Big Data Maturity in Australian Agricultural Industries

Andrew Skinner
Data to Decisions CRC

The Australian agricultural sector's big data maturity was assessed using the Data to Decisions Cooperative Research Centre 'Big Data Pillars of Success' model as a base and augmented with observations of big data maturity in adjacent sectors. In response to industry feedback taken during the Accelerating Precision Agriculture to Decision Agriculture (P2D) research project process (via workshops, interviews and the producer survey) the areas appraised for maturity were Strategy, data, analytics, culture, architecture, governance and training. Five maturity levels from ad hoc through foundational, competitive, differentiating to breakaway were evaluated, and the model was applied to the agricultural industries represented by the 15 Research and Development Corporations (RDCs) participating in the P2D project.

The participating industries demonstrated a low maturity level across all categories when evaluated against the project-developed maturity model. Many of the challenges and impediments to agriculture's big data maturity in Australia can be addressed through the creation of an actionable cross-industry digitisation roadmap. However, to move the sector from ad hoc to differentiating or breakaway across all categories will require focused cross-industry effort and cross-RDC collaboration.

The data maturity model for the Australian agricultural sector presented in this document was developed by Data to Decisions Cooperative Research Centre (D2D CRC). It is based on the D2D CRC Big Data Pillars of Success (Figure 1) and augmented with observations of big data maturity in adjacent sectors such as mining, utilities, health and manufacturing.

The resulting model evaluates six of the seven characteristics of the D2D CRC Pillars. It should be noted that based on industry feedback taken during the Accelerating Precision Agriculture to Decision Agriculture (P2D) research project workshops, industry interviews and the wider producer survey, the Data/Culture pillar has been evaluated as two separate entities in the model, whereas they are grouped in the original D2D CRC pillars. Additionally, due to the prevalence of established cloud services, infrastructure maturity has not been included in the evaluation as it was deemed unnecessary for the purposes of this paper.

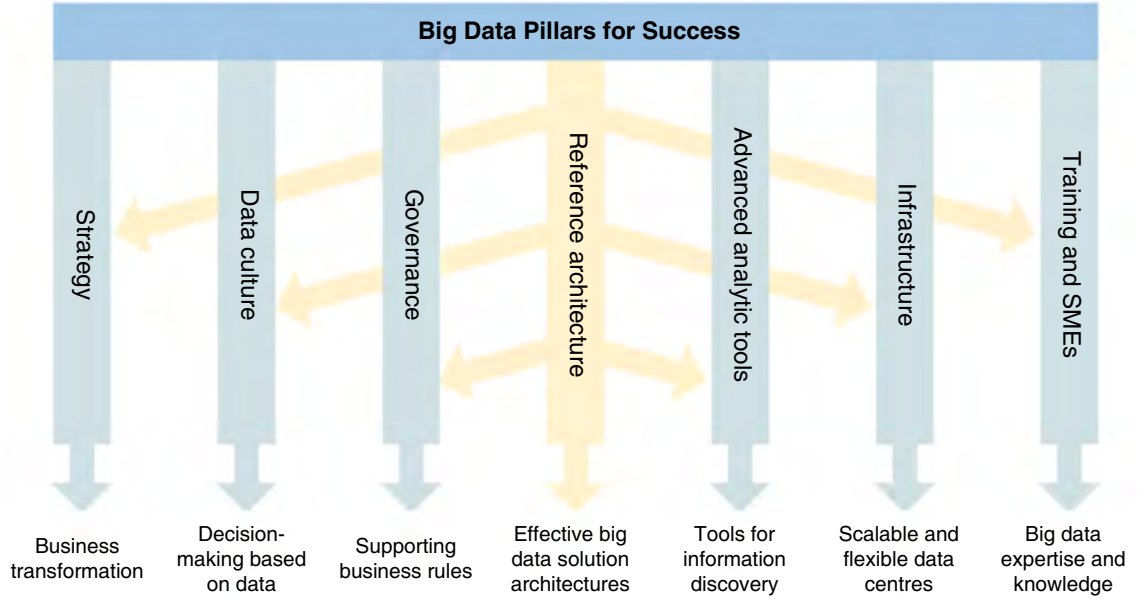

Figure 1: Big data pillars of success.

As a result, the areas evaluated were: Strategy, Data, Analytics, Culture, Architecture, Governance and Training. These categories were evaluated against five maturity levels from Ad hoc through Foundational, Competitive and Differentiating to Breakaway.

The model was applied to the agricultural industries represented by the 15 Research and Development Corporations (RDCs) participating in the P2D project.

The Precision to Decision Big Data Maturity Model

The D2D CRC Big Data Maturity Model comprises seven categories (detailed below) describing five levels of maturity. The model considers not only the technical maturity required for big data success but also (where appropriate) also considers on-farm business practices.

Elements of existing maturity models – eg those created by IBM in their evaluation of enterprise capabilities such as mobility, governance and service-orientated architectures – were also used in developing the model.

The evaluation outcomes were reached from observations taken during the eight P2D project regional producer workshops and supported with substantive desktop research, interviews with individual producers, nominated RDC representatives and supportive commercial providers.

Maturity Categories

Strategy

The first consideration with any advanced technology capability in agriculture (such as big data) is to recognise that its use must support demonstrable producer outcomes – all activities need to track through to increased profits. While robust technology is required to capture farm and industry data, and to execute computationally intensive analytics, deep industry and production expertise is needed to derive meaningful insights and then to use them to develop valuable whole-of-farm business outcomes. These outcomes are achieved by enriching value chains and driving operational improvements. To achieve this, producers and industry organisations need to transform their businesses to become more data aware, exploring data for new decision-making opportunities and leveraging these decisions to maximise yield and/or profit. Producers and organisations that are mature in the model have developed and communicated strategies enabling them to use all available data and apply analytics to innovate and improve their decision-making processes, maximising their value chain and opening up new market opportunities.

Data

Use of data to inform on-farm decision making is the base capability. However, agricultural value chains mature in big data analytics understand that data is a valuable asset. This data comes from many sources such as business records, machinery, data collected on-farm via the Internet of Things (IoT) and contributed third-party data from sources across the value chain as well as other external and open data providers. Industries mature in their approach to big data support producers by providing governed access to their data, and sharing it with third parties (where valuable) to give additional meaning and context.

Analytics

Mature use of analytics optimises on-farm decision making. The industry survey conducted as part of the P2D project showed that many producers are already using data to manage their finances and to report their regulatory compliance. However, analytics can also give producers a more detailed picture of why something has happened or can predict what is likely to happen in the future. The resulting insights help inform decisions, enabling producers to maximise yields and/or profit. Agricultural industries mature in their approach to analytics make data-driven decisions pervasive throughout their value chains. This requires timely insight in context.

Culture

Big data-aware industries recognise that use of analytics (as a tool to derive insights and aid decisions) is of limited value without the

engagement of producers, workforce and stakeholders along the value chain. To maximise the benefit of big data, producers must embrace analytics as a core tool of their trade, constantly seeking out and utilising decisions derived from their available data. This requires significant cultural change. To enable this cultural change, trust in these data-driven decisions is essential so that use of data becomes reflexive. To drive this cultural change, it is also important that producers have the ability to easily visualise analytic results and to provide feedback so that their decision support systems become 'go-to' business tools. Industries mature in their approach to big data culture promote a data or information-first approach to decision making and offer producers flexible, targeted data analytic services aligned to their production systems in order to embed data use into industry practices.

Technology

A planned, evolutionary, strategic technology approach to big data and analytics is essential for agricultural industries to establish a thoughtful, stable, scalable capability. This enables a managed approach to data access by end users, including producers, researchers, commercial organisations, government and industry bodies. This approach provides agility to address changes in production processes and across the value chain, as well as enabling interoperability. Mature industries will establish architecture that supports the primary three Vs of big data: volume, variety and velocity. They will support the three Vs through the creation of and reuse of shared cross-industry architectural patterns, data and standards. This includes supporting cross-cutting challenges such as security, governance and service levels.

Governance

With P2D workshop participants citing a lack of confidence in agricultural big data systems as a significant barrier to adoption, governance is a critical success factor for agricultural big data projects. Mature industries will have business models and policies in place that consider ownership, provenance, currency, data quality, foundational data and metadata, lifecycle management, security, privacy and ethical data use.

Training and SMEs

Big data solutions are worth little unless the relevant skills are available in industry to maximise the benefits of investment. A structured approach to building big data and data science expertise both through training staff and augmenting capability by identifying, evaluating and establishing trusted subject-matter experts (SMEs) is essential for big data success. A mature industry will recognise big data and data science as core competencies that offer market differentiation and build business value, and will also invest in people and partnerships to maximise the opportunities.

Maturity Levels

The project determined five levels of maturity for evaluation:

1. ad hoc
2. foundational
3. competitive
4. differentiating
5. breakaway.

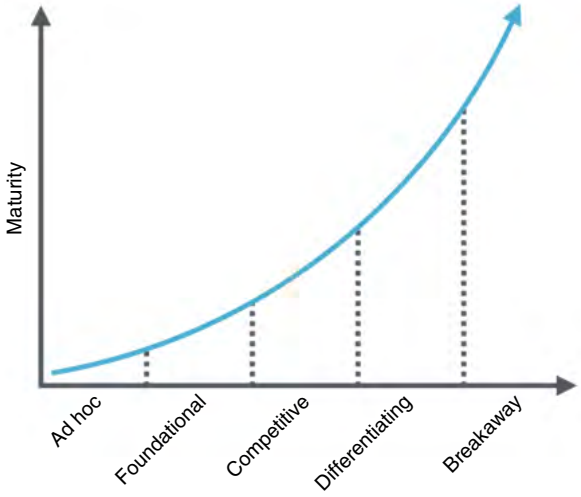

Figure 2: Maturity levels.

These levels illustrate maturity from basic (ad hoc) through to advanced (breakaway). The application of each is dependent on the category with which it is associated. The following matrix (Table 1) describes the maturity expected at each level for each category.

Table 1: Maturity matrix.

	Ad hoc	Foundational	Competitive	Differentiating	Breakaway
Strategy	Big data is discussed but not reflected in industry or farm business strategy. Use of data is limited to financial management and regulatory reporting.	Producers and industry bodies recognise that data can be used to generate profit. However, most data at scale is held by larger supply chain entities.	Industry strategy supports producer access to data throughout the value chain to support whole of farm decision-making processes.	Industry strategy realises competitive advantage for producers using data to make decisions.	Data drives continuous cross-industry and production innovation.
Data	Producers and the industry value chain use historical structured and unstructured data to make decisions.	Data is used to effectively inform producer decisions.	Data is applied across the value chain to improve industry and producer operational processes and decision making.	Data and predictive data driven decision making is used as an industry differentiator.	Data is used as a strategic asset driving producer profits and increasing market access.
Analytics	Analytics is limited to describing what has happened.	Analytics is used to inform decision makers why something on-farm has happened.	Analytical insight is used to predict the likelihood of what might happen.	Predictive analytics is used to help optimise producer decision making with linked farm management actions taken to maximise yield and/ or profit.	Predictive analytical insight optimises production processes and management actions are automated where possible.
Culture	The application of analytical insight is the choice of the producer and has little effect on profit or the wider industry.	The producer is aware of the insights available from data produced on farm but is largely resistant to adaptation required to take advantage of the insight.	The producer makes limited farm management decisions using analytical insight to improve operational efficiency and generate additional profits.	Producers are well informed with insight from analytics, and the capable of acting to maximise resulting yield/profits.	The producer and value chain continuously adapt and improve, using analytical insight to support their strategic objectives.
Architecture	Producers and the industry value chain do not have a single, coherent architectural approach to big data within their farm business.	A common big data framework exists but does not extend to new data sources or advanced analytics capabilities.	Best-practice architectural patterns for big data and analytics are defined and have been applied in certain areas.	Big data architecture and associated standards are well defined and widely adopted covering most of the volume, variety and velocity challenges for structured and unstructured data.	Established big data architecture fully underpins business strategies to enable complete market disruption with volume, variety velocity and veracity specifications applied.
Governance	Data governance is largely manual and barely sufficient to stand up to legal, audit and other regulatory scrutiny.	Understanding of data and its ownership is loosely defined and managed in a piecemeal fashion.	Policies and procedures are implemented to manage and protect core data through its life in the organisation.	The degree of confidence in data and resulting insights is reflected in making decisions.	Information governance is integrated into all aspects of the business processes.
Training and SMEs	There is little or no big data expertise within the industry.	Skills gaps have been identified but no formal strategy to expand expertise to fill gaps has been developed.	Training frameworks are implemented and have been applied to key areas of the value chain.	Training frameworks are mature and big data analytics expertise is established throughout the value chain.	Big data analytics is recognised a core competency within the industry generating measured returns.

Evaluating the current state of big data maturity

The project applied the model to the 15 agricultural industries represented by the participating RDCs. The eight P2D Regional Big Data workshops supplied the inputs to the model, and additional insights were taken from the CSIRO P2D Producer Survey results. While wide variances in capability both within and across industries were observed, the observations measured against the model were averaged to present a whole-of-Australian-agriculture view of big data maturity. The observed maturity results are presented for the seven categories below.

Strategy

***Observed as ad hoc:* 'Big data is discussed but not reflected in industry or farm business strategy. Use of data is limited to financial management and regulatory reporting.'**

None of the 15 RDCs canvassed were able to provide for review a published digital strategy that described the roadmap for the digitisation of their industry and specifically, where big data analytics mapped into their industry roadmap. Several RDCs had strategies in development and had identified big data as a source of value to their industry but were not able to assign a dollar value against their industry value chain adoption of the technology.

Of the producers attending the workshops, several of the larger organisations had developed commercial strategies for their businesses but were unwilling to share the detail of those strategies for consideration due to a perceived risk to competitive advantage. For smaller producers, 70–80% were keeping data for financial management and regulatory reporting purposes only, with the exception of higher than anticipated adoption of on-farm weather stations for climate-based decision making. Smaller farm businesses had no planned digitisation strategy and had not directly considered the benefits of big data/analytics use in their decision-making processes. Instead, they relied heavily on their consultants and advisers' strategic advice in this area. Of the consultants who attended the workshops, along with those interviewed as part of our research, none had developed a big data or analytics strategy or differentiated themselves with big data analytics within their industry.

Data

***Observed as ad hoc:* 'Producers and the industry value chain use historical structured and unstructured data to make decisions.'**

During the P2D project, it was observed across all 15 participating industries that producers are collecting increasing volumes of data. The CSIRO P2D Producer Survey supports this observation (see Figure 3).

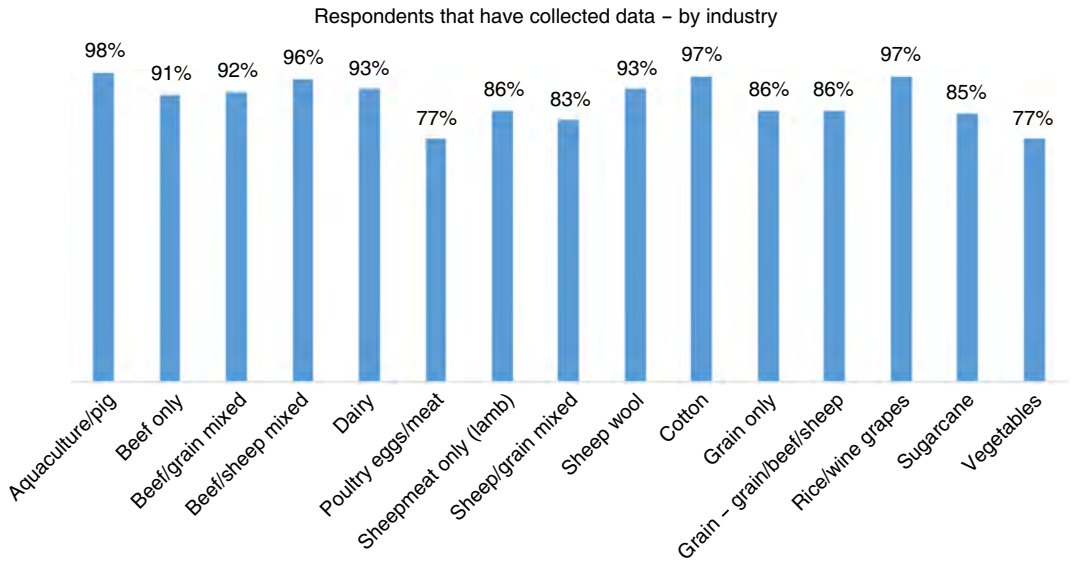

Figure 3: Survey results – data collection by industry.

Source: CSIRO P2D Producer Survey 2017.

Despite the limited data in use by smaller producers, the data collected on-farm and by value chain organisations was generally found to be more diverse than expected and included (but was not limited to) structured data representing climate, soil moisture, irrigation, yield and farm inputs. It is also important to note that within each industry there is a diverse spectrum of value chains – as such, they cannot be observed as one entity. By virtue of their ability to control all aspects of production, many vertically integrated or closed value chains were able to capture and interoperate much larger and diverse data sets than those value chains made up of independent organisations and individual producers.

It was observed that data is also generated at volume by on-farm machinery, labour and a growing number of other on-farm IoT sensors. In addition to structured data from machine telematics and sensor readings, this also includes data from unstructured sources such as normalised difference vegetation index (NDVI), geophysical, drone, satellite and other kinds of imagery also in use (although not always necessarily at scale). Additionally, many non-digital sources of data are available such as printed haulage dockets, receipts and log book entries, some of which are now being digitised.

Data quality is often an issue as data is frequently collected ad hoc or taken from poorly calibrated sensors, yield monitors and other sources. Governance to remedy this issue is not well established or absent.

Despite the growing availability of online farm management solutions, industry consultants and smaller producers generally use basic apps to work with data, along with tools such as Microsoft Excel, limiting their intuition to evaluate their data to make on-farm decisions for their customers.

Data representation and interpretation was also a commonly observed problem. During the P2D workshops producers reported that they did not know how to evaluate the data available to them (for example, soil sample results) to make effective on-farm management decisions.

Additionally, the data in use is normally limited to a single property, making it impossible to benchmark performance.

Finally, many producers cited a lack of interoperability of data between tools and services, again indicating an overall lack of maturity across all industries, specifically in identifying business models for data sharing.

For all industries, it was observed that many larger value chain organisations working with multiple producers do have the capability to collect, aggregate and benchmark data. This enables them to gain broader regional and national insights – however, data science capability and data culture were cited as challenge areas as the organisations struggled to extract these insights. Additionally, in most cases this data is not well governed, nor made easily available back through the value chain to producers in order to inform effective on-farm management decisions.

Analytics

Observed as foundational to competitive: **'Analytics is used to inform decision makers why something on farm has happened / analytical insight is used to predict the likelihood of what might happen.'**

At present, much of the analysis of data undertaken across industries is used to understand what has happened in the past informing management decisions to optimise future outcomes.

Many academic analytics models have been developed over the last 25 years of industry-funded research. In many cases these analytics also aim to predict future decisions. For example: given a set of known inputs, when will an animal group reach the weight for optimal eating quality and cut to be turned off to the processor? Academics are also now starting to apply machine learning techniques in some areas with varying success. Other examples of models available to support future decisions include pasture availability, yield and carcase weight. Again, not all industries have the same level of skill available to them in this category.

While these models exist, the P2D project observed that their uptake and support within industry is not yet widespread and adoption by producers is ad hoc.

Reasons for this include:

- a lack of automation – high degree of effort is often required to use them
- complex user interfaces provide cryptic feedback and no clear management actions
- the data producers can make available to the models is limited and of variable quality
- intellectual property issues prevent integration into third party farm management tools
- lack of ground-truthing data to support the accuracy of the outputs.

This has led to a lack of uptake and (in some cases) a distrust of model outcomes. As such, producers predominantly return to evaluating past season performance when making their management decisions. Maturity in analytics was evaluated as foundational to competitive because although first-generation models exist, they are not yet widely adopted.

Culture

Observed as ad hoc: **'The application of analytical insight is the choice of the producer and has little effect on profit or the wider industry.'**

The project found that big data use is limited to a small number of motivated producers. While some producers had seen increases in efficiencies and thus profits through the implementation of data-driven decision making, overall the ad hoc application of big data analytics had very little effect on their business. The use of data to make full management decisions has had very little effect on the wider industry, due to interoperability challenges throughout the industry value chains.

Data culture is also a challenge as producers are not generally aware of the insights that are available from the data that they produce in their businesses, with the exception of some motivated groups in each industry. The perceived cost of technology adoption also means many are hesitant to invest, particularly when potential benefits from the available insights are not well known.

It was observed that industry bodies such as the RDCs are also unable to clearly articulate positive, verified outcomes and impacts resulting from the adoption of big data analytics technologies.

Technology

Observed as ad hoc: **'Producers and the industry value chain do not have a single, coherent architectural approach to big data within their farm business.'**

Findings from the eight regional producer workshops, along with the follow-up interviews with producers, the RDCs and other value chain organisations, informed the P2D researchers that there is no common approach to big data across the participating industries. The P2D project – specifically, the big data reference architecture (BDRA) – will make recommendations based on cross-industry research to address this. While the current state of technical maturity for big data technology within Australian agriculture has been evaluated as ad hoc, cross-RDC implementation of the P2D project recommendations and the establishment of well-articulated industry digital strategies and change plans will quickly increase maturity in this area.

Governance

Observed as foundational: **'Understanding of data and its ownership is loosely defined and managed in a piecemeal fashion.'**

Each participating industry is aware that the data it creates has value. There is also a strong awareness regarding privacy and intellectual property. All the participating RDCs could provide data privacy policies. Many larger producer organisations (such as processors, supermarkets etc) also had established governance policies regarding data use in their organisations.

However, many organisations in the value chain had loosely defined governance in place for data and were unaware of specific data governance concerns such as copyright. When challenged, most were unable to clearly articulate the difference between private and public data in their businesses, and many were not clear on the current state of licencing agreements they had entered into with their consultants and machinery manufacturers. Other governance activities such as provenance, currency, data quality, foundational data and metadata, lifecycle management, security and ethical use had not been considered.

As a result, the maturity of the participating industries for big data governance was evaluated as ad hoc to foundational. Larger value chain organisations were demonstrably more mature in these areas than most of the industry producers.

Training and SMEs

Observed as ad hoc: **'There is little or no big data expertise within the industry.'**

When interviewing each participating industry, the lack of or absence of data science and big data capability was commonly mentioned. Of the industries interviewed, most had not identified or considered how to address the gap, while the two industries that had formally identified the need had not yet considered how to address training. This was also found to be true with producers. Those attending the regional workshops cited a lack of expertise in data, big data and analytics as a significant gap in their businesses. Additionally, they cited a lack of capability in their value chain and noted that their consultants often did not have the degree of capability required to help them make decisions from their data which could lead to demonstrable outcomes in their businesses. The results of the producer survey supported this view. As a result, the project evaluated the industry maturity in training and SMEs as ad hoc within the maturity model.

Summary

It is clear that the participating industries demonstrate a low maturity level across all categories when evaluated against the project-developed maturity model, with five of the seven categories observed as ad hoc, one as foundational and one as foundational/competitive.

1. All industries currently lack a clear strategic roadmap for developing big data capability both at industry and producer scale.

2. While data is being collected across industries, foundational data sets have yet to be defined, interoperability is an issue and good data management skills are still to be developed.

3. The business outcomes of data required to build a data culture are still to be validated, and the time required to do this may take many years in some industries, due to production factors such as crop rotation etc.

4. Culture change across agricultural sectors will speed the adoption and thus the realisation of the benefits of big data analytics technologies. Industry bodies are as yet unable to clearly articulate verified outcomes and potential benefits from digital adoption, and big data use is limited to a small number of motivated producers.

5. Whole-of-value-chain technology standards are yet to be developed and cross-industry collaboration to maximise spend is yet to occur. There are some industries who are developing or have developed data analytics capability but not at petabyte scale. Other industries are building substantial data sets but are unsure how to query the data for decision support insights.

6. Governance is in place at a macro level across industries in terms of managing data privacy, however deeper governance of data such as copyright, provenance, currency, data

quality, foundational data and metadata has not been established in any of the participating industries.

7. A structured approach to building big data and data science expertise is yet to be developed by the participating industries. Capability has been cited as being the leading market failure across industry value chains.

Many of the above challenges can be addressed through the creation of an actionable cross-industry digitisation/big data roadmap, quickly moving the participating industries towards maturity. However, to move from ad hoc to breakaway across all categories (if needed) will require focused cross-industry effort and cross-RDC collaboration.

About the Author

Over the last 18 years, **Andrew Skinner** has travelled the globe designing, developing and deploying digital tools and services into challenging physical environments for some of the world's biggest brands. More recently his focus has been on building tools, business models and techniques to help organisations derive value from information and analytics. Prior to a chance, career-changing meeting with some sheep and his subsequent appointment to the Integrity Systems Company he led the Data to Decisions CRC Innovation Exchange program, where he worked closely with the manufacturing and agriculture industries to develop a number of research based decision support prototypes. He also contributed to the Rural R&D for Profit Precision Agriculture to Decision Agriculture Project where he gained expertise in data use across the 15 major Australian agricultural industries.