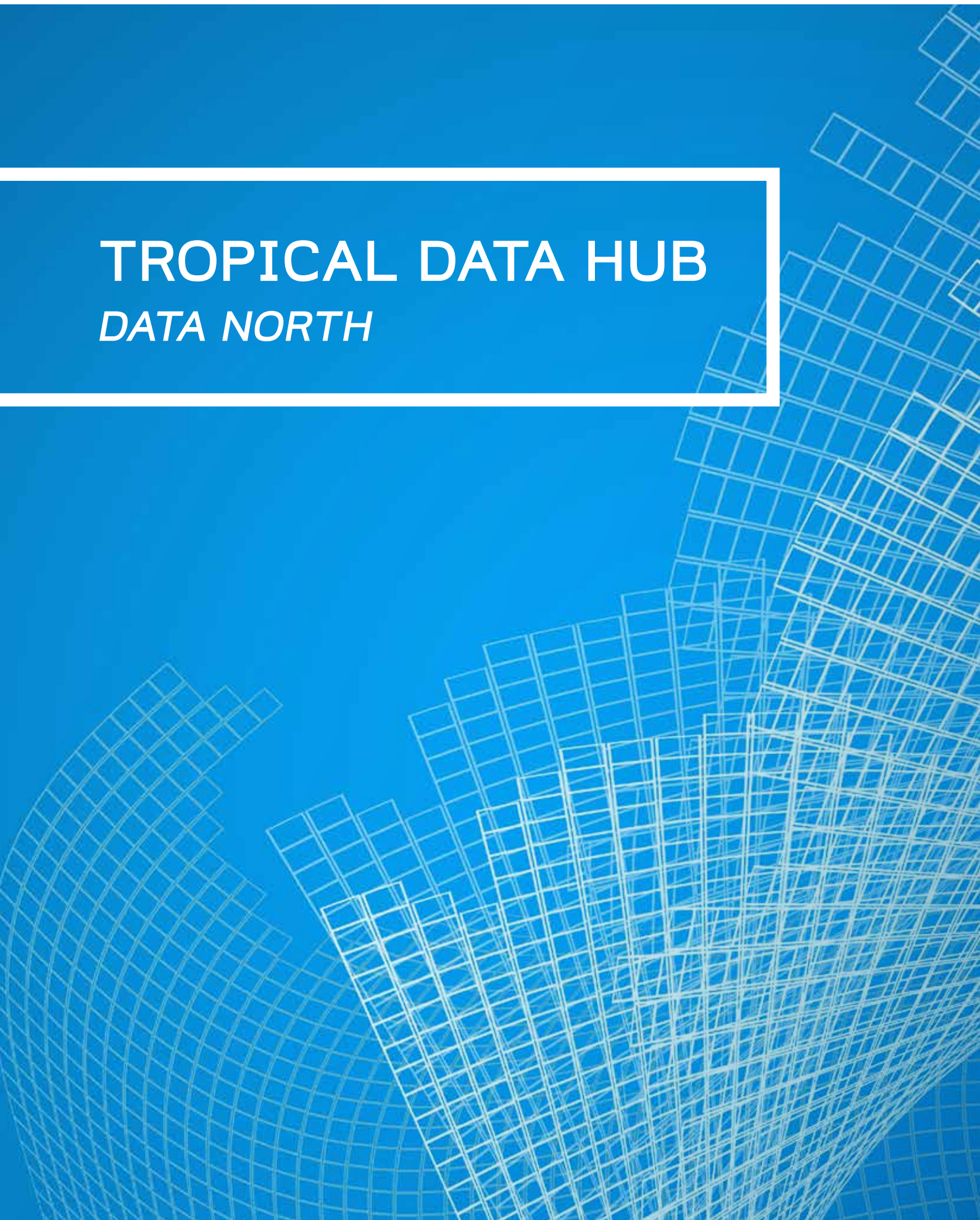
TROPICAL DATA HUB

JAMES COOK UNIVERSITY AUSTRALIA

TROPICAL DATA HUB

DATA NORTH



A GREATER PROPORTION OF THE AUSTRALIAN ECONOMY IS LIKELY TO BE SUPPORTED BY NORTHERN AUSTRALIA BY 2040.

THE TROPICAL DATA HUB

Australia's long-term security and regional stability depends on northern Australia, with northern Australia likely to represent a greater proportion of the Australian economy by 2040. To best capitalise on this growth potential we need to leverage existing knowledge, data and best practices - the products of past investments, to encourage new investments from both the public and private sectors. However, efforts to progress investment, innovation and change are hindered by the lack of a single repository from which coherent and linked data associated with tropical Australia can be accessed. Instead researchers, government and importantly private investors are forced to spend significant time and effort manually searching and synthesiseing disparate datasets, or worse, re-collect or re-create data. This unnecessarily complicates and hampers the process of discovering the core data needed to explore project viability – let alone innovation and change.

The next wave of productivity and wealth generation will be based on our ability to create, store, access and use data (the data cycle) in new, transformative and innovative ways. This is an international trend. Only five percent of Australia's population is sparsely scattered across the north. This, coupled with the stop/start nature of research and investment activities in the region, means that in the past there was not only a paucity of data, but also few centres or activities focused on retaining the very limited data available; this is no longer the case. James Cook University, the Australian National Data Service (ANDS), the Queensland Cyber Infrastructure Foundation (QCIF) and other partners have been developing a single digital repository focused on collecting, cataloging, storing and transforming research data and services related to the tropics: the **Tropical Data Hub (TDH)**.

This is a system that:

/ Enables open access to, and the use of historically generated data and associated knowledge by industry, researchers and community collaborators (such as innovators and entrepreneurs).

/ Reduces cost and technical barriers to the creation of and access to data and knowledge by third parties.

THE TROPICAL DATA HUB IS ABOUT

Increasing productivity in industry, government and research through greater speed, increased accuracy, reduced costs and reduced waste.

Leveraging existing data to foster **innovation** and change.

Mitigating investment risk, through access to credible, accurate data.

Transparency, accountability and leverage via the use of high quality, referenced data being translated into the core knowledge needed to support high value public and private investment decisions.

INCREASING PRODUCTIVITY

Historically, investment in the creation of data and knowledge has often generated siloed datasets, which have remained relatively invisible or just plain inaccessible. This not only represents a poor investment, but can result in countless dollars and hours being needlessly spent recreating data that (may) already exist. The TDH creates a path to value for this historic investment by translating it into time and dollars saved.

Productivity is gained through three immediate effects:

- A return is generated on historic investment.
- Waste is reduced or costs avoided as unnecessary repeat work is avoided and scarce resources can instead be directly invested into new activities.
- Data-dependent enablers of economic activity become faster and more efficient to generate.

Time is money; timely and cost effective data (for risk mitigation and project evaluation purposes, for example) can make or break an investment proposition. Avoidable data and analytics costs and delays caused by protracted regulatory evaluation processes that could readily be truncated with access to an agreed single source of truth, can make a project unviable. Delays can also cause projects to miss critical market cycle conditions, resulting in non-investment.

FOSTERING INNOVATION AND CAPABILITY IN THE NORTH

Increasingly, national innovation agendas are being influenced by arguments about the knowledge economy. While these may seem distant from northern Australia in many ways, **innovation and creativity will be essential to northern development**. Access to comprehensive digital data sets will provide a strong toolkit for national and international innovators to develop new startup firms to solve key problems to help unlock the economic potential of the north. Transport, logistics, commerce, design and consulting are all 'low-hanging fruit' for change and reinvention that could be tackled by new approaches to business.

Across the north in Cairns, Darwin and Townsville as well as other locations, there are emerging pockets of innovation. 'Start-ups' – the classic silicon valley model of small, agile and focused firms – are starting to emerge and embrace software and technology to solve real business problems in the north. The greatest assistance these firms can receive is access to data and services that drive the value of their products and increases market awareness. These developments would have a positive ongoing impact on regional capability and capacity.

MITIGATING RISK

Credible, validated data is pivotal to effective risk identification and mitigation. Large-scale investment activities depend on an **ability to mitigate risks**. Many of these risks are associated with environmental impacts, for example. The management of these risks - from both the public policy and private investor perspectives - hinges on access and analysis of credible data.

Existing data sets can be mined to achieve the productivity benefits outlined above. Should data gaps be identified which require the collection, production and analysis of new data, the raw data and derived knowledge can be added to the accumulated and accessible body of information. This ongoing process of *update, renewal and expansion* would ensure the relevance and sustainability of the data for risk mitigation and management purposes.

In addition, data waste and redundancy risk are reduced. When data is siloed and held in disaggregated environments, there are substantial risks to data loss. Data loss equals wasted funds and lost opportunity.

INNOVATION AND CREATIVITY WILL BE ESSENTIAL TO NORTHERN DEVELOPMENT.

ACCOUNTABILITY AND LEVERAGE

A readily accessible data source also has a range of ancillary 'public good' benefits. These are evident by opening up the work of researchers to the general public (directly, or indirectly via, for example, media organisations), whereby the broader public recognition and accountability of the value of investment in knowledge is promoted.

The TDH program has been building partnerships and collaborations across northern Australia since 2012.

Opportunities now exist to leverage data management and discovery needs across the north, with many organisations which are seeking fast, low barrier solutions to data storage and sharing. Programs, such as the National Environmental Science Program (NESP) and the proposed AgNorth CRC, will depend on large-scale data management and dissemination and present strong potential leverage points for a northern Australia data hub. The development of common knowledge stores represents a natural synergy with these significant Commonwealth investments and discussions are being held with the project proponents.

Commercial activity, based around planning and EIS developments, is dependent on fast access to data and represents an opportunity to explore long-term sustainable business models around data sharing and access.

PROOF OF CONCEPT IN PLACE: DE-RISKING INVESTMENT

JCU, in collaboration with a number of research infrastructure partners, has invested in building the proof of concept architecture to support data-driven productivity and investment growth. This pilot, called the Tropical Data Hub (TDH, http://www.tropicaldatahub.org) has been developed with a continuous improvement cycle embedded into the proof of concept architecture. The proof of concept is now ready for scaled development and deployment.

An early draft interface release of the TDH is available, that:

- Categorises several hundred data sets from a wide range of contributors into broad SEO Socio-economic objectives (SEO) areas (Business, Sciences, People and Economies).
- Allows for simple searching of information and access to primary and in some cases, processed data products.

With the expected shift towards more data/information centric economies and services In coming years, the data and analytic insights derived from information within the TDH would provide key insights for industries and governments wanting to realise economic opportunities presented by tropical regions.

It is important to note that while the TDH is a JCU initiative and currently hosted at JCU, it is intended to become an independent entity with an independent board. This independence is necessary to effectively partner with external organisations and give certainty to data providers and commercial partners.

Current infrastructure within the TDH now describes over 2,5,000 individual data sets and specific analytic tools have been developed by JCU researchers, as exemplars of the insights that can be gained by reusing and reanalysing existing data sets. Particular examples include the identification of critical future refuges for threatened species and automated data sets related to current and projected species distributions: key data for any EIS application.

Extension of the TDH to become the core, point of truth data repository for northern Australia would require investments in:

- Governance and Management
- Outreach and Data Access Agreements
- Up scaling existing software infrastructures
- Continuous development of analytics tools and services

The proof of concept work already undertaken has effectively de-risked the proposition of a northern Australian Data Portal and established a pathway for wide scale rollout and implementation.