

Final Report

MLA project code:

Prepared By:

L.GEN.1708

S. Fischer, S. Parry, P. Green Greenleaf Enterprises

Date Published:

PUBLISHED BY Meat and Livestock Australia Limited Locked Bag 991 NORTH SYDNEY NSW 2059

Final Report

This is an MLA Donor Company funded project.

Meat & Livestock Australia acknowledges the matching funds provided by the Australian Government and contributions from the Australian Meat Processor Corporation to support the research and development detailed in this publication.

This publication is published by Meat & Livestock Australia Limited ABN 39 081 678 364 (MLA). Care is taken to ensure the accuracy of the information contained in this publication. However MLA cannot accept responsibility for the accuracy or completeness of the information or opinions contained in the publication. You should make your own enquiries before making decisions concerning your interests. Reproduction in whole or in part of this publication is prohibited without prior written consent of MLA.



Executive Summary

Sheep Genetics hosts and manages a set of databases of genetic traits and resultant indexes and breeding values. These databases service much of the Australian Sheep industry as well as parts of NZ and the US. The sheep genetics data collection and analysis model is significantly established in terms of data sharing and collaboration. This represents a significant value opportunity that the Australian Sheep industry has over the Australian Beef industry which contains silos of data held by breed societies desperate to protect their IP at the expense of much greater levels of genetic gain than currently being seen.

The knowledge and capacity that exists in both Sheep Genetics and AGBU which have created the current model and associated database systems is significant, valuable and is key to moving forward in any solution. However, like many database systems which have existed and grown for 20 years or more, there comes a point where they need to be rebuilt due to outdated technologies, disconnected processes and multiple single points of failure. The recent resignation of a key manager has also highlighted the issues of knowledge retention being lost to the process.

This precipitated this review which worked closely with key Sheep Genetics, AGBU, MLA and Sheep CRC database managers, analysts and management to identify a number of key issues. This was followed by a 1 day workshop made up of these key proponents along with most of the Sheep Genetics council. The results of this process are detailed in section 3. A summary of the key issue areas identified through this review process are:

- Inflexible, unstable database design difficult to add new traits, other measurements and results
- Fragmented and disconnected data collection, slow validation of data inputs and outputs
- Poor, intermittent web performance, monitoring and management
- Outdated unsupported technology
- Reliance on a single person per process / single point of failure
- Inflexible non-standardised framework for solution providers, Sheep Genetics and AGBU
- Poor standards, documentation and unplanned approach
- Missed opportunities for collaboration between solution providers, breeders, Sheep Genetics, AGBU and MLA
- Limited Integration with other databases and consideration of existing and new data sources and analytics

Dairy Australia faced almost identical issues 20 years ago when considering the state of the ADHIS database and analysis process. The process they undertook and the costs involved were also not dissimilar to the recommendations outlined as part of this review. The process from here for Sheep Genetics should involve Diary Australia where possible to leverage off lessons learnt through the process as well as any documentation such as Database Functional Design Specifications Diary Australia are willing to share.



Collaboration will be a key to the success of Sheep Genetics moving forward and will involve breeders, solution providers, Sheep Genetics, AGBU and MLA as well as external organisations affected by the process. Because these changes have wide reaching ramifications affecting many stakeholders, putting in place a well-structured Change Management Plan is highly recommended. While outside the scope of this review, some of the key change management considerations which need to be put in place to assist Sheep Genetics in preparing, managing and reinforcing the changes recommended here have been highlighted in section 5.

The key components of the solution recommended by this review with associated expected costs are detailed in section 2. These are summarised below:

- Consolidated LambPlan / MerinoSelect / Resource Flock database
- Robust, high performance, secure, scalable cloud based solution
- Database Replication between Sheep Genetics and AGBU
- Analysis Results Warehouse enabling Integration with other data sources and powerful analysis tools
- Hosting of the Genomics Database
- Central Sheep Genetics/AGBU Data Elements, Security and Validation Administration
- Consolidated LambPlan / MerinoSelect / Resource Flock / Genomics website central data entry, on-farm management, reporting
- New API framework standardised business rules, validation and security
- Joint Sheep Genetics /AGBU database management team

This solution, while essential to address the issues highlighted in this review, also presents an opportunity to build a platform suitable for future growth which significantly enhances the breeder's experience from data collection to reporting, drastically reduces inefficiencies with processing, validation and reporting, develops standards, documentation and training; and opens the option for data to be shared externally. The key benefits outlined in section 3 can be summarised as:

- Fast reliable access to data for breeders, producers and researchers
- Fast standards based validation of data
- Fast reliable website and reports
- Fast, flexible website and reporting fuller access to results including genomics and resource flock, financial and other analysis
- Secure, fast and fully externally maintained database and servers
- Real time reporting and responses for data collection ability to rectify problems at the start of the process
- Standards, documentation & training developed for end users
- Ability to easily analyse data with other data sources such as NLIS and BOM
- Accessibility for data sharing with solution providers, R&D and commercial entities
- Early issue detection and simpler problem solving
- Easily scalable, manageable, self-healing, load balanced, geo-redundant technology
- Vetted, collaborative process



This will be a significant project expected to span at least 2 years covering database, website and API changes and will involve AGBU as well as other organisations such as SIL and AMSEA to adjust their current processes. It will also involve significant changes to the way all stakeholders work (i.e. breeders, database administrators, solution providers). However, it does present significant opportunities for each of these stakeholders to interact with Sheep Genetics databases and analytics which could significantly increase genetic gain in the Sheep industry.

One commercial participant in this review made the comment that if they had greater and simpler access to some the data and analytics produced by Sheep Genetics, they would be willing to invest their own commercial dollars in R&D and solutions which added real value back to breeders leaving industry dollars to be spent in areas where they could be better utilised. The solution recommended in this review enables this kind of data access and sharing model while giving full manageability and ownership to both industry and in particular the breeder.



Contents

E>	ecu	utive Summary2
С	onte	ents5
1	В	Background6
2	S	Solution Design and Costings8
3	ls	ssue - Solution - Benefit Map18
	3.1	Summary18
	3.2	New Single Database with security, validation and Web based Admin interface25
	3.3	New Website / API / Data Collection and Reporting Process
	3.4 trair	New Data and Analysis Management Team – Overseeing design, documentation, ning and ongoing management of issues, new requirements41
4	Т	imeline, process and detailed costings48
	4.1	Current state48
	4.2	Stage 148
	4.3	Stage 2
5	С	Change Management Planning55
6	А	
	6.1	List of Tables60
	6.2	List of figures60
	6.3	MOQ Digital High Level Design Specifications60



1 Background

Sheep Genetics hosts and manages a set of databases of genetic traits and resultant indexes and breeding values. These databases service much of the Australian Sheep industry as well as parts of NZ and the US. These databases have been developed over 20 years or more and have reached a point where they need to be rebuilt due to outdated technologies, disconnected processes and multiple single points of failure etc. The recent resignation of a key manager has also highlighted the issues of knowledge retention being lost to the process.

Following this, Sheep Genetics engaged Greenleaf Enterprises to:

- Conduct a review of Sheep Genetics databases and associated databases to assess functionality, efficiency and the ability to integrate with other systems.
- Scope out options for redeveloping Sheep Genetics databases at the strategic and architectural levels in order to create a more effective and efficient platform.

This review worked closely with key Sheep Genetics, AGBU, MLA and Sheep CRC database managers, analysts and management to identify a number of key issues. This was followed by a 1 day workshop made up of these key proponents along with most of the Sheep Genetics council.

After collating these issues along with proposed logical solutions, MOQDigital was engaged to help develop a high level architectural design. These key considerations in this design were:

- Consolidation and future upscaling of the Sheep Genetics Dbs.
- Collection and submission of data from breeders and Service providers to Sheep Genetics.
- Validation and analysis of Data.
- Data flows between Sheep Genetics, AGBU, MLA and Sheep Genetics' internal and partner's systems.

The purpose of the high level architectural design was to:

- Describe specifications concerning integration between the following systems; Breeders on farm software (e.g. Sapien and Practical Systems), Pedigree Wizard, Pedigree Master, Sheep Genetics Web site, Sheep Genomics and AGBU
- Consider integration with systems external to Sheep Genetics such as NLIS, LDL, BOM, LANSAT.
- Propose a new production platform based on Microsoft Paas Offering that will enable Sheep Genetics to integrate and leverage systems hosted currently on-premises into the cloud.

The key technical benefits of the high level architectural design are to:

- Establish a new workload model for scalability and performance.
- A central management system break the dependencies on different infrastructure teams for needed resources.
- Benefit from a managed service environment where data can be optimised for security and consumption.
- Better Agility rapidly prototype, build and deploy new changes while leveraging Microsoft Azure cloud.



The high level architectural design provided by MOQDigital was then incorporated with other findings attained throughout the entire review process including the Sheep Genetics workshop. These were then compiled into this report which details, at a high level:

- the solution from a technical point of view
- the tasks involved in achieving this solution both technical and managemental; and
- the estimated costs involved in achieving this solution



2 Solution Design and Costings

The key part of this review was to propose an overall solution and associated expected costs associated with not only resolving the current database issues but also set a platform for future growth of Sheep Genetics. Figure 1 depicts the recommended solution pictorially. Table 1 then follows on detailing each of the elements of the solution.



Figure 1 - Solution Architectural Diagram



Solution Component	Details
Robust, high performance, secure, scalable	Automated high performance monitoring and in-build
Soution component	 Automated high performance monitoring and in-build business continuity mechanisms Includes out of the box website and database monitoring and management which automatically detect performance anomalies, addresses load balancing, fault tolerances, availability etc ensuring a seamless and fast user experience Utilises high availability geo-redundant storage and disaster recovery for business continuity and protection against potential data loss No infrastructure or code to manage and maintain All components which make up the solution are either PaaS or SaaS based and as such do not require physical infrastructure Outsourced infrastructure and database/web technicians allow SG database administrators to better serve SG clients Scalable and extendable as needs grow Database and website loads, speeds and sizing can scaled up or down via a management console as parts of the business grow or shrink Additional Azure components such as machine learning and powerful BI reporting can easily be "plugged in" and leverage off the existing structures Leveraging the SQL azure capabilities for scalability as the business grows, maintaining high throughput performance
Consolidated LambPlan / MerinoSelect /	Synergy of resources skills and knowledge
Resource Flock Database	 Consolidation of LambPlan, MerinoSelect and INF/Resource Flock – allowing a central and single point of management while also facilitating effective data analysis and reducing inefficiencies like data validation duplication, reliance on multiple databases and multiple data management points; and better utilisation of costs. Replaces outdated technology serviced by single individuals
	New more agile data structure allowing:



	 Simple and efficient addition of new traits and results Maintenance of trait validation rules Ability to define other trait characteristics such as trait accuracy and financial benefits Reclassification of some traits such as weight at age vs age class to aid more detailed analysis 				
	Sets a baseline for future growth				
	automated data sync between those systems.				
	• Structure also allows for future incorporation of				
· · · - · · · ·	other research data				
Analysis Results Warehouse	Provision for Big data Azure data warehouse (vs SQL DB) is geared				
	towards analytical online processing at the scale				
	of hundreds of terabytes of data from different				
SQL	Full Analytical platform				
	Opportunity to incorporate in the future richer				
	analytics and data visualization of both structured				
	Analytical platform is tightly integrated with tools				
	such as Azure machine learning power BI and				
	HDInsight.				
	The data warehouse incorporates pause and				
	resume feature that will allows SG to load the				
	data a few times a month and then build new				
	pause it while cutting on computing cost and				
	keep the data intact in the Azure storage.				
Genomics Database	Separation of SG Database and Genomics				
	 Likely to grow to BigData vs SheepGenetics DB Commercial operation – potential ownership 				
	issues				
SQL	Better automated integration with Sheep Genetics				
	New Animals added to SG db for genomics tests				
	of new animals				



	 Updated list of SGA IDs/animal data transferred
	to Genomics DB
	 Genomics results transferred to SG DB
Integration with other data sources	Opportunity to map whole supply chain overlaying
MLA On-premises network	weather patterns
	• Integration and synchronization with other data
NUS Carcars Eaglback Animal Health	sources such as MLA's corporate network and
All S Carcass recovery Animal Health	external partners (eg. BOM, LANSTAT etc) for
	analysis.
Data Management	• Opens up ability for richer analytics such as
gateway	linking carcase feedback and animal health to
	genetics
	Secure, automated, seamless data synchronisation
	 Utilises Azure's data factory service and data
: r>	management gateway component to enable a
BOM()	secure channel data transfer between data
s s :	warehouse and external source.
e e e e e e e e e e e e e e e e e e e	• Allows for the transformation of raw data assets
FTP server	(both structured and unstructured) from different
°	data base systems into trusted information that
	can be then loaded in the AZURE sql Data
	warehouse and shared with BI and analytical tools
LANSAT()	
Replication between SG and AGBU	Priority performance access for both website users
	and high end R&D/OVIS Analysis
	• AGBU/R&D have the same access/response times
Data Mañagement gateway	as if they housed the database in house
•	• Gives priority access to end users in terms of
	performance not being affected by large
	R&D/OVIS queries being run
	Clear separation of data accountability
SG Replica AGBU Genomics Replica	Keeps accountability of data access where it
AGBU Network	appropriately belongs by giving:
	 AGBU write access to data elements
	managed by them such as adjustments and
	OVIS results; and read-only access to
	animals and traits.
	 SG write access to data elements managed
	by them such as animals and traits; and
	read-only access to adjustments and results.
Consolidated LambPlan / MerinoSelect /	Consolidates functionality of all genetics processes
Resource Flock / Genomics website	• Pedigree Master/Wizard, MerinoSelect uploads
 Control data antru 	SGAR Reporting SG Search







	 continuous model where records can be added/updated as needed Secure access to third parties Access to developers, third party providers and research companies (such as CSIRO) to submit and query the data by APIs based on their security access level. A security layer to protect the data from unauthorised access.
Central SG/AGBU Data Elements, Security and	Web based traits and validation rule maintenance
Validation Administration	 Allows SG / AGBU database administrators to add traits, set security access for end users, administer validation rules – pre/post analysis via web platform vs a reprogramming exercise.
	 Web based security maintenance Allows securing of whole associations (eg. Dohne,
	 SIL, AMSEA) Allows for greater data sharing opportunities in conjunction with the breeders
SG/AGBU database management team	Vetted, collaborative processes
 Development and documentation management Issues Register Promote/incentivise Solution providers 	 A new cross SG/AGBU database management team whereby all changes are vetted, impact considered, approved and scheduled and adequately communicated with all stakeholders including software providers and breeders. A more robust process should mean less bugs which need to be fixed on an ad hoc basis. Where such a need arises, pre-agreed escalation processes will already be in place with external development teams to allow fast tracked bug fixes to be undertaken still within a robust testing enterprise level framework. DB management team are responsible to outsourcing to service providers etc. They are also the interface with the end users, ensuring adequate help is available to them either via documentation or one on one training (mostly over the phone).
	 Standards, documentation and training The database management team will oversee the collating a full set of documentation including detailed requirements and design specifications, all database elements and their definitions



	training material pitched at various levels for end users. This process will also include breeders, producers and solution providers in the redesign of the website including them in focus groups, testing and enabling them as champions of the new website. The database management team will also oversee the redesign of the database and website. This will include performance improvements.
--	---

Table 1 Recommendations Explained



Implementation and Pricing options

The solution outlined above gives Sheep Genetics a platform which not only resolves the current issues identified but also provides opportunities for:

- Future scalable, manageable growth
- Integration with other data for easy and powerful analytics
- Accessibility for data sharing with solution providers, R&D and commercial researchers to add value back to industry.

Table 2 outlines expected costs associated with this solution (detailed costing assumptions with expected timeframes in section 4). These are based on an estimated number of days to complete on a base rate of \$1,500 / day assuming the entire project will be outsourced. A number of skills would be required from developers @\$60-\$80/Hour to Solution Architects @\$180-240/Hour. An early part of the project would involve a full functional specification and pricing which would detail costs. For this high level approach, the averaged rate of \$1,500 was used. The figures used here are not dissimilar to costs for the very similar ADHIS database rebuild project. In fact, these costs may even be considered quite conservative in comparison. Table 3 details how these costs were arrived at by detailing number of person days per task.

Sheep Genetics may also consider options of hiring staff or contractors as opposed to outsourcing the entire project which could reduce costs. However, given the previous failed attempt to consolidate LambPlan and MerinoSelect, an externally outsourced project which utilised SG key staff would be the recommended approach of this review.

A number of options have been detailed in the Table 4 which begin with the very minimalistic approach of consolidating LambPlan and MerinoSelect and running them on a server maintained in house at Sheep Genetics. Note that this is not the hosted solution recommended in the review. It is used merely to show the costs involved to resolve the key issues of:

- Unsupported Technology
- Reliance of a single person at multiple points of the chain
- A more robust database design

A number of enhanced solutions are also costed with the last line being the costings, including ongoing maintenance costs, associated with this recommended solution. Note that these are not all incremental. For example, building a database to be housed in house and then migrating to Azure at a later stage will incur duplicated costs of relinking all processes connected with the database.

The recommendation of this review is to implement the full solution in 2 stages:

- Stage 1 will be focused on building a foundational framework setting up an Azure cloud based network, scoping out, building and migrating one new consolidated database and setting up external organisations including AGBU, SIL, AMSEA to be part of the new Azure network. Current processes such as web site access and Pedigree Wizard uploads will remain unchanged from a user perspective apart from more reliable, faster access. This minimises the change control required and allows stage 1 to be focused on setting up the framework for stage 2.
- Stage 2 will build on the foundational framework established in stage 1. This stage will heavily involve breeders, solution providers and other end users of either the web site or APIs firstly because many of the existing processes they are used will dramatically change and they need to be involved and helped through that change. And secondly because their buy-in is essential to the project. In fact, they will be key designers of the website and API requirements in conjunction with Sheep Genetics and AGBU.



Stage	Total Setup Costs	Ongoing Annual Costs	Due Date
Combine LambPlan, MerinoSelect and Resource Flock, Create Analysis Data warehouse, Integration with other data sources, Build SG/AGBU Admin website - Move to Azure Cloud	\$1,195,500	\$137,651 p.a.	31-Mar-18
Build new combined Website, Build new API framework, Migrate Genomics database - Move to Azure Cloud	\$705,000	No New Costs	31-Mar-19

Table 2 Estimated costs by stage

Table 3 below breaks down these phases by number of person days to complete each major segment with expected due dates to achieve.

Stage 1 Tasks - Database consolidation and migration to Azure network	Due Date	Estimated Person
		Days
Set up of Azure network	30-Jun-17	7
Consolidate LambPlan / MerinoSelect / Resource Flock into one new database	31-Oct-17	285
AGBU Replication – From SG DB and Analysis Results -> SG Replica	30-Nov-17	60
ReMap databases for current API framework (current functionality only)	31-Jan-18	45
ReMap databases for current websites	28-Feb-18	25
Creation of database administration website	30-Jan-18	125
Migrate automated processes to use new API framework and database	28-Feb-18	60
Convert Pedigree Master process	28-Feb-18	35
Convert Emailed files process	28-Feb-18	45
Finalisation - Testing and Cutover	31-Mar-18	110
Total	31-Mar-18	797

Stage 2 Tasks - Consolidate website, new API framework and genomics migration	Due Date	Estimated Person Days
Migrate Genomics database	31-Mar-19	50
Extend API framework	31-Mar-19	180
New Website	31-Mar-19	240
Total	31-Mar-19	470

Table 3 Project stage breakdown



The table below represents a number of options Sheep genetics could take from the minimalistic approach of simply combining the LambPlan and MerinoSelect databases to an in-house server base solution; to the full solution recommended by this review which includes rebuilt website, API framework and cloud based analytics.

Options	Trait, Validation, Detailed Process Mapping, Functional design specifications. Database creation and migration Conversion of current processes (emailed reports, AGBU exports/imports) Testing and Cutover	API Framework	SG/AGBU Admin Website	Consolidated Website	Migration of Genomics Data	Hardware / Software required - Initial Setup and Build	Total Setup Number of days to implement each option	Ongoing Annual Costs Hardware/ Software required (including personnel to maintain)
Combine LambPlan and MerinoSelect Only (AGBU Results in DB vs Warehouse) - Keep On Premises	620	30	-	-	-	3	653	\$27,500 p.a.
Combine LambPlan and MerinoSelect Only (AGBU Results in DB vs Warehouse) - Move to Azure Cloud	620	30	-	-	-	3	653	\$43,439 p.a.
Combine LambPlan and MerinoSelect Only (AGBU Results in DB vs Warehouse), Migrate existing SG Website and APIs to cloud - Move to Azure Cloud	620	45	-	-	-	7	672	\$60,314 p.a.
Combine LambPlan and MerinoSelect Only (AGBU Results in DB vs Warehouse), Migrate existing SG/AGBU Website to cloud, Build new API framework, Build SG Admin website - Move to Azure Cloud	620	225	125	-	-	7	977	\$60,314 p.a.
Combine LambPlan, MerinoSelect and Resource Flock (AGBU Results in DB vs Warehouse), Build new combined Website, Build new API framework, Build SG/AGBU Admin website - Move to Azure Cloud	620	225	125	240	-	7	1,217	\$60,314 p.a.
Combine LambPlan, MerinoSelect and Resource Flock, Analysis Data warehouse, Genomics (incl. Replication), Build new combined Website, Build new API framework, Build SG/AGBU Admin website - Move to Azure Cloud	620	225	125	240	50	7	1,267	\$137,651 p.a.
Combine LambPlan, MerinoSelect and Resource Flock, Analysis Data warehouse, Genomics (incl. Replication), Integration with other data sources, Build new combined Website, Build new API framework, Build SG/AGBU Admin website - Move to Azure Cloud	620	225	125	240	50	7	1,267	\$137,651 p.a.

Table 4 Estimated days required and associated costs - options



3 Issue - Solution - Benefit Map

3.1 Summary

This database review has closely with key Sheep Genetics, AGBU, MLA and Sheep CRC database managers, analysts and management to identify a number of key issues. This was followed by a 1 day workshop made up of these key proponents along with most of the Sheep Genetics council. The results of this process are detailed in this section. A summary of the key issue areas identified through this review process are:

- Inflexible, unstable database design difficult to add new traits, other measurements and results
- Fragmented and disconnected data collection, slow validation of data inputs and outputs
- Poor, intermittent web performance, monitoring and management
- Outdated unsupported technology
- Reliance on a single person per process / single point of failure
- Inflexible non-standardised framework for solution providers, Sheep Genetics and AGBU
- Poor standards, documentation and unplanned approach
- Missed opportunities for collaboration between solution providers, breeders, Sheep Genetics, AGBU and MLA
- Limited Integration with other databases and consideration of existing and new data sources and analytics

The key components of the solution recommended by this review with associated expected costs are detailed in section 2. These are summarised below:

- Consolidated LambPlan / MerinoSelect / Resource Flock database
- Robust, high performance, secure, scalable cloud based solution
- Database Replication between Sheep Genetics and AGBU
- Analysis Results Warehouse enabling Integration with other data sources and powerful analysis tools
- Hosting of the Genomics Database
- Central Sheep Genetics/AGBU Data Elements, Security and Validation Administration
- Consolidated LambPlan / MerinoSelect / Resource Flock / Genomics website central data entry, on-farm management, reporting
- New API framework standardised business rules, validation and security
- Joint Sheep Genetics /AGBU database management team

This solution, while essential to address the issues highlighted in this review, also presents an opportunity to build a platform suitable for future growth which significantly enhances the breeder's experience from data collection to reporting, drastically reduces inefficiencies with processing,



validation and reporting, develops standards, documentation and training; and opens the option for data to be shared externally. The key benefits outlined in this section can be summarised as:

- Fast reliable access to data for breeders, producers and researchers
- Fast standards based validation of data
- Fast reliable website and reports
- Fast, flexible website and reporting fuller access to results including genomics and resource flock, financial and other analysis
- Secure, fast and fully externally maintained database and servers
- Real time reporting and responses for data collection ability to rectify problems at the start of the process
- Standards, documentation & training developed for end users
- Ability to easily analyse data with other data sources such as NLIS and BOM
- Accessibility for data sharing with solution providers, R&D and commercial entities
- Early issue detection and simpler problem solving
- Easily scalable, manageable, self-healing, load balanced, geo-redundant technology
- Vetted, collaborative process

The pages that follow summarise groups of issues, solution elements and benefits into 5 key groups. Each issue, solution element and benefit is then detailed.



SSUES:	I just want a database & processes that deliver my clients & me reliable, timely info I trust. so workshop Participant			
efficiency of processes, input, nalysis, output		Central SG/AGBU Data Elements, Security & Validation Administration Select / Resource Flock DB Genomics DB Replication between SG & AGBU		
agmented and slow validation f data inputs and outputs	Solution			
flexible, unstable DB design – fficult to add new traits, other easurements & results				
mited data accessibility, wnership & privacy		Fast, reliable access to data for breeders, producers & researchers		
oor standards, documentation planned approach	Benefits	Simple approach to add new validation methods, traits & results		
eliance on a single person per rocess / single point of failure		Clear ownership, management and accessibility of data for breeders, AGBU & SG Remove reliance on a single person per process / single point of failure		



a

F

In di m

Li

P(&

R/







L.GEN.1708 – Sheep Genetics Database Review – Final Report

ISSUES:	Make data collection easier for greater accuracy & confidence in the AGBUs used in the industry; Enhanced / less human intensive diagnostic tools; Less human-intense input & validation processes so Workshop Porticipants				
Limited web based data querying and analytic tools Lack of admin & user based	Solution		Consolidated LambPlan / MerinoSelect / Resource Flock / Genomics website • Central data entry • On farm management		
tools and logic Limited billing & other financial ability Fragmented data input- validation-reporting process	and logic ed billing & other ial ability eented data input-		Reporting API Network Real time reporting & responses for data collection – ability to rectify problems @ start of process		
Fragmented and disconnected data collection Slow, insecure, inflexible data access	Benefits		Less breeder intensive data collection – embedded with current processes Fast, reliable website & reports Flexible reporting – fuller access to results including genomics and resource flock, financial & other analysis User specific functionality & simpler access to invoices etc.		



ISSUES:

Inflexible non-standardised framework for solution providers, SG and AGBU

Consideration & use of new technologies & new or better data collection

Consideration of existing & new data sources & analytics

Poor collaboration between solution providers, breeders, SG, AGBU & MLA

Training, documentation & user collaboration in design

DB & website monitoring & management

Marketing, customer awareness

Missed opportunities to access joint & wider resources → Synergy of Resources

Solution

Benefits

SG/AGBU DB Management Team

 Development & documentation management

Promote/incentivize Solution Providers

Training

Set standards of what data to accept - not to

discriminate, but to maintain TRUST 55 Workshop Participant



Issues identified and resolved

Vetted, collaborative process

Standards, documentation & training developed for end users

Coordinated change management



3.2 New Single Database with security, validation and Web based Admin interface

3.2.1 Issues

3.2.1.1 Inefficiency of Processes, input, analysis, output

LambPlan and MerinoSelect store roughly the same information but on different animal types. A duplication of resources are required to manage the collection, pre-analysis and reporting of that data. Information Nucleus Flock (INF) / Resource Flock data is also stored in a separate database and, although has a slightly different focus, records similar information and is included in the breeding value analysis.

As well as being duplicated, current processes are also human resource intensive at a range of points in the process:

- Data submission
 - Breeders need to download and email data from their on-farm software. Because validation of uploaded data is not in real time, by the time incorrect data is identified, it is difficult to retrace where the problem arises. If their data is not correct, it will not be included in the analysis and they will not receive analysis results. Often breeders will enter incorrect information simply to ensure they don't miss the analysis run.
 - A second inefficiency exists because the breeder holds master data records. Every record is uploaded every time, every record is re-validated and adjusted on every upload
- Data validation
 - SG Database administrators who need to validate the same data multiple times, problems arising from incorrect data can take time between AGBU, SG and breeders to solve and is an inefficient process
- Data Analysis
 - \circ $\;$ AGBU Staff need to re-validate and re-transform data for each analysis
- Data retrieval
 - The time lag between submission and results can be long and often unreliable. Data retrieval is also difficult as it comes from multiple sources and multiple mediums in the form of website searches, emailed reports etc

Quote from Industry engagement

I just want a database & processes that delivery my clients & me reliable, timely info I trust.

- Improved efficiency in database use to release staff (SGA) for extension
- Potential for pre-analysis work to help manage run times



- Seamless transfer of data for analysis
- Data transfer (email) is not easy
- Seamless data export
- Seamless data import
- Pathway for data exporting from producers needs simplifying
- Manual data exchange process. Exporting files
- A more applied farm application \rightarrow real time data collection
- Ease of capturing data accurately & efficiently
- Ease of collection \rightarrow yard (farm) to database
- Efficient transfer of information between stakeholders without any confusion
- I just want a database & processes that delivery my clients & me reliable, timely info I trust
- Run times \rightarrow run turnover
- Problem solving needs to be easily done
- Efficiency of data set creation for analysis SG & RF & Genomics etc

3.2.1.2 Fragmented and slow validation of data inputs and outputs

The current validation process is built, maintained and run by a single person for LambPlan and a single person for MerinoSelect. While communication occurs between the two database administrators, there is no guarantee that the same validation is occurring. In fact both processes are completely different and are continually rebuilt often via updated programming code (such as "if animal Id=x then change to y). AGBU also have their own set of analysis they run pre and post analysis. While there are good relations and communication between AGBU analysts and SG database administrators, the processes are disconnected and disjointed.

Because no standard exists across all validation processes, diagnosis of issues can be time consuming and difficult. When errors are found and corrected within processes, previously reported index values change and raise questions on the validity of the process itself.

Quote from Industry engagement

We need a set standards of what data to accept - not to discriminate but to maintain TRUST.

- Easier diagnostics of issues in database
- Quality assurance for apps, reporting, data collection devices
- Set standards of what data to accept not to discriminate but to maintain TRUST
- Who should validate data
 - $\circ \rightarrow breeder$
 - $\circ \rightarrow \text{software}$
 - $\circ \rightarrow database$
 - $\circ \rightarrow AGBU$



- Validation process \rightarrow data quality & guidelines > 30% syndicates :(Bad data can still get in!
- Automatic data entry from computer after validation
- How is the validation of the incoming data happening
- Past data or old animals changing
- Less human-intense input & validation processes
- Differences in data between runs MS

3.2.1.3 Inflexible, unstable database design – difficult to add new traits, other measurements and results

New measurement technologies are continually becoming available and new analyses being discovered. Therefore the requirement to collect new traits is ongoing. The current process of adding new traits requires multiple changes requiring:

- Database redesigns in LambPlan i.e. the addition of new table columns
- New records to be added in MerinoSelect
- Changes to AGBU's results databases
- Import and Export changes for Lambplan, MerinoSelect, Sheep Genetics Website, Pedigree Wizard, Pedigree Master, AGBU analysis programs and solution provider's software
- Analysis changes and reporting changes for LambPlan, MerinoSelect and AGBU

Such changes take a long time to implement.

Already changes such as weight by age as opposed to age class have been flagged as required but difficult to implement in the current structure, particularly LambPlan which uses a single table approach with one record per animal with all traits as columns in that record.

Other requests include ability to enter a dollar value per trait for financial analysis, validity of measurement and date collected.

Quote from Industry engagement

We need to address process and communication pipeline with development and upgrades - new traits etc. all the way from software to report & use.

- Process and communication pipeline with development and upgrades new traits etc. all the way from software to report & use
- Able to add new traits
- Able to enter a \$ value for a trait to analyse financial impact
 - \circ sensitivity impact
 - o uniform units of measure
- Future data capture
- Flexibility/robustness of data when analysed work without problems



- Is aligning genetic recording and data transfer with non-genetic (ie management) a problem?
- Uniform ID: is it EID or 16 digit SG ID
 - o 1 # to link all data from lamb classing meatworks / yarn processors
- Analysis changes take a long time to implement

3.2.1.4 Limited Data Accessibility, Ownership and Privacy

There exists a growing desire to share data with a range of participants including:

- Breeders Additional information reported back from analysis
- Researchers for "industry good" functions
- Solution providers for ease of integration into apps for breeders and producers
- Increased information for breeder's customers, the producers
- Other third parties with a potential of purchasing the data

With this increased level of sharing, concerns are raised about who owns the data, who owns the right to share it as well as the separation of IP. For example between the raw data owned by the breeders and the breeding values produced by AGBU analysis.

Quote from Industry engagement

We need to give access to "outsiders" to data for analysis and "industry good" functions, but maintain privacy.

Issues identified in the workshop

- We need to give access to "outsiders" to data for analysis and "industry good" functions, but maintain privacy
- Access of reports by all stakeholders
- Access to data
- Protecting IP
- Who "owns" the data? How can it be used?
- Confidentiality of information
- Ownership of information
- Percentiles are not automatically included with BVs
- Grower ownership of data
- IP/security/ownership of data issues
 - o all owners need confidence that system is secure

3.2.1.5 Poor Standards, documentation, planned approach

Minimal or no standards currently exist between LambPlan, MerinoSelect, Resource Flock and AGBU Analysis Results for:

- Database design or technology
- Validation process



- Development approach or technology
- Reporting formats
- Import/export formats or gateways

There also exists minimal or no documentation for any of these. Were a single member of staff leave, as has recently been the case for MerinoSelect, the ability to replace that functionality requires analysis of program code, understanding potentially unsupported database and program languages let alone understanding the processes by which data flows from trait to result.

Issues identified in the workshop

- Standard format for reporting & submitting data
- Exclusions & standards
 - - pedigree
 - scanner accred
 - \circ measurements changed to avoid exclusions
- Future volumes of data

3.2.1.6 Lack of Monitoring and alerts

Breeders are currently given little feedback on when tests results are ready to be retrieved or when animals have invalid data and are likely to be excluded from tests.

Issues identified in the workshop

- Genomics office & SG communication → when genomic info is back, what sheep have tests pending
 - $\circ \rightarrow$ parentage tests (who, how many etc)
 - Data just appears

3.2.1.7 Limited Integration with other databases

Programs like INF and Resource Flock are paramount to genetic gain in sheep. However, potential exists for this to be done on a much larger scale and include other data such as animal health, BOM or LANSTAT data. The current state of databases do not easily allow these kind of interconnectedness of other data sources.

- Genomic pipeline needs to be well integrated with SG delivery / database
- Lack of carcase data to breeding decisions
- Ability to link to other databases
- Interconnectivity integrate NLIS, LDL and all new data sources?



3.2.1.8 Untapped Value of data and subsequent analytics

This follows on from lack of integration with other data sources. Once Sheep genetics data could be integrated with other data sources, who would have access to it and how would they use it for the benefit of industry.

Issues identified in the workshop

• Medium term: how to recognise value of data from different places animals, breeders?

3.2.1.9 Missed opportunities to access joint and wider resources, Synergy of resources

Each person within the process works effectively in silos. There is collaboration both within SG and between SG and AGBU. However all database administrators work with different database technologies which they are familiar with leaving little opportunity for cross development collaboration for example. There is also little connection with the MLA Information company who have a team of both database and web developers.

Quote from Industry engagement

Resourcing: development is a massive job. Ongoing support and development also required.

- Access to avenues of expertise to assist in data collating & management
- Any contingency plan in place?
- Better use of service providers to take pressure off SG
- Cost effective maintenance & modification
- Is outsourcing DB management an option? In-house?
- Resourcing: development is a massive job. Ongoing support and development also required
- Database consolidation:
 - o DNA
 - o CRC
 - o INF
 - o AGBU
 - o SIG
- Economy of scale
- Common platform for LP & MS
- Getting ALL R&D data into the one system AWI, INF
- To use the same DBMS for lambplan and merino
- Use the same system for lambplan & Merino
- Why not have 1 multi-point entry system, and 1 database team?
- Fragmented system
- Disconnected database



3.2.1.10 Outdated technology

The LambPlan database is built in Dbase III, a now unsupported database language. Pedigree Master and SGAR reporting are both 16 bit applications which severely limit the hardware they can be installed on.

3.2.1.11 Poor, intermittent web performance

MLA IT support SG with a range of virtual servers to house some of their databases and websites. This infrastructure seems reasonably inadequate for the task it is performing. While some redesign of the website itself could aid performance, simply moving it to a more robust, load balanced web server specific environment would likely address most issues.

3.2.1.12 Slow queries

The virtual servers which MLA IT supply SG with appear to be inadequately speced for the job they are built for. The database itself which the website uses could also be optimised for better performance.

3.2.1.13 Redundancy and backups reliant on single points of failure

While database managers maintain some level of backup and redundancy, these are widely managed with human intervention and performed at the instinct of each database administrator. There is no clear backup or disaster recovery plan in place giving clarity as to what level of disaster is acceptable and what the process of recovery would be from that disaster.

3.2.2 Solution

3.2.2.1 Consolidated LambPlan / MerinoSelect / Resource Flock database

LambPlan, MerinoSelect and INF/Resource Flock databases all record similar information - Animals and Traits. Consolidating them into a single database has a range of obvious benefits:

- Synergy of resources
- Simplicity of adding new traits.
- Single point access of information AGBU for BV analysis, Other researchers, the breeders themselves
- Allows for future incorporation of other research data
- Allows a central and single point of management while also facilitating effective data analysis and reducing inefficiencies, like data duplication, costs related to reliance on multiple databases and multiple data management points as well as allowing scalability as the business growth
- Leveraging the SQL azure capabilities for scaling and performance by providing high availability and disaster recovery.
- An automated linkage to other data sources for analysis by automated data sync between those systems.
- An out of the box azure HADR solution for business continuity and protection against potential data loss.



Breeder is still the 'master' of their data in that they can change/delete it but it is held in one central repository guaranteeing consistency, reliability, backups, lack of conflicts between data sources

The process of consolidating LambPlan and MerinoSelect was attempted with the current MerinoSelect database and was not successful. Anecdotal evidence would suggest that the design and management of this process was conducted by a single developer in isolation from the rest of SG and AGBU. As already stated, this is one of the key issues with the entire Sheep Genetics data collection / analysis and reporting process. Attempting to repair a disconnected process in a disconnected way achieved the result as would be expected.

It is the recommendation of this document that the consolidation of the LambPlan / MerinoSelect / Resource Flock databases be conducted in this way:

- The process to conducted by a group external to SG / AGBU
 - Expertise in Business Analysis, Database Design, Genetics and Change Management
 - Current database managers able to continue their current workloads while guiding the development process
- The external group would have the current SG and AGBU database managers and analysts as a key reference group

The database design itself will determined as part of the detailed functional specifications developed throughout the early part of stage 1. But it would be expected to be based on a simple but flexible design including the suggested list of tables below:

- Animal
- Trait list of traits and what they do, accuracy levels (e.g. Distinguish INF weight with a breeder entered weight), sources, \$ value (to perform financial analysis)
- Results list of results and what they do
- AnimalTraits and their associated attributes
- AnimalResults and their associated attributes
- TraitRules for input db. Produces anomaly reports
- Security rules around each trait and each measurement per animal

3.2.2.2 Central SG/AGBU Data Elements, Security and Validation Administration

Current validation is administered by a range of programming changes, database entries and manual procedures. This part of the solution would incorporate a web based administration tool by which SG/AGBU administrative staff would have access to:

- Administer validation rules pre/post analysis via web platform vs a reprogramming exercise
- Set security access for end users (table security will limit what SG and AGBU have access to change). Also allows securing of whole associations (eg. Dohne, SIL, AMSEA)

AGBU and SG would have different access rights to data. For example, AGBU would have access to:

- Add/delete/update adjustment records
- Add/delete/update OVIS result records
- Add/delete/update Resource Flock records



• Add/Update new Traits, Adjustments and Measurements

The validation process itself would be able to be administered. For example:

- Some obvious traits, such as excessive weights, may be excluded from being entered into the database particularly if entered via the website. It will be important to include solution providers at this point to enable to same functionality via APIs so that on-farm software can provide the same functionality.
- Some traits may issue a warning to the user based on criteria which makes them seem abnormal. However they may be accepted into the database and flagged as 'UnChecked'
- Traits which need to be validated as part of a group of animals or post OVIS runs, would be flagged as 'UnValidated' at that time and the user alerted that these animals have been incorrectly entered and will be excluded from all OVIS runs until fixed. SG Staff would have access to exception reports also in order to work with breeders to amend
- If a breeder changed the records of an existing animal, those details would fall into the same validation process again.

3.2.2.3 Analysis Results Warehouse

The Azure SQL data warehouse is a fully managed data warehouse as a service. The main reasons for separating it from the SQL azure database (of Animals and Traits) are the following:

- **Big data** By leveraging a different computing architecture than the Sql azure, The DW is more geared towards analytical online processing at the scale of hundreds of terabytes of data from different data sources (AGBU, NLIS, BOM...) whilst the sql azure will be mainly provisioned to store raw data.
- **Full Analytical platform** In mind to incorporate in the future machine learning, richer data visualization and both structured and unstructured data the DW Analytical platform is tightly integrated with tools such as: Azure machine learning power BI and HDInsight.
- **Cost** The DW incorporates pause and resume feature that will allows SG to load the data a few times a month and then build new reports for its customers while being able to pause it while cutting on computing cost and keep the data intact in the Azure storage.

3.2.2.4 Genomics Database

Genomics has been separated out into its own database for the following reasons:

- Likely to grow to BigData vs SheepGenetics DB
- Commercial operation potential ownership issues

This part of the solution will also involve better automated integration with Sheep Genetics such as:

- New Animals added to SG db for genomics tests of new animals
- Updated list of SGA IDs/animal data transferred to Genomics DB

3.2.2.5 Genomics results transferred to SG database Replication between SG and AGBU

One of the key reasons for the replication of SG data between SG and AGBU is to allow priority performance access for both website users and high end R&D/OVIS Analysis

• AGBU/R&D have the same access/response times as if they housed the database in house



• Gives priority access to end users in terms of performance not being affected by large R&D/OVIS queries being run

It also give clear separation of data accountability

- Keeps accountability of data access where it appropriately belongs by giving:
 - AGBU write access to data elements managed by them such as adjustments and OVIS results; and read-only access to animals and traits.

3.2.2.6 Integration with other data sources

This part of the solution involves integration with the MLA Information company data sources such and NLIS, Carcase feedback and Animal health. It will also involve integration with other data sources such as BOM and LANSTAT.

This integration and synchronization with Azure SQL server from within MLA's corporate network and external partners will be achieved by means of the data factory service component. This component will enable the following:

- Securely replicate and sync data between on premise data store (mainly AGBU) and SG's cloud data store (LAMBPLAN/MERINOSELECT, GENOMICS). This will need to be coupled with the data management gateway, which is an agent that enables hybrid data movement and processing. This agent will need to be installed on the same machine as the data store itself, or on a separate machine that has access to the data store.
- Support the transformation of raw data assets (both structured and unstructured) from different external system (LANSAT, BOM) into trusted information that can be then loaded in the AZURE sql DW and shared with BI and analytical tools.
- Connects directly to multiple DB platforms that Lamb Plan\Merino Select interact with (PostgreSQL, MySQL) without the use of intermediary files.

Hybrid connections - This Architecture leverages the azure data factory functionalities to support Hybrid connection with the SG's network and linkage to external system such as AGBU. The azure data factory will serve as a platform to link SG's data stores whether hosted on premise or in-cloud and link to internal and external system.

The Data management gateway will be deployed on premise on a dedicated server. All other components which make up the solution are either PaaS or SaaS based and as such do not require physical infrastructure

3.2.2.7 Azure cloud based solution

3.2.2.7.1 Security

These services will be protected by SQL Server Authentication and explicit white listing of IPs allowed to access the servers. MLA must secure the set of APIs through the AZURE API management using JWT key, token and IP filtering. + data management gateway

3.2.2.7.2 HA/DR/Backups

Azure Read-access geo-redundant storage (RA-GRS) to provide geo-redundancy. These backups are created automatically and at no additional charge. transaction log backups generally happen every 5



- 10 minutes. Active Geo-Replication to configure a database to have up to four readable secondary databases in the regions of your choice. These secondary databases are kept synchronized with the primary database using an asynchronous replication mechanism. This feature is used to protect against business disruption in the event of a data center outage or during an application upgrade. Active Geo-Replication can also be used to provide better query performance for read-only queries to geographically dispersed users

3.2.2.7.3 Easily scalable, self-healing, load balanced, geo-redundant

Azure has a range of out of the box functionalities which make this a robust, high performance, secure, scalable cloud based solution.

3.2.2.7.4 Technology evolving solution

As new technologies become available, eg. New SQL or Windows versions or new Azure functionality, these are deployed and made available by a team of Azure IT professionals seamless to the user.

3.2.2.8 Remove reliance on a single person per process / single point of failure

3.2.3 Benefits

A new Single Database with security, validation and Web based Admin interface provide the following benefits:

- Remove reliance on a single person per process / single point of failure
- Secure, fast and fully externally maintained database and servers
- Fast reliable access to data for breeders, producers and researchers
- Fast standards based validation of data
- Clear ownership and accessibility of data for breeders, AGBU and SG
- Simple approach to add new validation methods, traits and results
- Ability to easily analyse data with other data sources such as NLIS and BOM
- Early issue detection and simpler problem solving

3.3 New Website / API / Data Collection and Reporting Process

3.3.1 Issues

3.3.1.1 Fragmented and disconnected data collection

D

Quote from Industry engagement

A more applied farm application \rightarrow real time data collection.



Issues identified in the workshop

- Use similar / the same way of collecting data (using same file format)
- Why not have 1 multi-point entry system
- I have to send all data in rather than each test
- A more applied farm application \rightarrow real time data collection

3.3.1.2 Slow, insecure, inflexible data access

D

Quote from Industry engagement

Make data collection easier for greater accuracy & confidence in the AGBUs used in the industry.

Issues identified in the workshop

- Fast and secure access of data
- Being able to get the data either individually or in bulk
- Simple data extraction
 - o make it simple for the end user to get the data
- Instant updates of info located to search page. E.g. change listing or contact details
- Make data collection easier for greater accuracy & confidence in the AGBUs used in the industry
- Seamless transfer of data for analysis
- Pathway for data exporting from producers needs simplifying
- Percentiles are not automatically included with BVs

3.3.1.3 Limited web based data querying and analytic tools

D

Quote from Industry engagement

New ways for service providers to access clients data / reports / diagnostics etc.

- Automated genetic gain limiters at flock level (enhanced dashboard delivery)
- Enhanced / less human intensive diagnostic tools
- New ways for service providers to access clients data / reports / diagnostics etc



- Diagnostics more readily available to breeders & SP
- More automatic delivery of reports → website update

3.3.1.4 Lack of admin and user based tools and logic

D

Quote from Industry engagement

I need reminders on when to send data in.

Issues identified in the workshop

- Can renew subs & pay bills in single log in while on search site
- I need reminders on when to send data in
- Web knows which analysis a breeder needs once key log in
- Contact lists within SG.
 - o website
 - o SAP
 - o reports
 - o others
- Instant updates of info located to search page. E.g. change listing or contact details

3.3.1.5 Limited billing and other financial ability

D

Issues identified in the workshop

- Billing information coming out of the database needs to be able to capture CAP amounts automatically
- Can renew subs & pay bills in single log in while on search site

3.3.1.6 Fragmented data input-validation-reporting process

D

Quote from Industry engagement

Validation of data while collecting and before submitting data.

Issues identified in the workshop

• Manual data exchange process. Exporting files



- Ease of capturing data accurately & efficiently
- Ease of collection \rightarrow yard (farm) to database
- How is the validation of the incoming data happening
- Less human-intense input & validation processes
- Prevent the submission for wrong or invalid data using the system rather than the individual
- Automation when submitting data and receiving
- Seamless transfer data
- Simple data submission ie. One click
- Validation of data while collecting and before submitting data
- Who should validate data
 - $\circ \rightarrow breeder$
 - $\circ \rightarrow \text{software}$
 - $\circ \rightarrow database$
 - $\circ \rightarrow AGBU$
- Simple data extraction
 - o make it simple for the end user to get the data
- Being able to get the data either individually or in bulk
- I can't trust the data, everyone else makes up results
- Differences in data between runs MS
- Support: access to results (ie. APF, updates)
- Manual on data exporting for producers
- Ease of data transfer & entry
 - o Easy to learn & use
 - Low cost especially labour cost
- Ease of in-field data collection
- Manual on data exporting for producers
- More automatic delivery of reports \rightarrow website update
- I just want a database & processes that delivery my clients & me reliable, timely info I trust
- Exclusions & standards
 - \circ pedigree
 - scanner accred
 - \circ measurements changed to avoid exclusions
- Past data or old animals changing

3.3.2 Solution

3.3.2.1 Consolidated LambPlan / MerinoSelect / Resource Flock / Genomics website (Central data entry, on farm management and reporting)

Incorporates and replaces functionality of:

- Pedigree Master/Wizard, MerinoSelect uploads



- SGAR Reporting, SG Search
- Sheep Genomics search website
- Resource Flock website

Breeders access to:

- Enter new animals/ update existing ones
- View OVIS results, indexes and BVs
- View Genomics results, parentage
- Search across multiple db

New data input model – continuous data updates vs session based

Ability for users to set up alerts and reports via email, SMS, store preferences as to what they want to see when they log in.

Ability for users to update their contact details etc within website Identified requirements

• Being able to get the data either individually or in bulk

3.3.2.2 Secure, scalable website

The SG website is the front end of where users (Breeders, sheep producers ...) interact to submit, search and query data. By leveraging the MS Web apps on azure, SG will be able to benefit from the following:

- Scaling and fault tolerance scale up automatically based on a load balanced platform which will automatically manage all users traffic.
- Run the website reliably on a self-healing, auto patching cloud platform.
- Compliance Be Iso, SOC2 and PCI compliant.
- Security Automatic Support for SSL.
- Availability Guaranteed SLA and availability for 99.95% of the time.

3.3.2.3 Website monitoring and management

The Azure app insight will serve as a monitoring tool for SG live web application. It will automatically detect performance anomalies, diagnose issues and understand how users interact with the website. The telemetries types that can be reported, using the azure app insight, include but not limited to:

- HTTP request rates, response times, success rates.
- Exception traces from both server and client.
- Page view counts, user and session counts, browser load times, exceptions.
- Segmentation by client location, browser version, OS version, server instance, custom dimensions, and more.
- Diagnostic and analytic tools such the ability to query over different telemetries, compose different chart to monitor all the app component.



3.3.2.4 API Network

SG Middle tier will incorporate the following:

- A central management point where MLA can publish old and new APIs to support their business objectives.
- A central point in the solution that will contain all the Business logic to validate data entered by breeders.
- Access to developers, third party providers and research companies (such as CSIRO) to submit and query the data by means APIs and based on their security access level.
- A security layer to protect the data from unauthorised access.

Integration with SG must not only be limited through traditional web browsers but must also provide richer capabilities and interface allowing mobile users, internal employees and partners to integrate with it or allow community developers to create extensions to their applications.

The API management layer will be hosted in azure as well and will allow SG mainly:

- Perform all the business logic in one centralized point such as: Authentication, animal and traits validation, data query as opposed to the current fragmented process where it is complex to define a standard for data validation or format.
- Control access to SG APIs and ensure it performs well and meet service levels.
- Business insight & Analytics' into API data.

the rationale behind this design stems from the fact that data comes in today in different formats and different 3rd parties are enforcing different validation rules which propagates errors in later stages of the flows whereas common APIs will transform and validate the data into one standard format thus preventing any errors and resubmission of data later down the track

3.3.3 Benefits

A new Website / API / Data Collection and Reporting Process provide the following benefits:

- Real time reporting and responses for data collection ability to rectify problems at the start of the process
 - \circ $\;$ Eg re data validation at point of data entry via APIs or the website
- Less breeder intensive data collection embedded with current processes
 - Eg APIs used to embed in current on farm software
- Fast reliable website and reports
- Flexible reporting fuller access to results including genomics and resource flock, financial and other analysis
- User specific functionality and simpler access to invoices etc



3.4 New Data and Analysis Management Team – Overseeing design, documentation, training and ongoing management of issues, new requirements

3.4.1 Issues / Benefits

3.4.1.1 Inflexible non-standardised framework for solution providers, SG and AGBU

This involves providing a documented, industry standard framework for solution providers to aid their implementation but also to require them to meet a set of standards in terms of approach.

D

Quote from Industry engagement

Set standards of what data to accept - not to discriminate but to maintain TRUST.

Issues identified in the workshop

- Inconsistent outputs from on-farm software extracts
- Infield data upload to system so no double handling use of handheld portable devices
- Lack of mobility of access
 - \circ real time
 - o easy to read
- Quality assurance for apps, reporting, data collection devices
- Business rules to allow connectivity & interaction
- Third party authentication to operate on behalf of breeders
- Enable opportunities for commercial service providers to easily use analysis
- Service delivery issues
- Set standards of what data to accept not to discriminate but to maintain TRUST
- Standard format for reporting & submitting data
- Nice if the database could talk to SAP to allow one entry point of client information. One source
- New ways for service providers to access clients data / reports / diagnostics etc
- Software standard formats
- Business rules to allow connectivity & interaction

3.4.1.2 Consideration and use of new technologies and new or better data collection

Quote from Industry engagement



Use of mobile phones / cloud for infield upload & review so data (paper) isn't lost. Must integrate with existing software program.

Issues identified in the workshop

- New methods for capturing data required
 - e.g. big data such as weights from walk over weighing system
- Use of data collection devices & technologies we don't yet know about
- Use of mobile phones / cloud for infield upload & review so data (paper) isn't lost. Must integrate with existing software program
- Lots of potential add ons:
 - o apps
 - o decision support tool
 - o better use of data
- Limitations of software
- Suggested pathways to help improve data quality
- Are so m any traits required for outcomes? Does DNA assist
- Suggested pathways to help improve data quality

3.4.1.3 Consideration of existing and new data sources and analytics

This involves connecting other data sources such as LDL carcase feedback, BOM, LANSAT data and the potential for genetic gain. Considers how records will be aligned, data sharing arrangements

Quote from Industry engagement

Improve the feedback available to producers from the processing sector for genetic improvement privacy.

- Lack of carcase data to breeding decisions
- AN systems should make whole flock recording the default
- Ability to introduce new animals into database without penalising that animal > opens up breeding option & source of genetics & cross-breeding
- Improve the feedback available to producers from the processing sector for genetic improvement
- Interaction across systems vital
 - - SG
 - - INF/RF
 - o LDL/NLIS etc
- Is aligning genetic recording and data transfer with non-genetic (ie management) a problem?



- Diagnostics for breeding decisions
- Ability to link to other databases
- Percentile table consistent across all sheep types
- Ability for animals outside the resource flock system to be introduced & not penalised
- Fertility data improvement this data is captured outside of a production stage
- Recognition of multiple joinings/year
- Interconnectivity integrate NLIS, LDL and all new data sources?
- Diagnostics for breeding decisions
- Automated genetic gain limiters at flock level (enhanced dashboard delivery)
- Enhanced / less human intensive diagnostic tools

3.4.1.4 Poor collaboration between solution providers, breeders, SG, AGBU and MLA

This involves a single database team, a design team which involves breeders, producers and solution providers. Would probably also need external agencies – integration with MLA Information developers, MLA marketing etc. Also involves funding as it considers collaboration with other MLA programs

Quote from Industry engagement

Should technology providers be "in the room" e.g. eid, drafters etc.

- Should technology providers be "in the room" e.g. eid, drafters etc
- Why not have 1 multi-point entry system, and 1 database team?
- Systems
 - o reliance on individuals
 - - staff can have a holiday!!!
- Economy of scale
- Security of funding
- We must NOT forget the human element in the entire process. Logic only works some of the time!
- Improve team work from support services
- Support for breeding direction, genetic advice
- Is there value in a 'product & service development' team or forum with resources available to convert
- Redundancy & resourcing; Stephen, Ermias, Klint, Aaron +++ should be on or accessing same systems & have time for development
- Who owns housing, funding, reach, species
- Improved efficiency in database use to release staff (SGA) for extension
- Very reliant on individual people. Not system based.
- Access to avenues of expertise to assist in data collating & management



- Any contingency plan in place?
- Better use of service providers to take pressure off SG
- Is outsourcing DB management an option? In-house?
- Resourcing: development is a massive job. Ongoing support and development also required
- Is there value in a 'product & service development' team or forum with resources available to convert
- Problem solving needs to be easily done
- Obtaining appropriate engagement of all parts of the "chain" from breeders to users

3.4.1.5 Training, documentation and user collaboration in design

This involves working with breeders to discover what is useful to them in terms of website design, reports etc

Quote from Industry engagement

Process and communication pipeline with development and upgrades - new traits etc. all the way from software to report & use.

- Software better comm's
- Databases are hard to use
- I don't understand what the validation terms mean
- Indexes feel counter-intuitive black box
- Breeders' understanding of what services are available
- Value of services, info
- Process and communication pipeline with development and upgrades new traits etc. all the way from software to report & use
- Validation reports provide all info back, hard to tell what is new
- Overall goal needs to be very clearly spelt out
 - - maximise genetic gain
 - - maximise accuracy on animals
 - - lowest sustainable cost to users
- I need a mentor to get started practical info
- Uniform visual trait score. E.g. 1-5. 1 is worst, 5 is best
- How do I know if my issue is addressed?
- How the report should be presented to the end users
- New simpler ways to present results to industry
- Need documentation of process. Not reliant on one individual.
- Genomics office & SG communication → when genomic info is back, what sheep have tests pending
 - $\circ \rightarrow$ parentage tests (who, how many etc)



- Data just appears
- Difference between data submitted and what is said to have happened on the farm
 - management groups
 - \circ understanding their own data
- Ease of use. We often have unskilled labour so must be easy to learn & use & simple to teach
- Easy to use/understand reporting of relevant information
- Web knows which analysis a breeder needs once key log in
- Breeder end user education, guidelines
- Breeders don't know how data is being used
- Breeders don't know what data is needed
- Breeders' understanding of what services are available
- Improve grouping system of animals. Education
- Website is slow and hard to navigate
- Make information meaningful to clients

3.4.1.6 Database and website monitoring and management

This involves database sizing, web workload, backups, maintenance tasks. Also the current redesign of the database and processes to support current and into the future requirements. Also involves consideration of technologies which are fast to implement and database design which allows new traits, results etc

Quote from Industry engagement

Time effective maintenance & modification.

- Scheduled maintenance
 - $\circ \rightarrow$ website
 - $\circ \rightarrow$ analysis upgrades
- Increasing volume of data needs to be handled.
- Processing times getting longer
- Time it takes to upload (submit & validate) data to database Merinoselect
- Time effective maintenance & modification
- Bug fixes in MS take a long time
- Time from development to implementation
 - o we currently imp relatively quickly
 - \circ don't want to slow this down
 - avoid bottle necks
- Cost effective maintenance & modification
- Breeding values are complicated
- Speed analysis



- Increasing volume of data needs to be handled.
- Processing times getting longer
- Time it takes to upload (submit & validate) data to database Merinoselect
- Search site is slow and hard to use if not a seedstock breeder
- Website is too slow

3.4.1.7 Marketing, customer awareness

D

Issues identified in the workshop

• Facilitate a better understanding within industry of the benefits of selecting for genetic gain. Ie: create the lifestyle people want & guaranteed food supply

3.4.2 Solution

3.4.2.1 SG/AGBU database management team

A new cross SG/AGBU database management team whereby all changes are vetted, impact considered, approved and scheduled and adequately communicated with all stakeholders including software providers and breeders. A more robust process should mean less bugs which need to be fixed on an ad hoc basis. Where such a need arises, pre-agreed escalation processes will already be in place with MLA development teams to allow fast tracked bug fixes to be undertaken still within a robust testing enterprise level framework. DB management team are responsible to outsourcing to MLA, service providers etc. They are also the interface with the end users, ensuring adequate help is available to them either via documentation or one on one training (mostly over the phone).

3.4.2.2 Development and documentation management

The database management team defined in S#16 will oversee the collating a full set of documentation including detailed requirements and design specifications, all database elements and their definitions, training material pitched at various levels for end users. This process will also include breeders, producers and solution providers in the redesign of the website including them in focus groups, testing and enabling them as champions of the new website.

The database management team will also oversee the redesign of the database and website. This will include performance improvements.

3.4.2.3 Training

3.4.2.4 Issues Register

An issues register will exist within the web site which also allows breeders to log issues. The database management team defined in S#16 will monitor, reply to and update and provide feedback to users on the state of the issue. Allows for users to suggest new reports, updates to existing processes etc.



3.4.2.5 Promote/incentivise Solution providers

New structure, DB and APIs allow better software integration for more on farm device data collection/synchronisation. Need to find a way to promote/incentivise (by carrot or stick) software providers. Also involves providing clear documentation and requirements for valid data entry etc at the data collection end NOT just the database – i.e. stop the data corruption at the point of corruption. Clearer communication with software providers, breeders and other website users re website maintenance, analysis upgrades etc

3.4.3 Benefits

A new Data and Analysis Management Team – Overseeing design, documentation, training and ongoing management of issues, new requirements provide the following benefits:

- Vetted, collaborative process
- Standards, documentation & training developed for end users
- Coordinated change management



4 Timeline, process and detailed costings

4.1 Current state

Figure 2 depicts the current state of data flows within Sheep Genetics and between external organisations. It represents a complex set of processes involving multiple parties.



Figure 2 – Current state import-analysis-reporting Diagram

4.2 Stage 1

The purpose of stage 1 is to:

- Setup the Azure cloud network
- Scope out, build and migrate one new consolidated database
- Set up external organisations including AGBU, SIL, AMSEA to be part of the new Azure network

Current processes such as web site access and Pedigree Wizard uploads will remain unchanged from a user perspective a part from more reliable, faster access. This minimises the change control required and allows phase 1 to be focused on setting up the foundational framework going forward.

Stage 1 will involve the formation of the Sheep Genetics / AGBU database management team as well as the IT development team including Business Analysts, System Architects, database and web developers. The formation of this team will still need to be decided by Sheep Genetics and will involve operational as well as cost based decisions such as how the MLA Information company will be involved, the extent to which internal SG staff vs external contractors will be used or whether to outsource the entire project altogether.



The database design itself will determined as part of the detailed functional specifications developed throughout the early part of stage 1. But it would be expected to be based on a simple but flexible design including the suggested list of tables below:

- Animal
- Trait list of traits and what they do, accuracy levels, sources, \$ value (to perform financial analysis)
- Results list of results and what they do
- AnimalTraits and their associated attributes
- AnimalResults and their associated attributes
- TraitRules for input db. Produces anomaly reports
- Security rules around each trait and each measurement per animal

Figure 3 depicts the data flows within Sheep Genetics and between external organisations at the end of phase 1. Table 5 and Table 6 detail the tasks and estimated person days to complete stage 1 as well as the ongoing maintenance costs estimated for the Azure cloud network.



Figure 3 – End of Phase 1 import-analysis-reporting Diagram



Stage 1 Tasks - Database consolidation and migration to Azure	Due Date	Estimated
network		Person
		Days
Set up of Azure network	30-Jun-17	7
Databases	30-Jun-17	3
Web	30-Jun-17	2
Integration with other data	30-Jun-17	2
Consolidate LambPlan / MerinoSelect / Resource Flock into one new database	31-Oct-17	285
Mapping of traits, adjustments, results, processes	31-Aug-17	240
Create new databases	15-Sep-17	15
Migrate copy of data to new database	31-Oct-17	30
Setup migration tasks (to run incrementally), run and test		
Perform migration, set incremental migration tasks up to run manually		
AGBU Replication – From SG DB and Analysis Results -> SG Replica	30-Nov-17	60
Replicate Animals, Traits – SG->AGBU	30-Nov-17	15
This will replace the current OVIS Export file (SG->AGBU) process and require AGBU to adjust their OVIS run (Excludes		
Replicate Adjustments – AGBU->SG	30-Nov-17	15
This will affect the current AGBU own MySQL database process and require AGBU to adjust their OVIS run (Excludes		
Replicate Results – AGBU->Analysis Results	30-Nov-17	15
This will replace the current OVIS Result File process and require AGBU to adjust their OVIS run (Excludes AGBU time)		
Export to Analysis Results database		
Management of other affected processes	30-Nov-17	15
Genomics. These currently replicate information to AGBU's MySQL database. This process will need to be adjusted to replicate to the new SQL Replica (Excludes Sheep CRC Time)		
Resource Flock website – read from new SQL database (Excludes Sheep CRC Time)		
SIL – Azure Data Management Gateway installed on their servers and import export process adjusted (Excludes SIL time)		
AMSEA and DOHNE – Azure Data Management Gateway installed on their servers and import process adjusted (Excludes AMSEA and DOHNE time)		
ReMap databases for current API framework (current functionality only)	31-Jan-18	45
ReMap current API framework Based on current SG APIs and NLIS API framework	31-Dec-17	30



Map security and deploy via Azure API Management	31-Jan-18	15
ReMap databases for current websites	28-Feb-18	25
Sheep Genetics Website – Access database via new database This is an interim step before phase 2 is complete which includes extending the API framework	28-Feb-18	15
RamSelect – Point to new API framework (Excludes Sheep CRC time)	28-Feb-18	5
Resource Flock Website – Access database via new database (Excludes Sheep CRC time)	28-Feb-18	5
This is an interim step before phase 2 is complete which includes extending the API framework		
Creation of database administration website	30-Jan-18	125
Creation of website to allow SG / AGBI database administrators to add traits, set security access for end users, administer validation rules – pre/post analysis via web platform vs a reprogramming exercise	30-Jan-18	125
Migrate automated processes to use new API framework and database	28-Feb-18	60
MerinoSelect based batch system – Upload to new database (as well as existing one)	28-Feb-18	30
This is an interim step before phase 2 is complete which includes extending the API framework		
SGAR Reporting – Replicate process to extract data for reports from new database	28-Feb-18	30
This is an interim step before phase 2 is complete which includes extending the API framework		
Convert Pedigree Master process	28-Feb-18	35
Update to use copy files to a new Secure FTP location	28-Feb-18	5
Update current Merino Select batch system to accept and process Pedigree Master files from the secure FTP system	28-Feb-18	30
Convert Emailed files process	28-Feb-18	45
Set up process to move files to a Secure FTP location	28-Feb-18	15
Update current Merino Select batch system to accept and process Pedigree Master files from the secure FTP system	28-Feb-18	30
Finalisation - Testing and Cutover	31-Mar-18	110
Testing	31-Mar-18	90
Conduct OVIS runs against current and new databases and compare results. Note that as part of the database clean-up, it is expected that numbers will differ.		
Removing old websites and databases, repoint to new processes	31-Mar-18	20
Total	31-Mar-18	797

Table 5 Stage 1 tasks and estimates



Azure Component	Estimated monthly cost
	(AUD)
App service (2 X VMs with 4 cores,7 GB RAM, 50 GB storage)	\$955.21
Azure SQL Database (1 DB, 500 DTUs, 500 GB storage per DB)	\$2,624.94
Bandwidth (300 GB of outbound data from Azure per month)	\$51.85
API Management	\$890.38
Azure SQL Data warehouse (LambPlan/MerinoSelect/Resource	\$5,522.69
Flock) (600 DWU, 400 hour per month,2 TB storage)	
Data factory (~20 data movement per month)	\$83.43
Azure Automation	\$38.21
SQL Data warehouse(Genomics) (100 DWU,300 hours per	\$922.12
month,2 TB storage)	
Support Options (Includes Web incident submission, Phone	\$382.11
Support, Response time <2 hours)	
Estimated Monthly cost:	\$ 11,470.9 4

Table 6 Ongoing monthly estimated costs



4.3 Stage 2

Stage 2 builds on the foundational framework built in stage 1. This stage will heavily involve breeders, solution providers and other end users of either the web site or APIs firstly because many of the existing processes they are used will dramatically change and they need to be involved and helped through that change. And secondly because their buy-in is essential to the project. In fact, they will be key designers of the website and API requirements in conjunction with Sheep Genetics and AGBU.

Figure 4 depicts the data flows within Sheep Genetics and between external organisations at the end of stage 2. Table 7 detail the tasks and estimated person days to complete stage 2.



Figure 4 – End of Phase 2 import-analysis-reporting Diagram



Stage 2 Tasks - Consolidate website, new API framework and genomics migration	Due Date	Estimated Person Days
Migrate Genomics database	31-Mar-19	50
Migrate Genomics DB to Azure including remapping of all existing processes (Excludes Sheep CRC time)	31-Mar-19	40
Genomics replica to AGBU. Will require AGBU to adjust processes (Excludes AGBU time)	31-Mar-19	10
Extend API framework	31-Mar-19	180
Ensure all current website and data upload activity can occur via APIs	31-Mar-19	180
New Website	31-Mar-19	240
Create one new website connected via API network	31-Mar-19	60
Migrate functionality of Resource Flock website	31-Mar-19	0
Migrate functionality of Sheep Genomics website	31-Mar-19	60
Migrate SGAR Reporting functionality	31-Mar-19	60
Migrate Pedigree Master functionality to website	31-Mar-19	60
Total	31-Mar-19	470

Table 7 Stage 2 tasks and estimates



5 Change Management Planning

Managing the people side of change is critical to any major change initiative. Concerns raised by stakeholders in the Workshop can be addressed through a well-structured Change Management Plan, especially as the change is at an enterprise level. Engagement of all stakeholders including breeders, solution providers, AGBU, MLA and Sheep Genetics is essential. The following table shows the phases that individuals experience when undergoing significant change. The corresponding change management components are outlined and examples are included of how a dedicated Change Management Plan will assist Sheep Genetics in preparing, managing and reinforcing the changes.



Employee phases of change	Enablers	Description	Change Management Components
AWARENESS: of the need for change	Management communications	Workshop #1: Clear sponsorship, goals of workshop & DB rebuild project clearly communicated, workshop methodology enabled authentic feedback & valid data collection. Engagement achieved with all stakeholders at workshop. There was a clear	Communications Plan Sponsor Roadmap Coaching Plan
		sponsor/champion for the change.	Understand the nature of the change and the groups
	Stakeholder input	Fast follow-up after the workshop of the outputs kept people engaged and reminded them of their contribution.	impacted
	Marketplace changes	Stakeholders are aware of the benefits of changing the database as demonstrated by workshop comments	Creating the right sponsorship model & coalition
	Ready-access to information	Beyond the initial workshop report, a communications plan needs to be	Identifying risks
		their relevance to specific audiences must align to achieve engagement.	Developing special tactics
A Project Initiation wo from a Leaders communica	orkshop is held at which the DB hip/Sponsorship aspect, Project tion protocols within the Manag	Management Team is formed and includes SG, MLA & AGBU represent Management aspect and Change Management aspect. Change Mana gement Team and identifying Primary Sponsor/Champion and specific	tatives. The DB rebuild is considered gement may include defining roles and accountabilities.
Decide	Fear of job loss	Managed poorly, this element may result in high-risk situations with employees	Sponsor Boadman
DESIRE: to participate and support the change		Managed well, individuals may become proponents for change. Consideration and sensitivity are important approaches for each individual concerned.	Coaching Plan Resistance Management Plan
	Discontent with the current state	Major issues with current DB and related processes were highlighted by workshop participants (see Appendix for details). There is significant material here to support desire for change, however the key message is that support (both technical support and support for the overall change process) is needed.	Corrective action plans Transition management
	Imminent negative consequence	Stakeholders' highlighted the need for a contingency plan in place. This needs to be communicated and given consideration around how this fits into a deliberate communication plan.	
	Enhanced job security	Existing roles have an opportunity to be better utilized as a direct result of the DB rebuild. This needs to be identified with each individual and links made to how each benefit will flow on to other stakeholders as internal customers.	



	Enablers	Description	Change Management Components
	Affiliation and sense of belonging	Clearly message wins and benefits within a structured communications plan. Foster a sense of "esprit du corps".	
	Career advancement	Succession plans (as outlined by "hit by a bus" scenarios in the workshop) and career development must be supported.	
	Acquisition of power or position	Communication of data custodians must be clear.]
	Incentives or compensation	Anticipate gaps in participation from key stakeholders by providing solutions and incentives.	
	Trust and respect for leadership	Support strength of connection between leaders and stakeholders by building rapport and investing in 'emotional bank account'; further invest by demonstrating commitments "do what you say you're going to do".	
	Hope in the future state	Communication that conveys status of project and celebrates successes. Keep stakeholders aware of initial vision and mission.	
		WHAT THIS MIGHT LOOK LIKE:	
The project is under their contribution e Know	way. Focus groups are being hel mpowers these groups as key co vledge Capability and a represe	d that include breeders, solution providers and producers in order to b ontributors and champions of the new DB. Change Management Scopi ntative from the Dairy Industry has been engaged to share their experi	nest inform the rebuild. In addition, ng has revealed a gap in a group's iences and insights.
The project is under their contribution e Know	way. Focus groups are being hel mpowers these groups as key co vledge Capability and a represe Training and education	Id that include breeders, solution providers and producers in order to be ontributors and champions of the new DB. Change Management Scopi intative from the Dairy Industry has been engaged to share their experi-	pest inform the rebuild. In addition, ng has revealed a gap in a group's iences and insights. Coaching Plan
The project is under their contribution e Know KNOWLEDGE: on how to change	way. Focus groups are being hel mpowers these groups as key co wledge Capability and a represe Training and education Information access	Id that include breeders, solution providers and producers in order to be partributors and champions of the new DB. Change Management Scopi- natative from the Dairy Industry has been engaged to share their experi- Provide technological and resource support to those mostly impacted. Include ease of access to materials supporting change; this could include learning	pest inform the rebuild. In addition, ng has revealed a gap in a group's iences and insights. Coaching Plan Training Plan
The project is under their contribution e Know	way. Focus groups are being hel mpowers these groups as key co vledge Capability and a represer Training and education Information access	Id that include breeders, solution providers and producers in order to be contributors and champions of the new DB. Change Management Scopie intative from the Dairy Industry has been engaged to share their experience of the technological and resource support to those mostly impacted. Provide technological and resource support to those mostly impacted. Include ease of access to materials supporting change; this could include learning & development, training, training materials, one-on-one support etc.	eest inform the rebuild. In addition, ng has revealed a gap in a group's iences and insights. Coaching Plan Training Plan
The project is under their contribution e Know KNOWLEDGE: on how to change	way. Focus groups are being hel mpowers these groups as key co vledge Capability and a represen Training and education Information access Examples	Id that include breeders, solution providers and producers in order to be ontributors and champions of the new DB. Change Management Scopion tative from the Dairy Industry has been engaged to share their experi- Provide technological and resource support to those mostly impacted. Include ease of access to materials supporting change; this could include learning & development, training, training materials, one-on-one support etc. Provide examples on how other industries have successfully navigated similar change and where they are today (e.g. Dairy Industry)	est inform the rebuild. In addition, ng has revealed a gap in a group's iences and insights.
The project is under their contribution e Know	way. Focus groups are being hel mpowers these groups as key co vledge Capability and a represer Training and education Information access Examples Role models	Id that include breeders, solution providers and producers in order to be ontributors and champions of the new DB. Change Management Scopion intative from the Dairy Industry has been engaged to share their experi- Provide technological and resource support to those mostly impacted. Include ease of access to materials supporting change; this could include learning & development, training, training materials, one-on-one support etc. Provide examples on how other industries have successfully navigated similar change and where they are today (e.g. Dairy Industry) Include role models from successful case studies to support change.	est inform the rebuild. In addition, ng has revealed a gap in a group's iences and insights.



Employee	phases of change	Enablers	Description	Change Management Components
-	An audit has also sl	וסאה a particular group lacks A	bility Capability to use the new database, therefore the Training Plan h tailored training within their specific context.	as been adapted to provide more
ABILITY: to implement behaviours	: required skills and	Practice applying new skills or using new processes and tools Coaching Mentoring Removal of barriers	 Include frequent feedback and learning and development opportunities in initial stages to support uptake of new knowledge and skills. Identify appropriate and adequate coaching partners and sessions. Look to partner significantly impacted roles with industry mentors. Consider opportunities to remove barriers to change (e.g. create opportunities for breeders to engage with solution providers or other industry specialists) 	Coaching Plan Training Plan
-		Initial feedback has indicat	WHAT THIS MIGHT LOOK LIKE:	e website.
REINFORCE to sustain the	MENT: change	Incentives and rewards Compensation changes Celebrations Personal recognition	Implement corrective actions and celebrate successes	Communications Plan Sponsor Roadmap Coaching Plan Compliance Audit Reports Employee Feedback After Action Review
-			WHAT THIS MIGHT LOOK LIKE:	



Employee phases of change	Enablers	Description	Change Management Components
		An audit of Web Incident submissions	



6 Appendix

6.1 List of Tables

TABLE 1 RECOMMENDATIONS EXPLAINED	14
TABLE 2 ESTIMATED COSTS BY STAGE	
TABLE 3 PROJECT STAGE BREAKDOWN	16
TABLE 4 ESTIMATED DAYS REQUIRED AND ASSOCIATED COSTS - OPTIONS	17
TABLE 5 STAGE 1 TASKS AND ESTIMATES	51
TABLE 6 ONGOING MONTHLY ESTIMATED COSTS	52
TABLE 7 STAGE 2 TASKS AND ESTIMATES	54

6.2 List of figures

FIGURE 1 - SOLUTION ARCHITECTURAL DIAGRAM	8
FIGURE 2 – CURRENT STATE IMPORT-ANALYSIS-REPORTING DIAGRAM	48
FIGURE 3 – END OF PHASE 1 IMPORT-ANALYSIS-REPORTING DIAGRAM	49
FIGURE 4 – END OF PHASE 2 IMPORT-ANALYSIS-REPORTING DIAGRAM	53

6.3 MOQ Digital High Level Design Specifications

