activities that producers of the digital cadastre can undertake now to lay down the foundations for streamlined supply of data and prepare for the Semantic Web future.

Goal	Strategy	Targeted Actions and Research
Improved Data Supply	Automated Data Capture and Validation	Automated Spatial Transactions Automated Data Verification and Trust Automated Capture of Data Provenance (e.g. integration of Block Chain metadata)
	Collaborative Supply Chain Management	Supply Chain Strategy Integrated Supply Chain Management Framework Quality Models
	Multipurpose Data	Integrated Land and Property Model Cooperative Data Management and Maintenance (and automated sharing of metadata and provenance) Machine-readable Intellectual Property Management
Improved Information Value	Creating and Sharing Information	Domain Ontologies Schema Ontologies Sematic Web formatted Data
	Diverse Representations	End user profiling and machine learning 3 and 4 Dimensional Representations
	Information Release Readiness	Spatial Accuracy Dynamic Datum Transformations Linkable Data Data Release Ontologies
Improved Knowledge Utility	Query Execution	Open Query Interfaces (extending simple search and transitioning from a catalogue view to information retrieval) End-user Computer Profiling (including machine-learning)
	Query Processing	Natural Query Language Decomposition Semantic Queries Spatial Filtering Process Orchestration and Process Ontologies Spatial Analytics Case Studies Data Conflation and Federation
	Communicating Relevance and Reliability	Ranking and Rating Systems (including warrantability)

Table 1: Goals, strategies, and targeted actions and research.

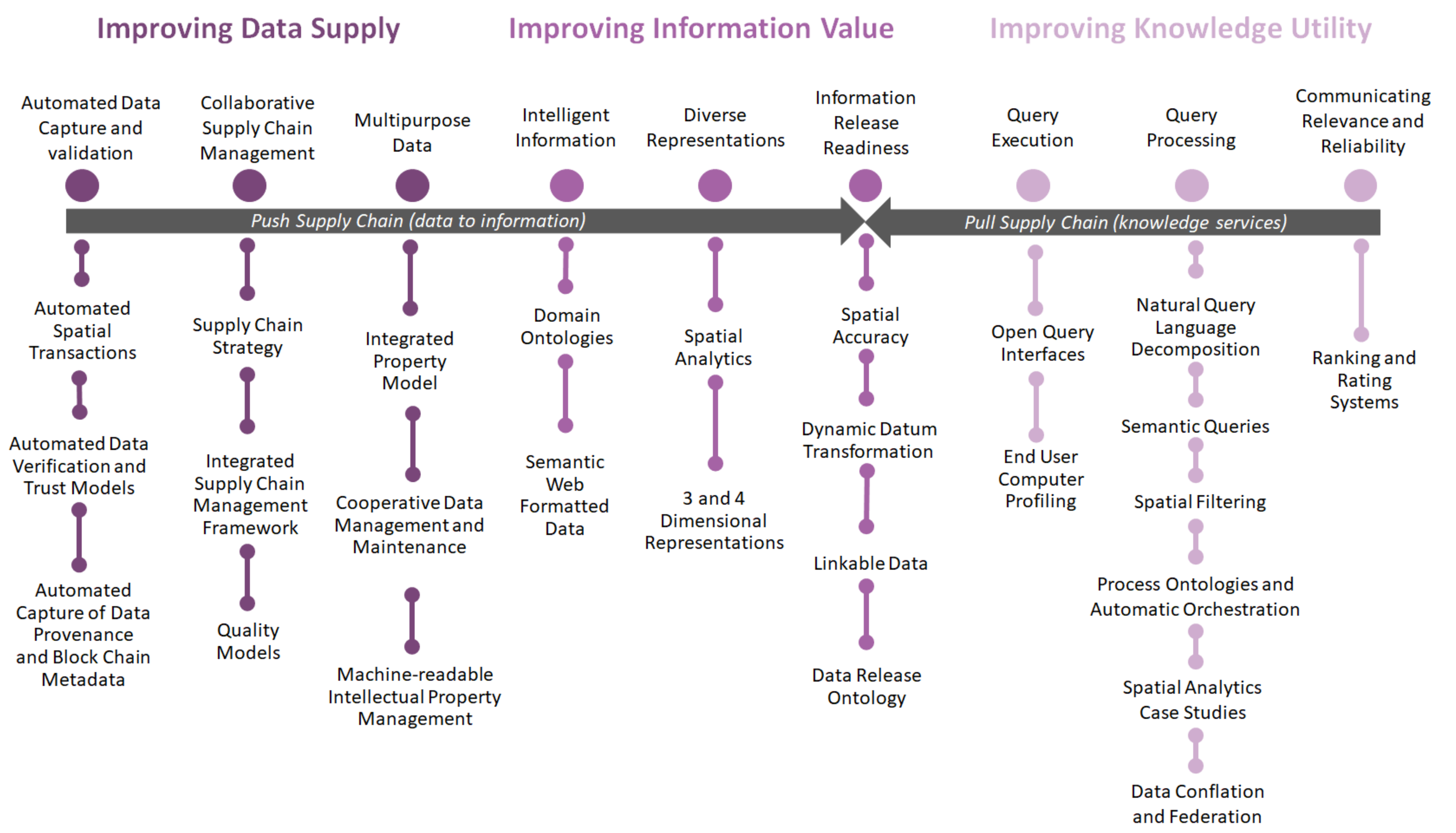


Figure 7: Targeted actions and research along the Push-Pull Spatial Data Supply Chain

7. Improving Data Supply

Improving data supply aims to automate and streamline data flows, better manage data inconsistencies, remove duplication, and transform data to enable interconnected information and knowledge creation further along the supply chain.

Machine-readable concepts and automation are at the core of the supply chain strategies, data models and data management activities discussed below; all of which aim to increase the value and utility of data as it moves along the supply chain. Improving the supply chain at point of data entry and making it machine-readable will create more streamlined and timelier service delivery, and cost savings for data suppliers, data managers and ultimately, the consumer.

7.1 Data Capture and Validation

Automated Spatial Transactions

Cadastral information supply chains are initiated by the cadastral or geodetic surveyor predominantly as part of the land development process. The precise positions of land boundaries (above and below ground) are then lodged with the land agency. Similarly, new street names and property street addresses are lodged by land developers, local governments and/or surveyors for approval. These lodgement processes are currently manual and data is often double handled, and different data may be required to be lodged with different agencies/levels of government.

Automation opens up opportunities for new information to be lodged and managed as part of an integrated digital cadastre workflow. This includes new surveyed land boundaries, geodetic positions, property street addresses and road names as well as volunteered information from property owners.

Automated Spatial Transactions are currently being examined using domain ontologies in conjunction with OWL Markup languages to develop rules to automate the spatial transaction process [19].

Local government and land developers can submit new road names online. Validation rules are derived from Geographic Naming Policy and Guidelines, and by harvesting expert knowledge. Rules sit on top of, and are separate from the data, rather than embedded within the data and discrete systems. This means that the Geographic Road Naming Ontology can be shared across multiple land agencies.

Automated Data Verification and Trust Models

Crowdsourcing, community mapping and data harvesting have not been seriously considered by government as a strategic sourcing option. Hesitation stems from concerns about the accuracy of content and the potential liability if information is wrong.

Current validation methods are extremely manual. Automated validation tools are required, along with trust models and crowdsorting methods, to enable land and property features, acquired from non-traditional sources, to be validated automatically.

Automated Capture of Data Provenance

Automated collection and validation methods that record and process data in machine-readable format will enable data provenance to be captured automatically at its point of origin, along with integration of Block Chain metadata.

Research is currently investigating methods to capture data provenance at each stage in the supply chain, including process enhancements and additions that are performed by downstream supply chain participants [20]. This is important as legal traceability and reproducibility are increasingly required.

> Improving the supply chain at point of data entry and making it machine-readable will create more streamlined and timelier service delivery.

7.2 Collaboration Supply Chain Management

Supply Chain Strategy

Spatial data management often operates in a business vacuum and is not tied to an organisation's business strategy [21]. Yet, cadastral data is often integrated with other land and property information and undergoes processing by value adders further downstream. Value activities often overlap between supply chain participants creating duplicate handling. It is important to get the order of activities right. A Supply Chain Strategy is required - one that considers both the discrete and collective needs of different groups within the data supply chain, as well as the roles and responsibilities of suppliers, producers, enablers and consumers.

Integrated Supply Chain Management Framework

An Integrated Supply Chain Management Framework is required to define, measure and improve supply chain processes and value activities. The framework needs to establish an end-to-end view that applies to all entities participating in inter-organisational supply chains, particularly national data product supply chains where governance models are evolving and there is a growing competition and increasing focus on innovation as an element of supply chain strategy.

Quality Models

Quality models are necessary to systematise supply chain activities and facilitate collaboration and cooperation between partners. Quality models consider the standards, processes, parameters and sequence of value activities essential to the provision of high performing and more viable datasets. For example, quality models can be used to manage and anticipate end-user spatial accuracy requirements and priority upgrades within the cadastral environment. Capturing quality requirements early in the supply chain means that the best representation of the digital cadastre is available to all supply chain participants further downstream.

7.3 Multipurpose Data

Integrated Property Model

Achieving an integrated Land and Property Model that will service the needs of multiple businesses is much sought after [2], and a number of land agencies have considered incorporating a provisional survey plan layer for the utilities and the telecommunications sector to meet their business needs. Furthermore, property models have moved beyond traditional data elements, such as address and ownership, to include rights, restrictions and responsibilities on land. More research is required to develop an integrated property model – one that considers property, parcel and land elements and their relevance and connectivity to broader location-based services and land-related concepts.

Collaborative Data Management and Maintenance Environments

There are several barriers to achieving a fully integrated multipurpose system that can be collaboratively maintained. This includes the inability to easily rationalise existing duplicate cadastral datasets and manage competing business interests. The problem is both financial and cultural. The former because data conflation is manual and therefore costly, and the latter because each dataset has well established business-specific improvements that need to be preserved.

Research is tackling data conflation to enable the best available data to be extracted from various similar online datasets to achieve a single source of truth [22] that can be maintained collaboratively or regenerated automatically as required. Research is also enabling the automatic federation of data to avoid the need for federated post-processing and to reduce infrastructure overheads [23]. Being able to automatically share metadata and provenance in federated environments will become an essential capability.

Machine-readable Intellectual Property Management

Cooperative data management environments raise questions around the management of intellectual property (IP) associated with different data elements contained in a multipurpose cadastre. Research is examining machine-readable policy to manage datasets where intellectual property belongs to more than one agency. Rules operate at the feature-level and not just at a dataset (macro level), meaning IP and related pricing and licensing can be managed at a micro level.

8. Improving Information Value

Improving information value is aimed at external business users, often spatial data specialists, who download information for planning and analysis and further value-adding. The digital cadastre, essentially a by-product of the land administration process, has supplementary value as an aggregated and integrated information product.

With integrated information, more complex spatial analytics can be performed leading to interpreted information that can be accessed to create new knowledge further along the supply chain. In this future, Linked Data and automated spatial analytics will become the new norm, as will domain ontologies that capture relationships and meaning between features in disparate datasets. In addition, more versatile data representations are required to create higher value products and services, and stimulate innovation and further value creation downstream.

8.1 Intelligent Information

Domain Ontologies

Domain ontologies (as well as vocabularies and data dictionaries) are commonly used on the Web to represent knowledge – ranging from large taxonomies categorising Web sites to the classification of products for sale e.g. Amazon.com [24].

A domain ontology is a set of concepts (things) and categories in a subject area. Ontologies make explicit the properties and relations (rules) between concepts, such as geographic features. Ontologies can be made machine-readable. W3C has developed a technology stack that allows computers to read, query and interpret the relationships between things [17].

One of the primary aims of domain ontologies is to share a common understanding of the structure of information among people and software agents. For example, an ontology can be used by computers to extract and aggregate land and property information from multiple sites to answer an end user query or input data to another application.

Ontologies separate domain knowledge from operational knowledge, meaning that ontologies can be shared and reused in multiple business environments. For example a Rights, Restrictions and Responsibilities Domain Ontology could be applied in the insurance sector to assess exposure risk, in the real-estate sector to assist buyers, and in the local government sector to ensure land use compliance.

There is a need to develop authoritative domain ontologies early, as there is potential for different ontologies in the same domain to result from the differing perceptions of organisations. A number of land and property-related domain ontologies are likely to arise and this may pose a problem as merging ontologies down the track is complex (24).

Semantic Web Formatted Data

With the evolution of the Semantic Web, organisations will need to consider adopting the Semantic Web RDF format for cadastral information. RDF supports BIG data processing, and advanced semantic querying.

The United States Geological Survey has successfully converted point, vector and raster spatial datasets from the National Map to RDF format to support direct querying [23].

The Semantic Web RDF format enables organisations to publish structured data so that it can be read by other computers. It does this by recording statements about things and how they relate. The format is commonly referred to as *triples* and the expression is rendered in the form of subject-predicate-object, where the predicate identifies the relationship between subject and object. For example, a **building** (subject) is located **within** (predicate) a **property** (object).

8.2 Diverse Representations

Spatial Analytics

Spatial analytics is central to creating new information from existing data sources. In addition to providing new insights, this new information contributes to the broader Information Resource Network. Growth areas include automatic 'Movement Analytics' to understand the movement of people, vehicles, livestock, disease and wildlife.

Integrated land and property information and complex spatial analytics have a critical role to play in enabling diverse decision-making applications; particularly in the insurance, retail, health and emergency service sectors, where real-time data analytics are required to reduce business and social risk, forecast change and take pre-emptive action.

3 and 4 Dimensional Representations

A more realistic representation of the digital cadastre and other land and property information will directly benefit land owners, business and governments through improved land and property management, climate change studies and urban planning. Research is currently addressing 3 and 4 dimensional models of reality to better represent the extent, adjacency, height, depth, volume and usage of property over time [26, 27], and enhanced visualisation capabilities using holograms are on the cusp of revolutionising how we interact with information [29]. Printers now have 3-dimensional capabilities that are being leveraged by the construction sector. Yet, the digital cadastre and other land and property related information, continues to be managed in 2.5D environments.

Moving to 3 and 4-dimensional representations poses some significant challenges including: (a) how to migrate data from a traditional 2-D environment; (b) how to manage changes over time and view this history in 4-D; (c) how to distinguish between changes made to the data (i.e. corrections, additions, deletions) and changes made to the cadastre itself, as the history of amendments made to the database is not necessarily the same as the history of the cadastre; (d) how to portray data dynamically for mobile technologies requiring datum translations; and (e) how to maintain visual (and mathematical) alignment of data across data themes, particularly when cadastral boundary data is progressively spatially upgraded.

8.3 Information Release Readiness

Spatial Accuracy

The digital cadastre is a fusion of 'variable' positional accuracies. The cost of spatially upgrading data is currently prohibitive and more automated methods are required. Demand-driven accuracy improvements in priority areas, with a focus on relative rather than absolute accuracy, is a practical alternative in the short term.

Dynamic Datum Transformation

With a dynamic datum being introduced across Australia [30], there is a need to develop tools and services that facilitate its use by the mass-market, and be able to link back to the source and provenance of the data. Importantly, in addition to the transformation itself, the propagation of coordinates between epochs must be carried across integrated datasets. This is crucial to future data release readiness and the versatility of information.

Linked Data

There is a core of recommended practices for exposing, sharing and connecting data using the Semantic Web [11]. These practises combine to make data machine-readable and able to be linked with other data. In terms of data release readiness, key components to making data linkable are: a) Uniform Resource Locators (URIs) that uniquely identify entities or concepts in the real world, b) HTTP that provides a universal mechanism to retrieve resources, and c) the Resource Description Format (RDF) graph-based data model used to structure and link data. While URIs can be assigned automatically using Minting processes, the actual linking of data is still manual [11] and new methods are required to generate data linkages on-the-fly.

Data Release Ontology

Custodians will also require a Data Release Ontology to manage licensing, pricing and data access levels automatically. The rules for release can then form part of the SKI technology governance framework. A call on a dataset will activate the rules embedded in the ontology. For example, if a cadastral data set is classified as commercial it may be encrypted, and a fee (and/or licence) may apply. The end-user's profile would then be checked for subscription service access rights or be directed to a one-off fee payment service.

9. Improving Knowledge Utility

Improving knowledge utility is about enabling the end-user to derive knowledge –when they need it and in their context. For the producers of the digital cadastre, this means being able to support complex end-user queries, such as enquiring on rights, restrictions and responsibilities pertaining to land and real property.

9.1 Query Execution

Open Query Interfaces

Currently, end-user query capabilities are hardcoded into application interfaces or simple catalogue search functions. These are often inflexible. The number of query possibilities are limited to those that can be pre-empted as required by end-users. In contrast the SKI of the future will require a more open query interface accessible on the Web, such as the Google™ interface, which gives access to a global information resource network.

End-user Computer Profiling

Interpreting an end-user query can be augmented by end-user computer profiling on an executed query, as it takes into account the characteristics and preferences of the end user (computer user). For example, "*will a location be flooded?"* has different connotations when considered from the perspective of an emergency responder, home owner and insurer. While Web search engines and machine-learning have exploited this concept for product marketing, there is a need to better understand the purpose and likelihood for which land and property information is queried in order to target query responses more appropriately.

9.2 Query Processing

Natural Query Language Decomposition

Natural language search engines target answers to questions. This is different from conventional search engines that simply locate data elements through keyword searching and ignore the question in the user statement.

Natural language statements can be decomposed to understand the nature of the question and then search

.

and return a subset of data that potentially contains the answer to the user's question. However, as with many knowledge domains, questions relating to land and property are many and varied, and when considered with real-time sensors and imagery, fall within the realm of Big Data.

The challenge ahead is to eliminate the ambiguity in natural language [31], cater for localisations (dialects and idioms), sentiment and emotion, and deliver reduced processing times.

Semantic Queries

Semantic Web research is focusing on semantic queries and analytics to retrieve both explicitly and implicitly derived information based on the syntactic, semantic and structural information contained in data [32]. The aim is to provide contextually appropriate results (and potentially a single correct answer) to a question, or to answer fuzzy and open-ended questions through reasoning algorithms. Semantic queries operate on the Web of Data using Linked Data, rules and RDF triples, to process the relationships between information and infer answers.

Spatial Filtering

Behind the open query interface, there is a need for far more sophisticated search capabilities than are available today. Research by [33] is investigating semantic and spatial search techniques in conjunction with query decomposition and filtering techniques to retrieve the necessary information for processing a response to an end-user's query. Spatial and semantic filtering is then applied, further increasing the relevance of results by extracting locations from the search terms and looking at similar locations.

Process Ontologies and Automatic Orchestration

Current geospatial workflows, such as flood prediction, are typically manual [34]. Workflows require an operator to initiate a process, set process parameters (i.e. tolerances) and once the process is complete, initiate the next process and so on… until all processes in the workflow are complete.

Responding to a complex end user query automatically requires the chaining together of resources, such as Web services and various data sources, to form an orchestrated *query processing* workflow. For example, "What are the chances of flooding in my area?" requires a series of geoprocesses to be executed.

The actual rules that control the workflow and execute and run the geoprocesses managed using a Process Ontology. A Process Ontology captures process knowledge for a particular domain, such as flood prediction modelling [35].

Spatial Analytics Case Studies

Spatial analytics case studies are continually evolving and provide the foundation for domain and process ontologies, which are necessary for on-the-fly data querying and allowing users to link information from disparate sources.

There is still much work to be done to develop spatial analytics workflows to understand the type, order and parameters of geoprocesses to respond to a query in a specific knowledge domain. Case studies relating to the land and property domain are required to develop the necessary knowledge for automatic orchestration of processes, such as workflows for responding to queries relating to rights, restrictions and responsibilities on land and property.

Data Conflation and Federation

Some geoprocesses are still extremely manual and cannot be easily implemented on-the-fly to answer a query that requires a regional or national perspective. Data Conflation and Federation are two such methods. In many cases, local governments and State and Territory Governments wish to retain autonomy over their cadastral information and continue to have their own models, standards and methods. This makes unification of datasets problematic. To combat this problem, the federation of individual data sets on-the-fly is being investigated to enable regional and nationwide datasets to be merged with consistency [36] and on-the-fly data conflation is being studied to create a *point of truth* data set from heterogeneous data sources on the Web [22].

9.3 Communicating Relevance and Reliability

Ranking and Rating Systems

Warrantability in query processing requires data provenance to be automatically extracted during the query processing phase so that responses can be classified and communicated according to their suitability for a purpose (rated) and listed according to relevance (ranked). Rating and ranking systems need to take into account the underpinning quality and lineage of the data used to process the query. However, the computations for ranking and rating systems are yet to be devised for responses derived through spatial analytics.

Ranking and rating systems need to be able to communicate the trustworthiness of derived knowledge to the end-user. Communication tools need to be simple and readily convey the usability of information for a particular purpose – be it for decision-making, navigation, precise measurement or prediction etc. Notations, pictograms and icons are typically used to convey meaning quickly and effectively, such as stars used to rank and rate hotels. However, a communication tool is yet to be devised for query responses in the context of an SKI.

Improving knowledge utility is about enabling the end-user to derive knowledge – when they need it and in their context.

10. Conclusion

The ability to analyse information through a global network of data will allow the government, business and community sectors to exploit the unique properties that new knowledge brings – be it a competitive advantage, delivery of new products, faster services or simply the ability to make sound decisions from having access to new insights.

Cadastral information is likely to be a major contributor to decision support systems and open query knowledge systems of the future. Understanding location, ownership (including RRR) and land value is a powerful asset that has potential to advance Australia's knowledge economy

Growth in the knowledge economy will intensify with advancements in digitalisation and data analytics. However, the digital cadastre, like many spatial information themes, is not adequately positioned to deliver the virtual world necessary to develop, monitor and preserve our physical world.

The next level of capabilities in knowledge-based services are likely to stem from Semantic Web, machine-learning and Artificial Intelligence tools. The CRCSI has been exploring these emergent methods in terms of an interconnected system for knowledge discovery – referred to as the Spatial Knowledge Infrastructure.

The delivery of knowledge-based services requires revolutionary methods for computerised workflows in the spatial domain, improved usability of spatial information, and enriched spatial analytics for query processing. Combined, these capabilities will take industry to the next level of automation, and end users to the next level of knowledge on-demand.

Currently, the focus of mapping organisations is on producing and using existing content and presenting this data as new information. However, it is only when information is queried and used that it becomes knowledge. Developing mechanisms to make knowledge more accessible and reliable is paramount.

In the past, technological advancements have been hindered because knowledge discovery and data supply have traditionally been researched as mutually exclusive problems in the spatial domain. The CRCSI has tackled this duality more holistically by; a) examining aspects of 'push' supply chains and the interrelationships between value activities performed by supply chain participants; and b) developing 'pull-based' query methods to enable end users to easily extract knowledge from published data and understand its trustworthiness.

For spatial information, such as the digital cadastre, to be relevant in the future, a step change in the way information is acquired, managed, analysed and published is required.

This paper has introduced an SDI-SKI research agenda and transitional methods to encourage a strategic, systematic and coordinated approach to the step change required over the coming years.

The ability to analyse information through a global network of data will allow the government, business and community sectors to exploit the unique properties that new knowledge brings – be it a competitive advantage, delivery of new products, faster services or simply the ability to make sound decisions from having access to new insights.

The characteristics of the SDI-SKI transformation relate to the policies and technologies, and change in thinking towards the development of open 'pull-based' query systems, adoption of Linked Data formats, increased levels of supply chain collaboration and cooperation, and the implementation of mechanisms that stimulate trust in community data and derived knowledge.

The step change research agenda introduces several technologies to extend the general Web architecture to interrelate data using a common framework. For the land and property sector, a common framework means that the digital cadastre, in various formats and locations, can be integrated to deliver knowledge-based services seamlessly through an open query Web interface that uses domain-specific search engines to improve the accuracy of query responses [35].

The outcome of the SDI-SKI step change envisages an Australian Spatial Data Infrastructure as a Knowledge infrastructure, where suppliers of land and property information have improved data supply mechanisms, an increase in information value and enhanced knowledge services.

Future planning will need to consider the necessary governance frameworks with which the broader land and property information fabric will operate. The implementation of this fabric will be different – moving from paper-based policies to machine-readable rules for the validation, management and delivery of information.

The future SKI holds much promise for suppliers and managers of land and property information. The ability to share this knowledge with others is what affords value to the knowledge economy and thus economic good. Being able to link concepts, data and processes will enable the delivery of knowledge services, such as enquiring on interests on land and real property.

.

References

[1] Matt Duckham, Lesley Arnold, Kylie Armstrong, David McMeekin, Daren Mottolini (2017) Towards a Spatial Knowledge Infrastructure, White Paper, Australia and New Zealand Cooperative Research Centre for Spatial Information, [Online] Available at http://www.crcsi.com.au/spatial-knowledge-infrastructure-white-paper/, accessed June 2017

[2] ICSM (Intergovernmental Committee on Surveying and Mapping (2015) Cadastre 2034, Powering Land and Real Property: Cadastral Reform and Innovation for Australia – A National Strategy, [Online] Available at http://www.icsm.gov.au/cadastral/Cadastre2034.pdf, accessed September 2016, pp.40. [2]

[3] Naim Hamdija Afgan, and Maria G. Carvalho (2010) The Knowledge Society: A Sustainability Paradigm, Cadmus Article, October, 2010, Available at http://www.cadmusjournal.org/node/14, accessed January 2017

[4] Keloğlu-İşler, Esra İlkay ; Bayram, Özlem Gökkurt (2014) Commodification of Knowledge Communication Mediums: From Library to Social Media, Procedia - Social and Behavioural Sciences, 25 August 2014, Vol.147, pp.550-553 [Peer Reviewed Journal]

[5] CISCO (2016) Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015-2020, 3 February 2016, [Online] Available at http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.pdf, accessed Jan 2017.

[6] Eaton, K (2016) A Smarter Phone Becomes a Personal Assistant, New York Times, 19 May 2016, Page B6.

[7] Ablimit Aji, Fusheng Wang, Hoang Vo, Rubao Lee, Qiaoling Liu, Xiaodong Zhang, and Joel Saltz, (2013) Hadoop-GIS: A High Performance Spatial Data Warehousing System over MapReduce, Proceedings VLDB Endowment, August, Vol6, No 11: p1109, 2013.

[8] Voinov, A; Kolagani, N; McCall, M; Glynn, P.D; Kragt, M.E; Ostermann, F.O; Pierce, S. A and Ramu, P (2016) Modelling with stakeholders - Next generation, Environmental Modelling & Software, 77 (2016) 196-220, Elsevier Science.

[9] Peter Woodgate, Isabel Coppa, Suelynn Choy, Stuart Phin, Lesley Arnold and Matt Duckham (2017) The Australian Approach to Geospatial Capabilities; Positioning, Earth Observation, Infrastructure and Analytics: Issues, Trends and Perspectives, Volume 20, 2017 - Issue 2: Special Issue: Perspectives on the Nature of Geo-Spatial Information. International Journal of Geo-Spatial Information Science (GSIS). Published online: 22 Jun 2017; http://dx.doi.org/10.1080/10095020.2017.1325612, Guest Editor: Prof. Dr. J.L. van Genderen.

[10] Tom Heath and Christian Bizer (2011) Linked Data: Evolving the Web into a Global Data Space, Synthesis Lectures on the Semantic Web: Theory and Technology, February 2011, Vol. 1, No. 1, Pages 1-136

[11] Tim Berners-Lees (2017) 5 Start Linked Open Data, [Online] Available at http://5stardata.info/en/, accessed August 2017.

[12] ANZLIC. 2014. "The Australian and New Zealand Foundation Spatial Data Framework." April. http://anzlic.gov.au/sites/default/files/files/One_ANZ_Foundation_Spatial_Data_Framework_Booklet.pdf.

[13] Sehgal (2009) Enterprise Supply Chain Management: Integrating Best in Class Processes, John Wiley and Sons, pp 201

[14] Landgate (2011) Interest Enquiry System, [Online] Available at https://www2.landgate.wa.gov.au/web/guest/interest-enquiry, accessed Jan 2017.

[15] W3C (2016) Semantic Web, [Online] Available at http://www.w3.org/standards/semanticweb/, accessed January 2017.

[16] Ibrahim, Nora Y; Mokhtar, Sahar A. and Harb, Hany M. (2012) Towards an Ontology based integrated Framework for Semantic Web, International Journal of Computer Science and Information Security, Vol. 10, No. 9, 2012 pp 91-99.

[17] W3C (2012) OWL 2 Web Ontology Language Document Overview, 2nd Edition, December 2012, [Online] Available at http://www.w3.org/TR/owl2-overview/#Overview, Accessed January 2012.

[18] Hunt, P. Darwin (2003) The concept of knowledge and how to measure it, Journal of Intellectual Capital, Vol 4, No. 1, pp100-113

[19] Premalatha Varadharajulu ; Geoff West ; David A. McMeekin ; Simon Moncrieff and Lesley Arnold (2016) Automating Government Spatial Transactions, in proceedings of the 2nd International Conference on Geographical Information Systems Theory, Applications and Management - Volume 1: GISTAM, 157-167, 2016, Rome, Italy

[20] Sadiq, M. A., McMeekin, D. A., & Arnold. (2016). Spatial Data Supply Chain Provenance Modelling for Next Generation Spatial Infrastructures Using Semantic Web Technologies. In *Proceedings of GEOProcessing 2016 The Eighth International Conference on Advanced Geographic Information Systems, Applications, and Services* (pp. 146-153). Venice, Italy: IARIA.

[21] Arnold, L.M. (2016) Improving Spatial Data Supply Chains: Learnings from the Manufacturing Industry, GeoProcessing 2016: The Eighth International Conference on Advanced Geographic Information Systems, Applications & Services, Venice, Italy

[22] Yu F, West G, Arnold L, McMeekin DA, and Moncrieff S. (2016) *Automatic geospatial data conflation using semantic web technologies*. In Proc. Australasian Computer Science Week (ACSW), doi:10.1145/2843043.2843375

[23] Fa, Jeremy Siao Him; West, Geoff; McMeekin, David; and Moncrieff, Simon (2016) Brokered Approach to Federating Data using Semantic Web Techniques, The Eighth International Conference on Advanced Geographic Information Systems, Applications, and Services, pp 146-153, Venice, Italy.

[24] Natalya F. Noy and Deborah L. McGuinness (2001) Ontology 101: A Guide to Creating Your First Ontology, Stanford Publications, [Online] Available at http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html, accessed January 2017.

[25] E. Lynn Usery and Dalia Varanka (2012) Design and Development of Linked Data from The National Map, Journal of Sematic Web – On linked spatio-temporal data and geo-ontologies, Vol 3, Issue 4, October 2012, pp371-384

[26] Hong, Jung-Hong and Kuoa, Chiao-Ling (2015) A semi-automatic lightweight ontology bridging for the semantic integration of cross-domain geospatial information, International Journal of Geographical Information Science, 2015 Vol. 29, No. 12, 2223–2247, Taylor and Francis.

[27] Behnam Atazadeh, Mohsen Kalantari, Abbas Rajabifard, Serene Ho (2017) Modelling building ownership boundaries within BIM environment: A case study in Victoria, Australia, Vo. 61, Part A, January 2017, Pages 24-38.

[28] Behnam Atazadeh, Mohsen Kalantari, Abbas Rajabifard, Serene Ho, Tuan Ngo (2016) Building information modelling for high-rise land administration, Transactions in GIS, March 2016

[29] Bright, Peter (2016) Hands-on with HoloLens: On the Cusp of a Revolution, [Online] Available at http://arstechnica.com/gadgets/2016/04/hands-on-hololens-on-the-cusp-of-a-revolution/, accessed January 2017.

[30] Nic Donnelly, Chris Cook, Joel Haasdyk, Criagg Harrison, Chris Rizos, Craig Roberts and Richard Stanaway (2014), Progress Towards a New Datum for Australia, Eds. S. Winter and C. Rizos (Eds.): Research@Locate'14, Canberra, Australia, 07-09 April 2014, [Online] Available at http://ceur-ws.org, accessed January 2017.

[31] Anjali M K, and Babu Anto (2014) Ambiguities in Natural Language Processing, International Journal of Innovative Research in Computer and Communication Engineering, Vol.2, Special Issue 5, October 2014, pp 392-394

[32] Marcelo Arenas and Jorge Peres (2011) Querying Semantic Web Data with SPARQL, PODS'11, June 13–15, 2011, Athens, Greece.

[33] Reed, T; Gulland, E-K; West, G.A.W; McMeekin, D; Moncrieff, S; and Cox, S. (2015), GeoMeta: A modern geospatial search engine, *Proc. 11th Int. Conf. on Innovations in Information Technology*, Dubai, UAE, November.

[34] Bing Tan, Chet; West, Geoff; McMeekin, David; and Moncrieff, Simon (2016) CIAO-WPSS Automatic and Intelligent Orchestration of Geospatial Web Services using Semantic Web Ontologies, 2nd International Conference on Geographical Theory, Applications and Management (GISTAM) 26-27th April, Rome, Italy.

[35] Trillo, Raquel ; Po, Laura ; Ilarri, Sergio ; Bergamaschi, Sonia ; Mena, Eduardo (2011) Using semantic techniques to access web data¸ Information Systems, 2011, Vol.36(2), pp.117-133.