



# Visualising Australasia's Soils: Extending the soil data federation

Final report for research project 2.3.002

**Peter Dahlhaus, Andrew MacLeod, Alison Ollerenshaw,  
Heath Gillett, Alicia Reynolds, Nathan Robinson, Megan  
Wong, Helen Thompson, Richard MacEwan, Paul Feely,  
Richard Archer, Kelly Hartwig, Indishe Senanayake  
& In-Young Yeo**

September 2024



## PROJECT PARTICIPANTS

Federation University Australia, Centre for eResearch and Digital innovation

University of Newcastle, School of Engineering & School of Environmental and Life Sciences

Southern Cross University, Faculty of Science and Engineering

Manaaki Whenua – Landcare Research, New Zealand

Birchip Cropping Group

Burdekin Productivity Services

Central West Farming Systems

Farmlink

The Facey Group

The Liebe Group

Mackillop Farm Management Group

North Central Catchment Management Group

Riverine Plains Inc.

Southern Farming Systems

WA No-Till Farming Association

West Midlands Group

Wheatbelt NRM

Wimmera Catchment Management Authority

## DISCLAIMER

Any opinions expressed in this document are those of the authors. They do not purport to reflect the opinions or views of the Soil CRC or its partners, agents or employees.

The Soil CRC gives no warranty or assurance and makes no representation as to the accuracy or reliability of any information or advice contained in this document, or that it is suitable for any intended use. The Soil CRC, its partners, agents and employees, disclaim any and all liability for any errors or omissions or in respect of anything or the consequences of anything done or omitted to be done in reliance upon the whole or any part of this document.

## PEER REVIEW STATEMENT

The Soil CRC recognises the value of knowledge exchange and the importance of objective peer review. It is committed to encouraging and supporting its research teams in this regard.

The author(s) confirm(s) that this document has been reviewed and approved by the project's steering committee and by its program leader. These reviewers evaluated its:

- originality
- methodology
- rigour
- compliance with ethical guidelines
- conclusions against results
- conformity with the principles of the [Australian Code for the Responsible Conduct of Research](#) (NHMRC 2018), and provided constructive feedback which was considered and addressed by the author(s).

## KEYWORDS

Soil data, Data management, FAIR data, Soil data sharing, Data value propositions.

## CITATION

Dahlhaus, P., Macleod, A., Ollerenshaw, A., Gillett, H., Reynolds, A., Robinson, N., Wong, M., Thompson, H., MacEwan, R., Feely, P., Archer, R., Hartwig, K., Senanayake, I. & Yeo, I-Y. (2024). Visualising Australasia's Soils: Extending the soil data federation. Phase 2 (2021 – 2024). Project 2.3.002 Final Report. Cooperative Research Centre for High Performance Soils, Callaghan, NSW.

## ATTRIBUTION

All material presented in this publication is provided under a Creative Commons Attribution 4.0 International licence ([www.creativecommons.org.au](http://www.creativecommons.org.au)), with the exception of the Soil CRC logo and any content identified as being owned by third parties. <https://creativecommons.org/licenses/by/4.0/>



Creative Commons Attribution 4.0 International Licence is a standard form licence agreement that allows you to copy, distribute, transmit and adapt this publication provided that you attribute the work. Unless otherwise stated, all images (including background images, icons and illustrations) are copyrighted by their original owners.

# CONTENTS

1. Introduction .....	6
2. Background.....	7
3. Methodology .....	12
4. Results.....	14
5. Discussion .....	38
6. Conclusion .....	42
7. Recommendations .....	43
8. Acknowledgements .....	45
9. References .....	46
APPENDIX A    Interview questions .....	49
APPENDIX B    Data sub-contracts.....	51
APPENDIX C    Data input templates.....	61
APPENDIX D    Data aggregator .....	76
APPENDIX E    VAS system API .....	108
APPENDIX F    Soil sensor data: architecture development .....	113
APPENDIX G    VAS Data Transfer Agreement .....	120

## EXECUTIVE SUMMARY

The Visualising Australasia's Soil (VAS) project aspires to create a research data federation, based on agreed governance and data stewardship frameworks, that allows relevant data from the private and public sectors to be discoverable through intuitive-to-use internet portals. The research is being undertaken over three phases and this document reports on Phase 2 of the project. The aims of this phase of the research are to: co-develop the functionality of the VAS online portal to demonstrate value for the participants (both farming groups and researchers), maximise the reuse potential of the federated soil data for research innovation and project participants' purposes, and transition to an agreed governance and business model that ensures that the future of the VAS data federation endures beyond the life of the Soil CRC. Research significantly progressed the first two aims, but the third aim has been deferred to Phase 3.

A variety of methods were implemented that reflect the eclectic research required to meet the stated goals. These include social science research to assess the project impact, technology systems for data aggregation and interoperability, soil data input and output tools, data stewardship, and literacy education and communication.

The most rewarding conclusion to take from this research to date is the clear demonstration that Australasian soils data, sourced from both the public and private sectors, can be made findable, accessible, interoperable and reusable (FAIR), and shared subject to the access rules set by the data custodians. The research has developed and implemented a functional and useful soil data federation system, accessible via the soil data portal, that was co-developed with project participants to meet their needs and provision standardised data to users, subject to the data custodian's consent. A clear value proposition for the farmer groups and catchment managers is access to a trusted, supported, web-based spatial soil data management system that suits their purposes and is relevant to their location.

As Phase 3 of VAS commenced in August 2024, research will focus on establishing a governance framework, business case and clear pragmatic value propositions for end-users in order to meet the goal of a creating self-sustaining and enduring Australasian soil data federation.

## OBJECTIVES

The end vision of this research is that VAS becomes an Australian soil data federation established by the Soil CRC as an ongoing component of a national and international soil data federation. Arguably the most significant benefit from the VAS project is that it can empower Australasian farmers, agricultural industries, researchers and governments with comprehensive soil and land data to support decisions. These data can be exported into existing and emerging tools for analysis and visualisation.

The potential highest value of VAS for researchers is in supplying well-described, structured, standardised, harmonised, place-based data with which to undertake novel research and make new discoveries. Data is the fuel for new decision support engines, such as machine learning and artificial intelligence models that could benefit the Australasian agricultural industries. Good quality data that is sourced from trusted repositories and is provided in an international standard machine-readable format, with reliable metadata, will be required to generate credible modelled output.

## RESULTS

The most obvious lesson from the research undertaken in VAS Phase 2 is that the challenges in creating and implementing a successful and enduring public–private soil data system are considerable. They include the social challenges of finding the value proposition for the project participants and end-users, the technical challenges of making it all work in an intuitive-to-use, seamless and effortless manner, and creating governance and business model that provides the project stakeholders with confidence and a vision that VAS will prosper beyond the lifetime of the Soil CRC.

A key finding of the social research is that if the VAS soil data federation can provide pragmatic value propositions through an intuitive-to-use web portal, then it will most likely continue to be supported and actively used by the data providers and consumers. Access to a trusted, supported, web-based spatial soil data management system is considered very useful and highly desired by the participating farmer groups and catchment managers.

At the commencement of Phase 2 of the VAS research there were 19 farmer groups, and two catchment management authorities engaged in the project. As time progressed, seven of the groups withdrew for various reasons, such as staff turnover, time constraints, concerns around data security and perceived risk, and scepticism that the project would endure or provide value for the return on their time investment.

As of July 2024, the VAS soil data aggregator holds contributions from 11 farmer groups, four catchment managers and two universities. This amounts to 55 datasets, representing 3,043 sites with 10,919 samples on which there are 218,375 observations. These data span the period from 1988 to 2024. In addition, soil moisture and temperature probe data have been included from one farmer group, amounting to over 600 million observations. Included in these datasets are some that have been made publicly available or available on request.

Although the project participants believe that better data leads to better decisions and can provide proof of soil stewardship for compliance with environmental standards and sustainable development goals, these potential rewards are yet to be demonstrated. Hence a strong focus in the next phase of the project is collaboration with participants willing to pilot use-cases and new initiatives, such as seamlessly feeding their soil data to existing tools (e.g. calculators, decision support, etc.) that add value to their data.

## NEXT STEPS

Phase 3 of the project commenced in August 2024 and will comprise five work packages, guided by a Community of Practice (VAS-CoP) made up of the project partners and end-users of the research.

WP1: Governance and business model

WP2: Data sharing

WP3: Linked data

WP4: Metadata harvesting

WP5: Enhanced reporting

## TIMING

Phase 3 commencement: 1 Aug 2024

VAS-CoP establishment: 30 Sep 2024

Shared soil and ancillary data: 30 Jun 2025

Online training courses: 31 Dec 2025

Data analysis tools: 30 Jun 2026

Trial governance model: 30 Jun 2026

Documented project impact: 1 Feb 2027

Documented architecture: 1 Feb 2027

Independent VAS: 30 Jun 2027

# 1. INTRODUCTION

The Visualising Australasia's Soil (VAS) project is being undertaken through the Co-operative Research Centre for High Performance Soils (Soil CRC) and aspires to create a research data federation, based on agreed governance and data stewardship frameworks, that allows relevant data from the private and public sectors to be discoverable through an intuitive-to-use internet portal. The project is being undertaken over three phases. This document reports on Phase 2 of the project.

## 1.1 PROJECT AIMS AND OBJECTIVES

Following on from Phase 1, which designed and implemented the technical and information architectures required to establish the VAS soil research data federation, the overall purpose of Phase 2 is to extend the data federation through data additions and data sharing.

### 1.1.1 Aims

Three aims were proposed at the commencement of Phase 2:

- Aim 1. Co-develop the functionality of the VAS online portal to demonstrate value for the participants (both farming groups and researchers).
- Aim 2. Maximise the reuse potential of the federated soil data for research innovation and project participants' purposes through enabling data that is discoverable, accessible and well-described (semantically interoperable) to both humans and machines.
- Aim 3. By 2024, transition to an agreed governance model that ensures that the future of the VAS data federation is supported by the members, and enduring.

### 1.1.2 Objectives

The aims were designed to lead to the long-term objectives of the VAS project, stated as:

- Long-term Objective 1. By providing a range of benefits for research, on-farm decision making and policy development, motivate Australian soils and agriculture data custodians to make their data Findable, Accessible, Interoperable and Reusable (FAIR).
- Long-term Objective 2. Align with other soil and agricultural data initiatives from the local (e.g., farming data co-operatives) to national (e.g., NatSoils, National Soils Information Framework, National Soil Monitoring Program) and international (e.g., Global Soil Partnership, ESIP soil data information cluster) scale that are aiming to maximise soil data discovery and reuse through the FAIR framework.
- Long-term Objective 3. Co-develop and implement an enduring Australasian soils knowledge system that is based on principles of data democracy, self-sustaining and inherently useful for research and education.

During Phase 2 of the project, not all of the aims were fully achieved, despite substantial progress being made. This report details the research progress and discusses the successes and failures of the research towards meeting the aims.

## 2. BACKGROUND

The VAS project commenced in February 2019 with the aspiration to ensure that all research within the Soil CRC builds on the best available and most current datasets, regardless of their pedigree. The intention is to motivate Australasian soils data custodians to make their data findable, accessible, interoperable and reusable (FAIR), by providing a range of benefits for research, on-farm decision making and policy development. It aligns with other soil data initiatives – local, national and international – to maximise soil data discovery and reuse by adopting the FAIR data guidelines. The ultimate goal is to co-develop and implement an enduring Australasian soils knowledge system that is based on principles of data democracy and is self-sustaining by offering pragmatic value propositions for end-users across the agricultural sector, including industry, research and education.

As the project progressed the global soil data ecosystem also developed, with the result that there are now many current initiatives working on soil data sharing, mainly delivered by government organisations and research agencies and focused on public sector open data. Within this ecosystem the VAS represents a novel research experiment in soil data supply that federates both public and private sector data across the Australasian agricultural sector, government agencies and researchers. Also novel is that VAS supports custodians of soils data (in both private and public sector groups, agencies and institutions) in maintaining their soil data collections using best practice FAIR data guidelines.

### 2.1 VAS PROJECT PHILOSOPHY

The underlying philosophy of the VAS project stems from the data – information – knowledge – wisdom model, or DIKW model (Ackoff 1989). In most representations it is pyramidal (Rowley 2007) with data at the base and wisdom at the top (Figure 2.1).

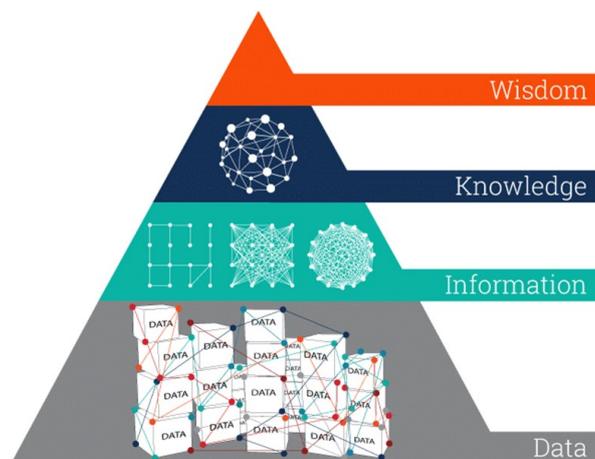


Figure 2.1: The DIKW model (after Ackoff 1989)

In these representations, data are unstructured items, typically measurements, observations or empirical facts, stored in and delivered by transaction processing systems (i.e. data infrastructures). In the second tier, information results from structuring these data into collections using management information systems and undertaking analyses. In the third tier, knowledge results from applying scientific laws, rules, algorithms and models to the information using decision support systems. Finally, wisdom at the top of the pyramid is gained from the application of the knowledge through expert systems.

In agricultural situations, conceptualising the DIKW model as one in which ‘better data leads to better decisions’ requires the insertion of human elements (Figure 2.2), since objective data becomes subjective wisdom, that is, subject to individual bias. To gain wisdom, the knowledge is applied to a decision (an action), which in turn leads to an outcome. By evaluating the

outcome of the decision (based on data-information-knowledge), subjective wisdom is gained by each individual actor and returned back to the cycle.

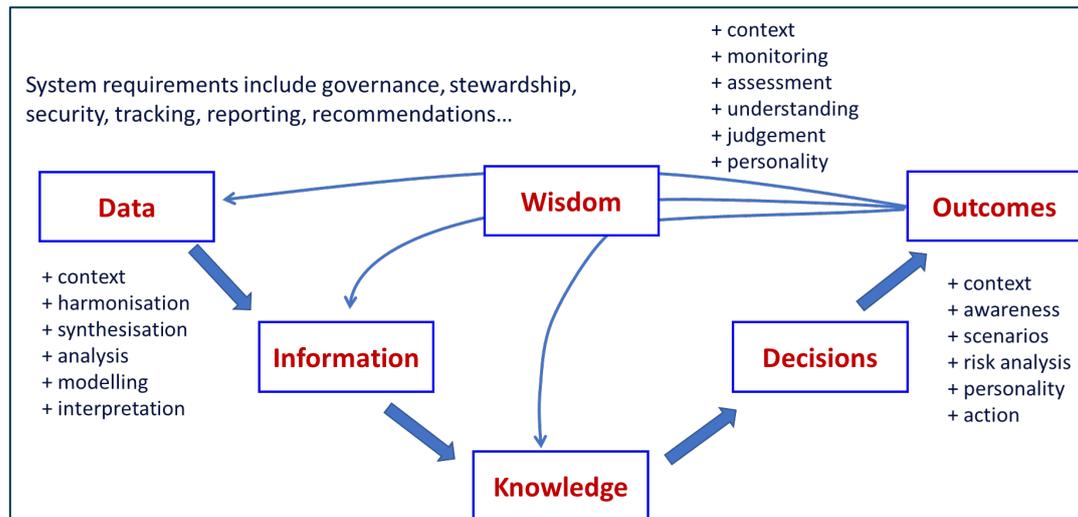


Figure 2.2: A conceptual model translating better data to better decisions, based on the DIKW model.

The assumption that better data leads to better decisions has been the foundation for mobilising a global ‘data revolution’ to make data more visible and equitable for sustainable development (United Nations 2014). Data is not only required to monitor progress towards the United Nations (UN) Sustainable Development Goals (SDGs) (United Nations 2015) and hold governments accountable, but it enables both the public and private sectors to make better informed decisions and close the gap between information-rich and information-poor societies. Globally, many governments have responded to the data revolution by increasing availability of open data served through publicly accessible data infrastructures and standards (Kitchin 2021).

In agriculture, examples include the Global Open Data for Agriculture and Nutrition (GODAN) initiative, which is committed to developing Open Data policies to support sustainable agriculture and food security across the globe (GODAN 2023). Initiated as a result of the G8 Summit in 2012, GODAN sprang from a global conference on Open Data for Agriculture. Similar initiatives are the Open Geospatial Consortium (OGC) Agriculture Domain Working Group (OGC 2017), and the Research Data Alliance Agricultural Data Interest Group (RDA 2018). Specific to soil data are the Global Soil Partnership Pillar 4 (FAO 2018) through their Global Soil Information System (GLOSIS), the International Union of Soil Sciences (IUSS) Soil Information Standards Working Group (IUSS 2020), the ISRIC World Soil Information project (ISRIC 2021), and the European Soil Data Centre (ESDAC), hosted by the European Commission's Joint Research Centre (JRC) (Panagos et al. 2022).

In the Australian context, the National Soils Strategy, a 20-year strategy on how Australia will value, manage and improve its soil, was launched in 2021 (DAWE 2021b). The Strategy included an objective to make Australian soil information and data available, leading to the development of the Australian Soil Information System (ANSIS)<sup>1</sup>, which was launched in June 2023. Designed and developed by CSIRO, ANSIS operates as a place to discover or find standardised soil data from multiple providers. VAS has a key role in interoperably exchanging data with ANSIS, according to the access controls allowed by the soil data custodians.

<sup>1</sup> <https://ansis.net/>

### **2.1.1.1 VAS project evolution**

The VAS project evolved over the past five years into three distinct phases of development:

Phase 1: Visualising Australasia's Soils: a Soil CRC interoperable spatial knowledge system.

Phase 2: Visualising Australasia's Soils: extending the soil data federation.

Phase 3: Visualising Australasia's Soils: building a legacy.

### **2.1.2 VAS Phase 1**

Phase 1 commenced in February 2019 and was reported in June 2021 (Dahlhaus et al. 2021). The project research participants were Federation University Australia (Lead), Manaaki Whenua – Landcare Research New Zealand, the University of Tasmania, and the University of Southern Queensland. Sixteen industry participants initially participated in Phase 1: 13 farmer groups, two catchment management authorities, and one agricultural service provider. Total project budget was \$1,179,060, made up of \$514,800 cash from the Soil CRC and \$664,260 in-kind from the participants.

The main project outputs from Phase 1 included:

- Establishing the project governance and Steering Committee.
- Social engagement to understand the expectations and intended uses for the VAS soil data federation. This involved a project kick-off meeting over four days and follow-up visits to each of the industry partners to explore their data and discuss the VAS research aims. The outcomes were reported by Sexton (2020) and contributed to the design of the technical system and information model.
- Co-design, development and implementation of the VAS soil data information model. The chosen model conforms to international data exchange standards and can aggregate a variety of soil (and other) data, value adding to the data by making it FAIR. The information model is fully documented in Appendix B of Dahlhaus et al. (2021) and is publicly available<sup>2</sup>.
- Building and implementing the soil data aggregator and VAS web portal as the public-facing component of the project. The public view has a selection of openly available soil data, and the login allows VAS project participants to view their own (non-public) data. Some functionality to select and filter data, graph data, and download data were included.
- A discussion paper to explore options for the VAS system governance, data governance and data stewardship was prepared and presented to the project Steering Committee (Dahlhaus 2019).
- Educational videos, newsletters, conference presentations and media communications to promote understanding of the VAS soil data federation.

The main lessons from the Phase 1 research were:

- Australasian soils data, sourced largely from the private sector, can be made FAIR and shared, subject to the access rules set by the data custodians.
- A clear value proposition for the project participants is access to a trusted, supported, web-based spatial soil data management system that suits their purposes and is relevant to their location.
- For almost all of the soil data owners or custodians, the value proposition for sharing their data remains obscure.

---

<sup>2</sup> <https://soilcra.com.au/technical-reports/visualising-australasia-soils/>

- Data stewardship and data literacy among soil data custodians is generally poor, especially with respect to metadata.
- The social architecture challenges in creating and implementing a successful and enduring public–private soil data system are considerably larger than anticipated.

### 2.1.3 VAS Phase 2

Phase 2 commenced in July 2021 and is reported in this document (to the end of July 2024). The project research participants were Federation University Australia (Lead), Manaaki Whenua – Landcare Research New Zealand, the University of Newcastle, and Southern Cross University. Initially, 21 industry participants participated in Phase 2: all 19 farmer groups of the Soil CRC, and two catchment management authorities. Total project budget was \$2,369,971, made up of \$944,297 cash from the Soil CRC and \$1,425,674 in-kind from the participants.

Both the end of Phase 1 and the commencement of Phase 2 were impacted by the Covid-19 global pandemic, which resulted in staff turnover among the project participants and the VAS project team. Manaaki Whenua – Landcare Research New Zealand discontinued their membership in the Soil CRC but have remained in the VAS project to a limited extent. Seven of the industry partners ceased participation in the project, for a variety of different reasons detailed in Section 4 of this report. Staff turnover and illness impacted members of the VAS team, including a seven-month sick leave period for the Project Leader, and retirement, resignation or redeployment of key team members. These unplanned events inevitably led to delays and project setbacks.

Significant events in the evolution of the VAS project during Phase 2 were the implementation of the National Soil Strategy and the subsequent initiatives announced in the 2021-22 Federal Budget. The initiatives included:

- The Historical Soil Data Capture Program that committed \$21.3 million to provide payment to farmers (or Data Owners) in exchange for sharing access to privately held soil data (DAWE 2023a). The intention was to increase the amount of soil data available for use by the Commonwealth and public relating to soil condition, and to contribute soil data to a national soil data repository (i.e. ANSIS).
- The Pilot Soil Monitoring and Incentives Program that committed \$54.4 million over two years to provide incentives for approximately 5,400 Land Managers via subsidised costs for soil tests, in exchange for Land Managers sharing their test data in agreement with data sharing terms provided by the Department of Agriculture, Water and the Environment (DAWE 2023b).
- The commitment of \$15M for the development of ANSIS.

These initiatives aspired to support the National Soil Strategy by increasing the amount of data available to assess the nation’s soil condition, both historic and current. It also aimed to support Land Managers with the soil information they need to make informed decisions to increase productivity and profitability and monitor soil health, as well as improve the quality of data from soil testing (DAWE 2021a, 2023a, 2023b).

These Australian Government initiatives coincided with a growing global awareness regarding the value of data and the potential implications of data misuse and theft (e.g., Brown et al. 2023; Kaur et al. 2022; Zhang et al. 2021), and the emergence of farm data codes of practice (e.g. NFF 2020, 2023). In addition, the rapid uptake of artificial intelligence (AI) tools, especially neural network-based language prediction models built on the transformer architecture (such as GPT – Generative Pre-trained Transformers) has raised the VAS participants’ awareness of their soil data as a commodity, that may have monetary or commercial value for use by others.

### 2.1.4 VAS Phase 3

The third and final phase of the VAS project commenced in August 2024 and finishes in June 2027. The project lead remains with Federation University, and the University of Newcastle remains a collaborating research partner. All current industry participants have been invited to remain with the project and can elect to participate in a Community of Practice (VAS-CoP) to co-develop VAS portal functionality based on their use-cases. The total project budget is \$2,401,387, made up of \$800,000 cash from the Soil CRC and \$1,601,387 in-kind from the participants.

The overall purpose of Phase 3 is to *‘...create an enduring soil research data federation, based on agreed data stewardship and governance frameworks, that allows Australasian soils data from the private and public sectors to be discoverable, accessible and useable to Soil CRC participants through an intuitive-to-use internet portal. The research is guided by a Community of Practice to ensure a clear value proposition for all industry and research stakeholders. It is intended that the legacy VAS system will remain available to the soil research data federation members beyond the Soil CRC term.’*

The stated aims of Phase 3 are:

Aim 1. Develop and implement a business case, value propositions and governance framework for post-CRC management of the soil data federation. The business case requires agreement from all participants, and be based on strong value propositions that demonstrate the unique role that VAS has in the ever-evolving international and national agricultural data sharing ecosystem. Implementing the business case and governance framework before the end of this project will test the potential for an enduring self-sustaining VAS system beyond the life of the CRC.

Aim 2. Implement enhancements for agricultural decision-making support using soil data analytics and integration via new and developing artificial intelligence tools. This aim intends to build the value propositions that will be required to demonstrate the unique place that VAS has in the international e-research data-sharing ecosystem and provide the foundation for the business case to sustain the project beyond the end of this project.

Aim 3. Implement auto-data-capture capacity for soils information through seamless metadata harvesting to improve data management via the self-serve system. The intention of this aim is to make the VAS system as simple and seamless to use, as is possible. If the VAS system is easy and intuitive to use, it has the best chance at succeeding as an enduring technology.

Hence, the foremost goal of Phase 3 is to transform the current VAS research into an enduring technology platform that is a valued and valuable component of the Australian soil data and information ecosystem.

## 3. METHODOLOGY

A variety of methods were implemented in Phase 2 of the VAS project that reflect the eclectic research required to meet the stated goals. These have been grouped as follows:

- Social science research: project impact
- Technology: data aggregation and interoperability
- Soil data input (data stewardship)
- Soil data output (reporting tools)
- Soil data literacy (education and communication)

### 3.1 SOCIAL SCIENCE RESEARCH: PROJECT IMPACT

Between 2022 and 2024, semi-structured interviews were conducted with staff ( $n = 18$ ) from 14 industry partner organisations. Through in-depth interviews, participants were asked about their involvement with the VAS project, their use and interactions with VAS, their perceptions of VAS and its functionality, and its contribution to soils data sharing. Questions were also asked about the challenges and opportunities for data sharing using VAS, whether the project had met expectations, and for suggestions for improving VAS and facilitating future data sharing. Two versions of the same interview schedule were prepared. One interview schedule comprised 12 questions and was used with organisations that had partnered on the VAS project since Phase 1. Another version of the interview schedule, containing fewer questions ( $n = 8$ ), was adapted for staff at organisations that had commenced partnership on VAS during Phase 2. Both interview schedules are included in Appendix A.

The impact researchers met with project partners during online inaugural meetings and followed up with email invitations to key contacts (CEOs, managers, project staff) at each of these partner organisations to participate in the research. The first set of interviews ( $n = 9$ ) were conducted between February to September 2022. Interviews ( $n = 5$ ) with remaining project partners were conducted between December 2023 and April 2024.

The interviews were recorded and later transcribed in preparation for coding and thematic analysis. Interviews were first examined for consistent patterns and codes within the data using the NVivo 12 software (QSR International). Sentences and statements from the participant narrative were coded to identify meaningful and consistent ideas. Themes were then generated from the coded data for broader meaning.

This method is consistent with structural approaches for analysis when using qualitative data, whereby key phrases or content from interviews are aligned with the research focus and questions. This approach is well suited to semi-structured methods of data collection and hypothesis testing (Saldana 2016). Analysis of the interview data was conducted independently by two experienced qualitative researchers. Themes were cross-referenced to ensure reliability and accuracy.

Approval to conduct this research was received from the Federation University Australia's Human Research Ethics Committee (project approval number: A16-137).

### 3.2 TECHNOLOGY: DATA AGGREGATION & INTEROPERABILITY

The technology that underpins the VAS project is a critical element in the research success to date. Much of this was achieved and documented in Phase 1 (Dahlhaus et al. 2021). The methods adopted in Phase 2 were largely implementing system improvements that were required to support the current users and new requirements for data interoperability with ANSIS, as well as maintain the currency of the system (technology updates, bug fixes, etc.). These methods are fully documented in Milestone Report 14.3 (MacLeod et al. 2024).

The main changes were alterations to the way that data is delivered to other systems via an application programming interface (API). Other technology improvements were made to the VAS aggregator architecture, the self-serve system, and improved data reporting.

### **3.3 SOIL DATA INPUT (DATA STEWARDSHIP)**

Soil data input and improving data stewardship has been a significant focus in Phase 2, building on methods established in Phase 1 (Dahlhaus et al. 2021), with a rigorous approach to soil data collection, metadata collection, and mapping data to the international standards.

Soil data collection was improved through the application of detailed sub-contracts with the project industry partners who elected to participate, in which partners nominated specific soil and soil sensor datasets that would be provisioned, the extension activities that would be undertaken and periodic reporting responsibilities. Partners were paid up to \$15,000 to cover their time and costs. The generic data sub-contract is included in Appendix B. Additional data added to the VAS data aggregator includes Federation University soil datasets, and non-Soil CRC members who are affiliated with Soil CRC industry groups and/or contributors of ANSIS third tier data. These include Corangamite Catchment Management Authority (CMA) and Glenelg Hopkins CMA, both of which collaborate closely with Southern Farming Systems (SFS), Monash University, and the University of Sydney who are contributors of ANSIS data.

Metadata collection was improved by the co-development and implementation of improved data input templates in the self-serve system (Appendix C). They are designed to reduce the iterative data checking between the VAS team and the industry data custodians required to clarify data stewardship details such as ownership, custodianship, licensing, funders, project names, collection dates, site names, sample names, laboratory, and analytical procedures.

Mapping the soil data loaded into the VAS aggregator by the partner contributors to the interoperability standards has been significantly improved, largely in response to the changing global soil data ecosystem. The development and implementation of ANSIS has improved the ontology for Australian soil data and established community-controlled vocabularies (Cox and Wong 2022). The VAS information model and the information model adopted by the ANSIS project<sup>3</sup> are very similar, specifically in the ANSIS Ontology<sup>4</sup>.

An initial use-case to include SFS soil moisture and temperature probe data in the VAS portal was undertaken. The methods built on a previous project undertaken by CeRDI for SFS, which established a portal known as Probetrax in which the public can view the current plant available water (PAW) as a percentage value (%) at each probe site as a coloured dot. Additional methods implemented include data from Agriculture Victoria's soil moisture probe network in the VAS portal. Methods for adding soil probe data into VAS are still evolving.

### **3.4 SOIL DATA OUTPUT (REPORTING TOOLS)**

Improvements were made to reporting the metadata summaries of the soil datasets uploaded via the self-serve system by each participant. This included basic metadata on each dataset and statistics (count, average, mean, maximum, minimum, standard deviation) on the content of each dataset. Improved displays for each dataset and improved graphing tools were also developed.

### **3.5 SOIL DATA LITERACY (EDUCATION AND COMMUNICATION)**

The VAS web portal was upgraded with a new web address and a series of content pages. For the users of the VAS aggregator, a series of 'how to use' short videos were produced and published on the website. These web pages and videos remain works in progress.

---

<sup>3</sup> <https://ansis.net/data/information-model/>

<sup>4</sup> <https://raw.githack.com/ANZSoilData/def-au-domain/main/doc/domain.html>

## 4. RESULTS

The results of the research in the Phase 2 of the VAS project are detailed below.

### 4.1 VAS PROJECT IMPACT RESEARCH

The data collected by the social research interviews (Section 3.1; Appendix A) classified into four key themes and related ideas (Figure 4.1), which are presented in the following sections.

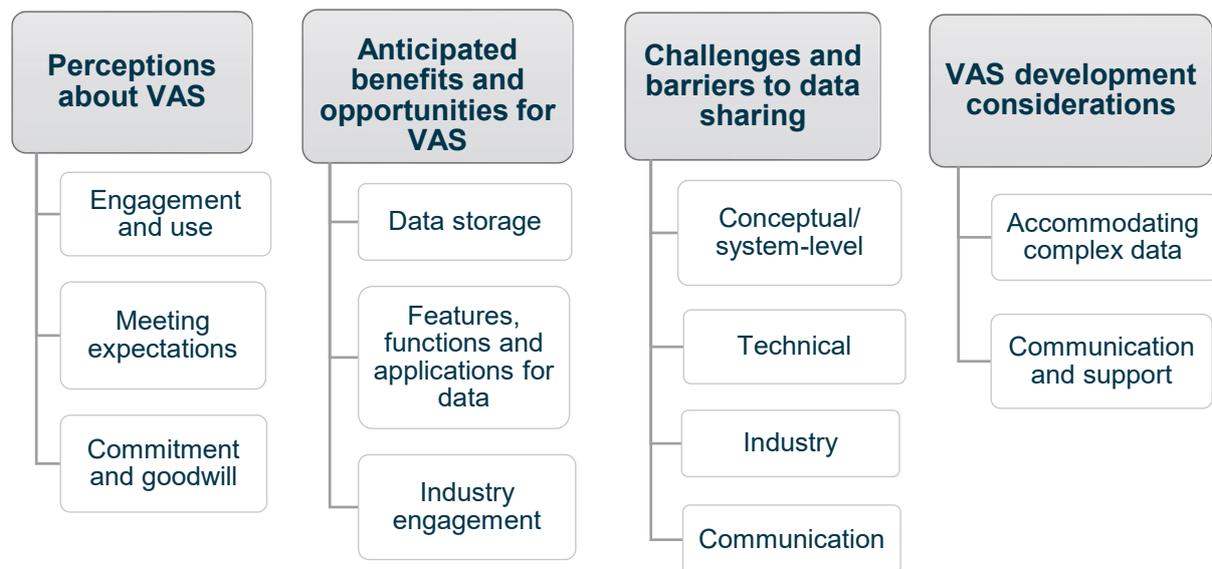


Figure 4.1: Themes and ideas identified through analysis of partner interviews.

#### 4.1.1 Perceptions about VAS

Three factors informed participants' perceptions about VAS. These included:

- Engagement and usage of VAS.
- Whether VAS was meeting their expectations.
- Commitment and goodwill towards the project.

These are examined further, using quotes and excerpts from participant interviews to demonstrate and support these ideas.

##### 4.1.1.1 Engagement and use

During the first round of interviews, most participants reported low to moderate interaction with VAS, and one participant had not yet used VAS. The main reasons for not using VAS included:

- inability to interact with VAS due to their own time restraints.
- lack of available data on VAS from their organisation.
- lack of experience in using and sharing data via VAS restricted their ability to interact with VAS.

The following quotes are representative of these points:

*Uh, but I haven't looked at it since. It's probably, a contribution of the other project work that we've been doing and the lack of, I guess, additional free time that we ... probably a lot of people experience to be able to delve into the depths of it. And I guess also I was sort of waiting for it to be to have more in it and our data in it before sort of having a good play through it (Interview participant, Time 1).*

*So really looking at the portal from our end, at the moment we're just clicking on a site and seeing the data. There's sort of no interpretation or added visualisation, it's just the visualisation is a dot on a map and a table with the data in it. So that's all we've, you know, we still feel that we're pretty much at the infancy of using the site as it stands for us at the moment (Interview participant, Time 1).*

*We probably haven't been using it and the reason, we haven't been given a login or anything like that and I think that's what you need. I have looked at the site and I see that there's some data that we've collected through our monitoring trials that have been put on there, that was previously on the <name of another> website ... but when you actually click on the data there's no data in it (Interview participant, Time 1).*

*I can go to, you know, a particular grower's place and instantly see where all the soil tests were taken and very quickly click on a dot and see it. So, the information is more accessible than our current storage system. It's just that we're not really using it (Interview participant, Time 1).*

During the first round of interviews three interviewees commented that their involvement with the VAS project had prompted their organisation to reflect upon and undergo preparations and activities relevant to the research data they now collect. One informant acknowledged that their involvement with VAS, and the insights gained through the project, had assisted them with other data sharing projects they were now involved with. These insights are captured in the following quotes:

*The data sharing practices that I have been exposed to from being a part of the VAS project has benefited me in navigating other projects that are to do with data sharing and what is classified as publicly available data or from projects that are funded by particular organisations that within the contract should, would have the fact that it's written as publicly available data (Interview participant, Time 1).*

*Up to this point, the main thing I've been using it for is, I suppose, as a bit of a point of reference and a bit of a source of inspiration to help our team think through some of the issues and the challenges we been dealing with as part of a soil moisture probe and weather station network project we've been rolling out (Interview participant, Time 1).*

Thematic analysis from the most recent round of interviews reveals new insights from participants about VAS. Interviewees confirmed having a foundational understanding of VAS and had conducted a preliminary review of it. Moreover, some interviewees had commenced entering their organisation's soils data into VAS.

*I've had a little bit of a play looking around at the different layers and that sort of thing, but I haven't necessarily understood well how I might use it (Interview participant, Time 2).*

*I've obviously looked at the website, we've entered some data into it, that sort of thing so we're at a fairly basic level, but and that's sort of what the project's about a bit is us getting familiar with it (Interview participant, Time 2).*

#### **4.1.1.2 Meeting expectations**

Despite varied interactions with VAS, most participants in the first round of interviews agreed that the project had, to date, met their expectations. The build and launch of VAS were primary reasons given by informants. For others, while the initial expectations of VAS had been met, the subsequent issues they had encountered with VAS, such as challenges associated with

its use and functionality, had somewhat tempered their feedback and expectations. These ideas are highlighted in the following representative quotes:

*I guess it did meet my expectations in terms of a portal that can store data, soils data, and provide a mapping interface, like it's got a spatial component to it. So yeah, that ticked boxes for me. In terms of its functionality that's probably not what I expected. It's yeah, it ... I don't know how to describe it, but it just seems clunky to use and yeah, difficult to use and not as smooth and as I would have expected it to be (Interview participant, Time 1).*

*It was really getting something built in the first instance. There is something there now that wasn't there before. So, I'm happy that we've been successful to date. This next stage now is all around building it to do something (Interview participant, Time 1).*

*I think the expectation has been met that there's a built system (Interview participant, Time 1).*

During the first round of interviews, interviewees indicated that their expectations for VAS was now focused on the outcomes associated with the current phase of VAS developments. For some informants, it is only at the conclusion of the current expansion and development phase that the overall opportunities for VAS, and the expectations for data sharing would be fully realised. These ideas are captured in the following comments:

*I suppose the second phase is where the most exciting phase is for me because we had to get a database setup and working. So we've done that. So it has it has met expectations, but the next steps are probably the most important bit. And that's what we haven't actually done here (Interview participant, Time 1).*

*I mean, the biggest challenge with building and spending so much money on one of these things is what happens at end of project to this and if the more we put in the, the more we've got it then getting [the data] out if the project ends and the product doesn't go anywhere (Interview participant, Time 1).*

In the second round of interviews there was widespread agreement that VAS was meeting expectations. For some, their expectations for VAS had been met through its ease of use, and the tools available for displaying and mapping soils data across regions. The expectations for VAS appeared strongest amongst interviewees that (a) had been 'long-term' partners on VAS (since Phase 1 of the project), and (b) reported higher usage of VAS. Through their engagement with VAS, and by entering soils data into the system, a deeper understanding and appreciation of VAS appears to have formed. These ideas are evident in the following statements:

*I've been fortunate to see [VAS] develop from its inception so I guess I kind of knew what its capability was, and I guess my expectation was that it was functional and usable, which it is. So I guess it has met those expectations (Interview participant, Time 2).*

*My first thought was that it's actually really good. You know, I like the layers of data that that are in there. It's really easy to visualise the soils because it's all mapped out and you can click into the different locations to look at the different parameters so that's really good (Interview participant, Time 2).*

A small subset of participants across both interview rounds commented that their expectations for the project had not been met or had only partially been met. In the first round of interviews, the reasons given include a lack of achievement linked to milestone developments for VAS. It also included a lack of functionality and methods quantifying the value for data, that this partner had anticipated but not yet achieved. In the second round of interviews, the expectations for VAS by one informant had not been fully realised. However, this same informant acknowledged that the contribution of VAS for other stakeholder groups was a strength of the project. These points are reflected in the following quotes:

*Yeah, I guess expectations are tricky to manage on people's behalf and so I guess we're three years into the project and I don't think my expectations have been met really. But again, maybe*

*I had too high an expectation of what might have been delivered. I had hoped that, three years in that I would have more than a dot on the map with the, with the table of data there. I'd hoped that we would have some traffic lights or some, a bit more of that visualisation, that sort of feedback about what levels, you know, some of the soil test results are at (Interview participant, Time 1).*

*I can't see a huge amount of benefit that we are going to get from VAS directly, but it's worthwhile us being involved with because anything that helps our farmers is something that that we have bought that we have we will stand behind (Interview participant, Time 2).*

#### **4.1.1.3 Commitment and goodwill**

Ongoing commitment for, and goodwill towards, VAS was consistently documented in the first round of interviews. The reasons for this are associated with the potential value and expectations for VAS that have been largely facilitated through discussions between the VAS team and project partners. There was also a strong sense of anticipation and achievement in the skills of the VAS team to achieve success with VAS. This is captured in the following interview excerpts:

*But the other part for me ... the soils database is gold, in a way, so even if the farmers don't use it, it's useful for the government, useful for the universities, and CRCs and everything else (Interview participant, Time 1).*

*...one of the things I've always liked about, the potential about VAS is to pull a lot of the other old sort of, a lot of the general datasets that are in the states that, you know, it could give some sort of, just that little bit extra information. I think there's something like over 4,000 maps and datasets that are sitting in government departments that, you know, if you had access to that, then it just adds all of these different layers. And so what is one of my thoughts as around looking at having as many layers as we can, and then I actually being able to feed those layers into something like an AI or machine learning type scenario in the future because it, you know with all sorts of machine learning, it doesn't matter how much data you've got, it's never enough. So, having all of these extra layers may just add little percentage increments onto the outputs that you'll get, which will, you know, just enrich our knowledge about our farming systems (Interview participant, Time 1).*

One informant described their experience during the first phase of VAS as having established a baseline for understanding the anticipated benefits and challenges associated with their involvement in the extension phase of VAS:

*I think in the delivery of the first phase and getting something built, there were a lot of teething things that we had to go through and learn. I think both we learned and then also, you know, [the VAS project] team has also learned as well. And so, going into the next phase I think again we're a lot wiser and we kind of know where we need to land, what challenges we have to conquer (Interview participant, Time 1).*

#### **4.1.2 Anticipated benefits and opportunities for VAS**

Informants frequently referred to the prospective opportunities and anticipated benefits associated with the development of VAS. Three key benefits and opportunities for VAS were identified during both rounds of interviews, including: (a) data storage; (b) features, functions and applications for data; and (c) industry engagement. Each of these ideas will be examined in detail below.

##### **4.1.2.1 Data storage**

The majority of informants across all interviews indicated that the data storage and data management capabilities of VAS are a key benefit and outcome for VAS. The capacity to store new and existing data within VAS was important in supporting their own organisation's data storage needs, and overarching research capabilities. Interviewees, particularly in the recent round of interviews, valued having their organisations' soil data on VAS and accessing this

through a centralised system whereby their soils data would be available in the future. The tools available within VAS to identify key trends in the soils data over time will deliver broad benefit to organisations and industry, enabling the monitoring of soil changes over time. These ideas are represented in the following interview extracts:

*One of the ... levels that we're looking at is being able to organise and recall our data and being able to make that usable on a spatial scale (Interview participant, Time 1).*

*...it'd just be so nice for us to have a single point where everything's housed and make us realise where our gaps are in terms of what we're recording and how we're recording it. Which is a good thing and possibly highlight some of our shortcomings (Interview participant, Time 1).*

*We're also looking at building in a soil sampling program around the moisture probe network. Again, that's something that could potentially be linked into the VAS portal as that unfolds (Interview participant, Time 1).*

*Our own operations require some good systems to store soil data. So, and I think, you know, we've been going for many years as an organisation. I don't know how many soil tests we would have taken over that time, but a lot of them get lost ... as the project finishes or whatever ... it would be nice to have something that keeps all that together and in one spot. So, we can go back, you know, 20 years later and see how that soil has changed over time... (Interview participant, Time 2).*

*Storage and management of data. But then also the ability to pull out trends and potential ideas for projects going forward and having that dataset to be able to clarify and support applications that are in the fields that align to soil health and things like that (Interview participant, Time 2).*

#### **4.1.2.2 Features, functions and applications for data**

In the first round of interviews, some participants referred to the practical benefits associated with VAS. This includes the benefits of using the VAS functionality and technology to search for specific soils data:

*I guess I'm looking at the functionality of it long term and in terms of how, whether it's a time saver ... in terms of how we use, collect and collate our soil data at the moment. So, I'm very much driven by expectations at the technical end of things to this point (Interview participant, Time 1).*

Several interviewees during the first round of interviews anticipated using VAS for data extrapolation and application across a range of organisation-specific activities. For some informants, the accessibility to more detailed data and information offered the potential for new insights (through interrogation of organisational data) to inform their research and data-relevant activities, including research reporting and funding applications. It was also noted also that the data could be used to convey key soil information to their member farmers. Using data available through VAS for these informants will provide additional mechanisms for communication and information sharing within the agriculture sector. These ideas are represented in the following interview quotes:

*So I mean, we'd love to be able to get to a point that we can take it to the community and say to farmers, look, this is the state of the soils in your area at the moment and this is how they've changed over time, to give them an idea of how their practices might be impacting the state of their soils, and then also to be able to use it to report to funding organisations so that we can say this is the state of our soil at the moment, and these, you know, we need some funds to address these issues or we could also then report back to say look at the work that we've done and how we've changed the state of our soils as a result of the funding that we've been given. So, yeah, it would certainly make our reporting and communication of soil data to locally, nationally, much easier or robust, and yeah (Interview participant, Time 1).*

*In a perfect world, if you had a lot of data on there you could say I generate profile data of this and this soil type straight down the profile or pH down the profile and you could do all that on*

*the platform. I don't know if I'll ever get to that point, but I spend a lot of time on Excel manipulating stuff at times so potentially you could do that on the platform (Interview participant, Time 1).*

*I think that there'll be potentially, if people see the data that we've got, we might get projects out of it. That we could attract new funding opportunities I suppose. Because I guess it shows that we're active in the area where the people come to if they're wanting soil data collected (Interview participant, Time 1).*

Participants in the latest round of interviews shared insights about the perceived contribution of VAS. These included the provision of soils management and the security of having their soils data in VAS. The application of VAS for discoveries in the data was identified as having broad benefit for the organisation, and the industry. Creating reports and visualising their soils data were additional benefits of VAS. These ideas are represented in the following quotes:

*I think management, security, but also just visualisation. Reporting. Like being able to just see all that data, you know, pull out pH or pull out organic carbon or, you know, just one element and be able to report or visualise where they are. I think that will be quite important for reporting (Interview participant, Time 2).*

*I think that our contribution to this project will have bigger benefits for our producers at the end of the day, than for us, because even if their data is being used for research, that research will hopefully come back to benefit farmers (Interview participant, Time 2).*

During both interview rounds, participants outlined their expectations for VAS. The opportunity to conduct extensive reviews of their own data was an organisational benefit that would provide insights that would inform the future research and planning activities. Informants also commented on the potential for VAS to enabling knowledge transfer within organisations, and between organisational staff:

*Just from a data storage consistency perspective, within the organisation, I think that'll be really beneficial to just troll through everything and make sure that we've got as much as we can and that we can then re-interrogate it later on (Interview participant, Time 1).*

*I'll say I was interested in this program mainly because of the visualisation part of it. Like we get all these datasets and stuff, and they just sit in some database, and nobody ever looks at them, and having a tool you can go into and actually see what's happening, I think that's quite unique (Interview participant, Time 1).*

*... some people have a general knowledge of what's happening within soils and stuff like that, but there's new stuff. There's changeover of staff, all that sort of transfer of knowledge becomes a lot simplified when you have this sort of dataset, you can pull on and easily access, then it's ready in there to go. Yeah, it's not lost when someone leaves (Interview participant, Time 2).*

The efficiencies associated with sharing data in VAS was another identified benefit. This was highly regarded by organisations that collaborate on projects involving multiple stakeholders:

*... we're always very collaborative in what we do, we're working with different project partners, and I guess an example of that is so with a soil project where <name of organisation> are involved. Lot of the file sharing is just done via email. It's primitive but I think if we could have this place, this sort of platform as storage for that, it would make it a lot easier to share ... rather than back and forth with email (Interview participant, Time 2).*

### **4.1.2.3 Industry engagement**

Interviewees identified opportunities for industry engagement through VAS. The contribution of generating industry-relevant information and the prospect for industry engagement, including research collaborations, was a benefit associated with VAS:

*... when we are designing projects, and we're usually trying to access grant funds to do that, [VAS] could be very helpful to us in potentially identifying partners in different parts of Australia*

*that might be worthwhile working with, because if we can see that they have similar soils and things like that, that might help us to develop an idea about what other areas have similar problems that need to be addressed to the ones that that we do (Interview participant, Time 2).*

For some, the project was an opportunity to engage with farmer-stakeholders (particularly the member-base at the partner organisation level) about VAS. This would necessitate the involvement of farmers to supply and/or enter their data on VAS. As evident in the following quotes, establishing farmer trust in VAS is an important consideration for partners in relation to the project's growth and expansion:

*Me personally, get farmers involved, and get them to have some level of trust with using the tool (Interview participant, Time 1).*

*... hopefully again through this little project in our region, we're helping to contribute in a positive way to encouraging more people to think more positive way about sharing various data from their property... it's a major issue, and I think sometimes when big projects (are) conceived not enough time and energy is put into just nutting out exactly how farmers are going to be engaged and supported to. Think through the issues of data sharing (Interview participant, Time 1).*

*I'm always looking for things that might be of interest to our farmers or can be provide some value. And I guess when this project first came up, I thought ... it's going to be really useful and valuable (Interview participant, Time 2).*

At a broader industry level, VAS provides an opportunity for involvement with state and national government agricultural organisations through the sharing and accessing of industry-relevant data. Several interviewees commented on their expectations in accessing data from other organisations (including data from government agencies), through access to VAS. This would fulfill an important need for some, in providing them with access to industry-relevant data and information. These ideas are highlighted in the following representative quotes:

*One is accessibility, having data and the different layers, and one of the things I've always liked about the potential about VAS is to pull a lot of the other old, sort of, a lot of the general datasets that are in the states that, you know, it could give some sort of, just that little bit extra information. I think there's something like over 4,000 maps and datasets that are sitting in government departments that, you know, if you had access to that, then it just adds all of these different layers (Interview participant, Time 1).*

*So, I think if the VAS project can ... continue to be developed with that intent and over time evolve in a way that it does, hopefully link seamlessly with other things like the ... other Australian government data gathering projects ... and other bits and pieces that Agriculture Victoria are trying to do (Interview participant, Time 1).*

*I'd still like to see perhaps some more connection with perhaps our state agencies, and I know they have got some data in there, but things like soil moisture and that sort of thing (Interview participant, Time 2).*

### **4.1.3 Challenges and barriers to data sharing**

Without exception, all informants identified multiple barriers associated with the VAS project. These are categorised into four broad areas: conceptual/system-level; technical; industry; and communication. Key ideas representing each of these areas is captured in Figure 4.2.

During the first round of interviews numerous barriers were identified. These were wide ranging and represented actual and perceived barriers with VAS. Many of the reported barriers were representative of, and an expected outcome linked to, the current and ongoing developments of VAS and its technologies. Fewer challenges were raised during the second round of interviews. In this section, a brief exploration of the barriers is presented. These insights are supported by representative quotes included in the Milestone Report associated with this research (Ollerenshaw et al. 2024).

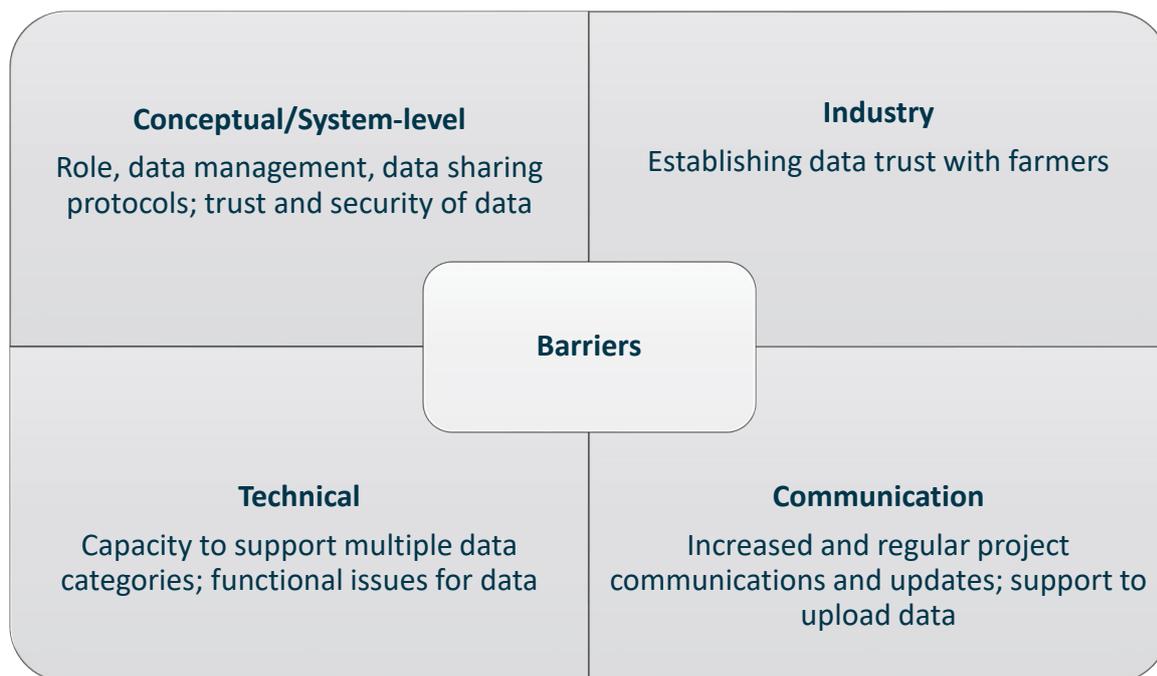


Figure 4.2: Barriers associated with VAS identified through analysis of interviews.

#### 4.1.3.1 Conceptual/system-level barriers

Conceptual/system-level barriers refer to discrepancies in understanding about the overall project, its operations, objectives, and expectations. During the first round of interviews, half the respondents were unclear about their role – and that of their organisation – in managing soils data, and establishing agreed soil data sharing protocols at their organisation and in their role as data custodians (for their members, including farmer-producers). Some participants also raised concerns about trust and security of farmer data entered in VAS. Overall, there was consensus that the data sharing and data trust challenges would be addressed in this next stage of the project, and that some of these barriers could be addressed through the provision of information and advice about how partner organisations establish their data governance arrangements.

Participants across both interview rounds were unclear about their organisations' capacity to share data in VAS, given existing agreements and expectations about data use with funding bodies, members, and farmers. In the recent interviews, these uncertainties were predominantly associated with soils data that had been collected in collaboration with other organisations, funding bodies, and with farmers.

Another barrier identified by a small subset of informants during round one interviews was a lack of understanding about data sharing arrangements. This was widely discussed by informants in the second round of interviews where participants reported uncertainty around assigning data sharing permissions. These included complexities associated with assigning data sharing permissions (detailed in Appendix B of the Milestone Report by Ollerenshaw et al. (2024)). However, there was an exception to this, with a recent interviewee commenting that the data sharing agreement in VAS was a useful resource and contained useful information that they had applied to other projects and activities:

*Well, I am hoping that I have actually asked to see some of the data sharing agreements that VAS has put together, because data sharing is, as you know, something that we're being encouraged to do more and more for good reason. But privacy and legalities around that are a complex area, and so I'm dealing with this in a number of different projects at the moment and so I am hoping that VAS has gone further down the track than I am able to do, and that I might be able to actually utilise some of their some of the intel out of their data sharing agreements,*

*that they've developed, or that you've developed, that would be very beneficial to us because it's an issue that has come that comes up a lot (Interview participant, Time 2).*

Another issue raised during the first round of interviews was about the project's longevity. A small subset raised concerns about the practical implications of accessing their organisation's data, should VAS cease to be funded long term. Another respondent commented that longevity for VAS was not only reliant on ongoing funding, but also on continued stakeholder 'buy-in', whereby VAS becomes the industry 'go to' place for data. Other concerns included the ongoing management and maintenance of their data within VAS, and security and potential technical issues that may occur with VAS in the future. Furthermore, the provision of information about the role of VAS within the broader national soils context, and the federation of data and its maintenance over time, was raised by some project partners across both interview rounds. There remains some ambiguity by a small subset about the role of VAS within the broader national soil context.

#### **4.1.3.2 Technical barriers**

During the first round of interviews, a range of technical barriers, and functional issues with VAS were identified. These included uncertainty about using the online system (and associated functionality) and uploading data into VAS, which for some, was not 'user-friendly'. Another significant challenge noted by informants in the first round of interviews was the scope and categorisation of data for uploading into VAS. Interviewees remarked that their organisations' data was wide-ranging, complex, and detailed. As such, some project partner data was unable to be fully supported by the current, limited parameters and data categories in the data import template for VAS. This was a substantial barrier that was limiting the breadth of data that partner organisations could include on VAS. This issue also had research implications: the lack of complex research data was viewed by some as limiting the scope and breadth of data extraction and discovery via the VAS data federation. Most informants commented that they had conveyed these issues to the VAS project team, who were currently working with some of the partners to find ways to accommodate and broaden the scope for reporting detailed, specific soils data.

Fewer technical challenges were reported by participants in the second round of interviews. However, participants reported challenges associated with entering complex soils data into VAS and uncertainty with shared data permissions in VAS while retaining anonymity and value in the data.

#### **4.1.3.3 Industry barriers**

Industry barriers refer to the challenges associated with issues that have links to the agriculture industry, specifically the sensitivities identified by partner organisations associated with uploading their farmer-members' data in VAS.

Recognising and addressing issues of trust associated with data management and sharing of farmer data was a persistent issue identified by participants across both rounds of interviews. The security of soils data shared in VAS and maintaining the trust that exists between partner organisations and their member-farmers were dominant concerns raised across interview rounds. This is both a key concern and priority for partner organisations. Consistently informants commented about the lack of farmer trust in sharing their soils data. Establishing the protocols within the VAS project to enable project members to seek permission and agreements to use farmer data, and thus promote trust in the data between partner organisations and their members, is important.

In the latest round of interviews, the enthusiasm for VAS is tempered by uncertainties associated with the sharing sensitive soils data from farmers/producers and their land. A lack of time to fully engage with VAS was an additional challenge identified by one informant during the recent round of interviews. Notwithstanding these concerns, the following comment demonstrates the deeper understanding that recent interview participants have of VAS. It

identifies that, as data custodians, they have the capacity to restrict public access and this, in turn, enhances their trust in the system:

*Farmers are potentially reluctant to share their data in a public forum, and you get others that are quite happy to say, well, look, if it's going to be, you know, used for research and I'm not going to be identified, more than happy. But there are different levels of access as we know, and different farmers have different ideas about what level of access they are willing to provide. And that's where data sharing agreements become really important, because if you don't have these agreements in place and there is a mistake made, you know, you need it properly agreed up front and properly documented that, you know, if they have allowed access to the researchers that they've done that on this basis and this is what they've agreed to and these are the things that they haven't (Interview participant, Time 2).*

#### **4.1.3.4 Communication barriers**

The lack of project communication was widely discussed by participants in the first round of interviews. There was widespread agreement requesting regular and increased in communication about the project. Informants commented on the lack of continuity and regularity in communication about the project. This was identified as a barrier pertaining to the sharing and conveying key developments and progress about the project, restricting their expectations about the project relating to key milestones and points. It was acknowledged by some informants that the Covid-19 pandemic may have contributed to the irregular communications about the project, and in promoting its usage.

Another barrier with communication was raised during the first round of interviews. Some informants had experienced problems with entering their data into VAS. This had prompted them to seek assistance and support from the VAS project team. Having a dedicated person to assist with these issues was considered an essential service that could facilitate and contribute to the ongoing success of the project.

Few barriers to communication were identified in the recent round of interviews. Only a small subset requested assistance with using VAS and for entering their soils data.

#### **4.1.4 VAS development considerations**

Interviewees provided suggestions for developing VAS. These include recommendations to address some of the challenges and barriers with VAS, previously identified. Two key areas where VAS could be further developed include: (a) accommodating complex soils data; and (b) increased project communications and support. These ideas are examined in this following section.

##### **4.1.4.1 Accommodating complex data**

The accommodation of complex soils data in VAS and increasing the ease of access were identified areas for VAS development and improvement. Several interviewees in the second round of interviews recommended that VAS accommodate complex soils data. Areas for expansion includes soil data categories and region-specific soil data. The following representative statements capture these ideas:

*I guess we're conscious of when we're soil sampling or collecting data we're conscious of the fact that we do, we will be wanting to move it through into the VAS portal. So, making sure that it hopefully will be in a state that is easily transferable to the portal (Interview participant, Time 1).*

*We were approached to be involved [in VAS] and we were obviously very keen to say yes because we do think that a central place to put soils data is pretty critical for Australia because there's a lot of data that exists all over the place and lots of different, you know, in lots of different formats, and lots of different organisations but it's all very varied; it's widely distributed is what I'm saying, and so it's inaccessible as a result of that (Interview participant, Time 2).*

*I think there's opportunities to use it to understand trends over time. And, you know, maybe even segregate out different types of cropping systems and that sort of thing to, you know, look at soil data as a collective against a system (Interview participant, Time 2).*

A number of interviewees called for the expansion of available soils data in VAS:

- integrating the unique characteristics and measures of existing soils data in VAS
- enabling soil data to be entered into VAS from different sources/formats
- additional data 'overlays' (i.e., topographic data; yield; trends over time; yield maps) and the inclusion of data from other organisations (government departments; statewide organisations).

These suggestions were considered essential for increasing the knowledge and insights from the soils data in VAS, and minimising data entry (where soils data can be directly extracted from other reports and formats). These ideas are captured in the following participant comments from participants in both interview rounds:

*I suppose one of our main uses that we will use it for is ... comparing things like soil pH through time. So, if you've got two points in time, like two soil pH points in time, you want to know, like, if there is a change and, yeah, why? Is it because there was ten tonnes of lime put on it, or what? Yeah, that'll be useful (Interview participant, Time 1).*

*I was thinking about some historic data we've got, which is probably only down to paddock scale. It's just understanding what, like, if we were to import that data, how do you reflect the differences between a known point and versus a known area and reflect that ambiguity geographically, that was all (Interview participant, Time 1).*

*It's probably not possible, because I imagine that everything in the background of VAS is set up based on a particular template for data upload, but it would have been good if the output that you get from say, APAL or Nutrient Advantage or CSV was a format that you could use to upload and you could just have some extra columns you've got to fill in (Interview participant, Time 2).*

*I'd still like to see perhaps some more, I guess, connection with perhaps our state agencies, and I know they have got some data in there, but things like soil moisture and that sort of thing (Interview participant, Time 2).*

#### **4.1.4.2 Communication and support**

The importance of communicating about the project was widely discussed by participants in the first round of interviews only. During which time there was widespread agreement for increasing project communication between partner. This was considered essential for the ongoing delivery of milestones associated with this next phase of the project. These ideas are reflected in the following participant quotes:

*When [the VAS project leader] has reached out and held meetings with us, it's been really informative and they've been, they've kept us in, but there has been a long time between drinks when it comes to communications and so, understanding where they're up to and what phase they're up to and what things they're implementing and when the next version is ready for us to test (Interview participant, Time 1).*

*I think all those things that we have discussed about clearer communication and milestones in regard to expectations for each of the groups perhaps, and I appreciate that there were some things that they probably had to work out along the way. So, there might have been questions that I was asking that they sort of couldn't answer because until you get into something you don't really know how it's going to go and that sort of thing (Interview participant, Time 1).*

There was also uncertainty, or at the least a need for further clarity, about how VAS would accommodate farmers and farmer data. Further information about establishing data agreements and including farmers in these discussions was recommended. Some informants

proposed that this could be achieved through the mechanism of a workshop and/or conference about data management and agreements which would be facilitated by the VAS project team:

*In terms of the, you know, data management and governance, I guess, that would be a great workshop there, but possibly not yet. In terms of how, you know, giving advice to farming groups on how to manage that. Yeah, if we can find a conference or something like that where it could be a good group discussion that would be worthwhile, giving examples showing how others are doing VAS, and the types of contracts people are putting in place (Interview participant, Time 1).*

Furthermore, the opportunity for learning about the project and communicating this through involvement and discussion with other project partners was recommended by some interviewees. The sharing of insights between project partners about their experiences, interactions and use of VAS, including how they were uploading and interrogating data, was an approach that could be informative. This is highlighted in the follow quote:

*Being able to present how VAS is presenting the data and visualising it I think would be encouraging to others to share their data hopefully so if they could see the advantages of sharing their data and how it's reported and how they can use it and those sort of things hopefully that will reduce some of the data sharing barriers (Interview participant, Time 1).*

Multiple informants requested a key contact person from the VAS project team they could contact about questions and queries linked to the self-service system in VAS:

*So I think ... if the VAS can continue to have someone to call to help in the end stages of the project would be very good (Interview participant, Time 1).*

## 4.2 VAS SYSTEM FUNCTIONALITY

Based on the feedback from the partner interviews and from requests made by the users of the self-serve system, the VAS team, and collaborating research partners, several improvements have been made to the VAS system functionality.

### 4.2.1 Revised VAS web portal

In 2024 the VAS project web portal was completely revised with a new web address – [vas.soilcra.com.au](https://vas.soilcra.com.au) – and a series of web pages (Figure 4.3). A link to the new portal has been added to the top-bar menu on the Soil CRC website.

Three 'cards' are displayed on the landing page: About VAS, Data portal, and Contact us.

The menus along the top-bar include:

- Home: navigates to the landing page.
- About: navigates to a page about the project which includes a link to the project partners, videos, and documents. Links to the project Fact Sheet and the data portal are also included.
- Data portal: navigates to the web-mapping portal to view and access the soil data.
- Resources and user guides: navigates to a page that links to the About the project page, About the data, and User guides (includes the 'how to use' instructional videos).
- Contact us: navigates to Contact information, a Feedback form, Newsletters, and Mailing list subscription.

All the content and navigation elements were reviewed by the Soil CRC before publishing. It is envisaged that this website will develop during Phase 3 to include a considerable volume of educational material to improve the project participants data literacy and be a resource for their members and stakeholders.

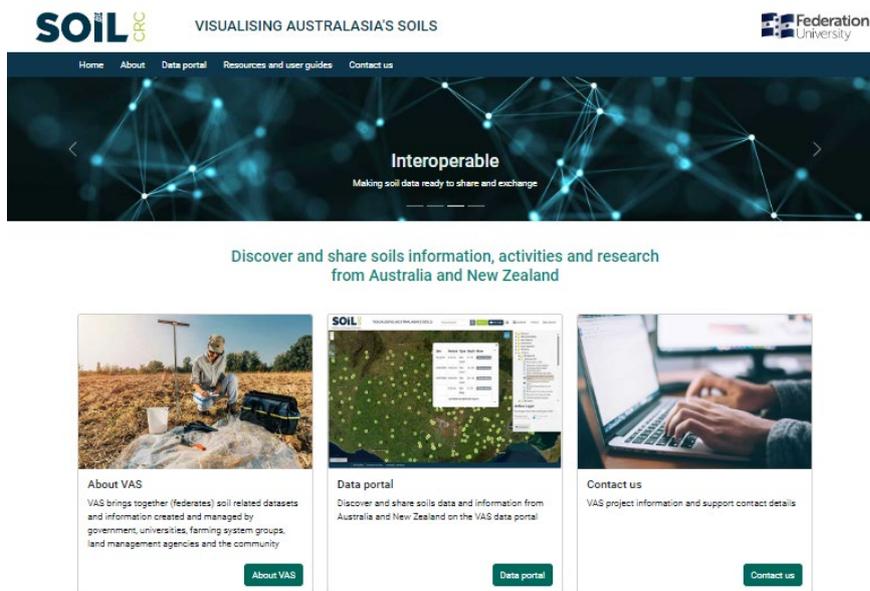


Figure 4.3: The landing page of the revised VAS website.

#### 4.2.2 Improved data reporting

Data reporting now includes more information on each soil dataset in the VAS aggregator (Figure 4.4), including data statistics, date, extent, ownership and licencing details.

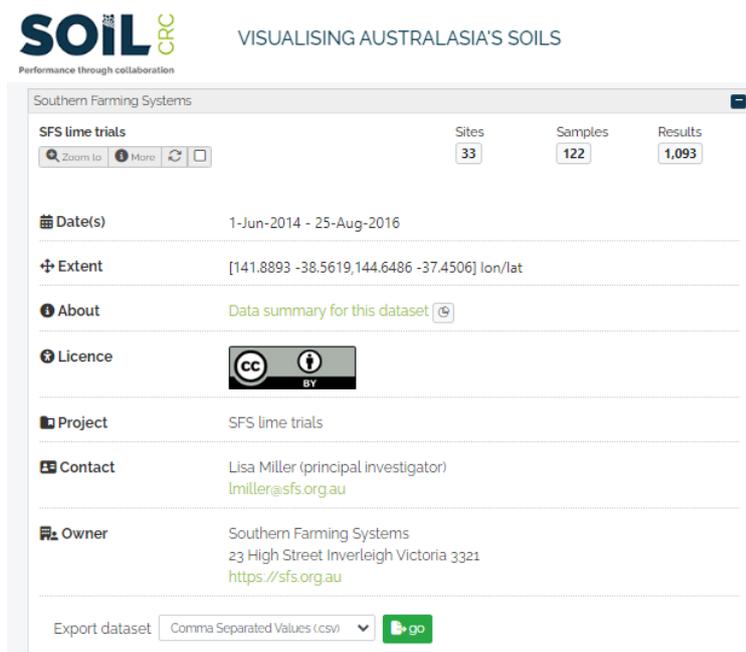


Figure 4.4: An example metadata report for a dataset.

The data summary link opens to a new page which provides reports on the number of sites, samples, observations, properties measured, analytical procedures, and sample depths (Figure 4.5a). Also included is a histogram of the number of analyses and results by sample depth (Figure 4.5b), and a summary table for each depth of the soil properties by analytical procedure and the basis statistics (Figure 4.5c). This can be displayed for each depth or for all depths using a toggle button. When the mouse cursor is placed on the procedure listed in the summary table (Figure 4.5c) it shows a hover box to name the analytical procedure used.

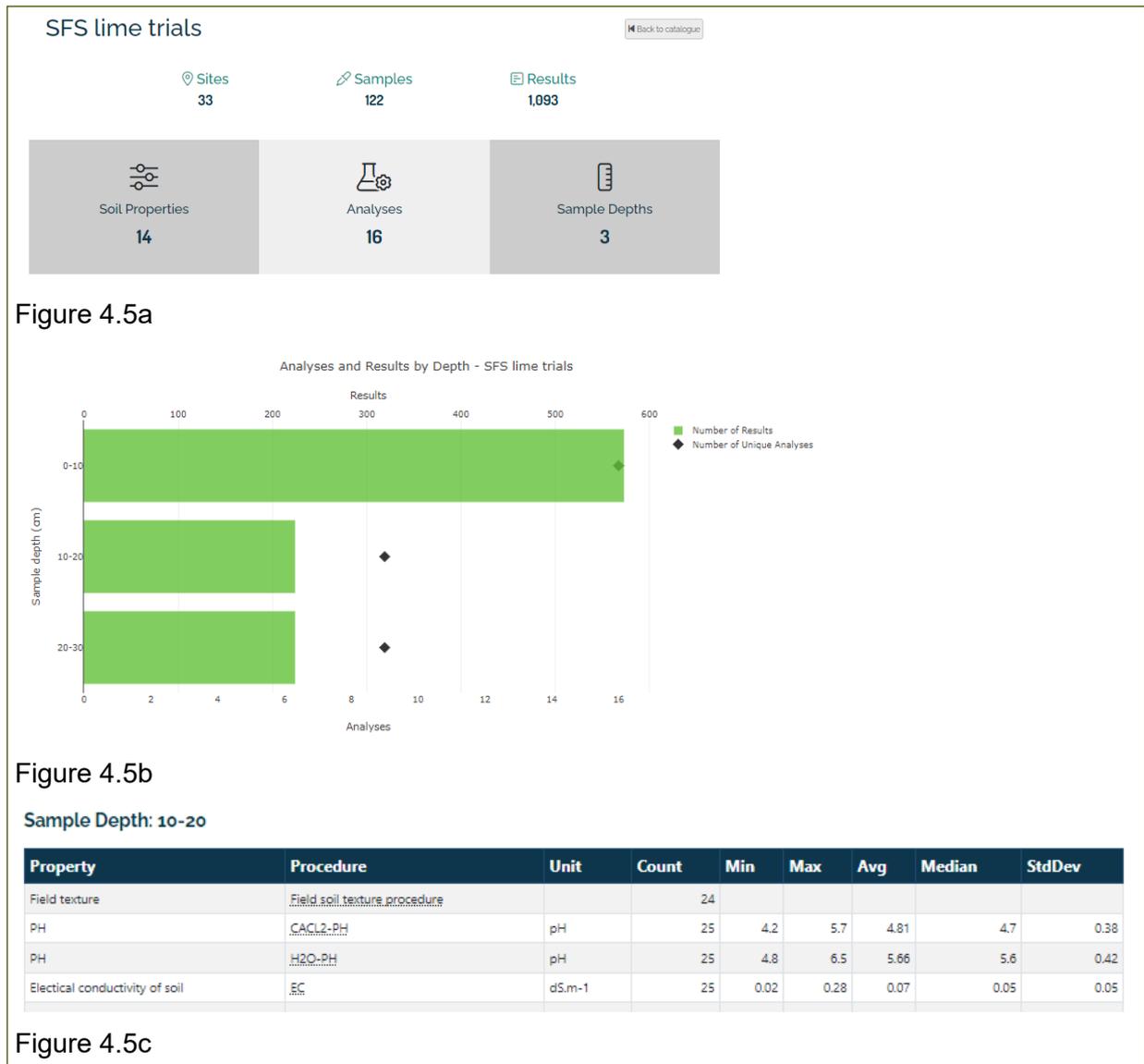


Figure 4.5: An example of some elements of the dataset summary reporting.

Where datasets have repeat samples at the same location, trends over time can be graphed for repeated soil property observations (Figure 4.6). It should be noted that this is a relatively rare situation, as most of the data are collected from one-off sampling events. Soil properties that are repeated down a soil profile for a sampling event can also be graphed (Figure 4.7). In both cases, the properties to be graphed are selected by the user.

Further enhancements to data reporting are anticipated in Phase 3.

**Site: SMP 94**

Project	Sample depth	Date(s)
Corangamite CMA	0-10 cm	Feb 2015 May 2018 May 2023
Sampled feature	Land use	Sampling method
Soil Layer	Grazing modified pastures	Core drilling

electrical conductivity of soil

Electrical conductivity (EC) of 1:5 soil/water extract

Add Clear

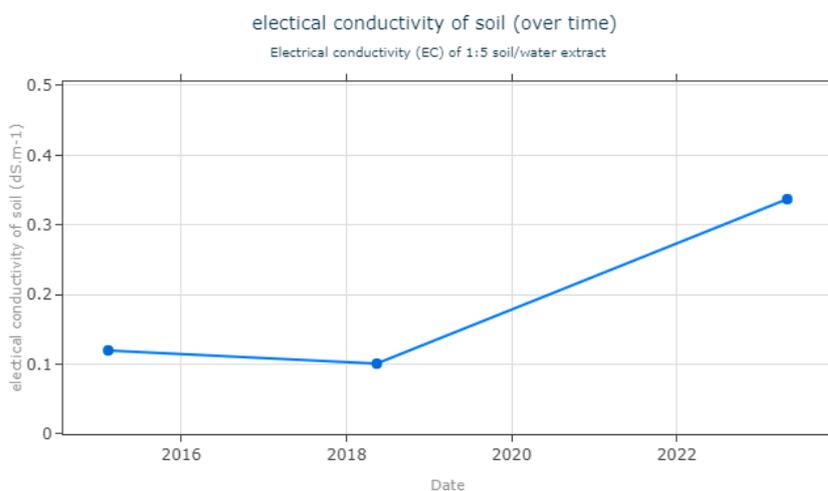


Figure 4.6: An example of charting trends over time for one parameter at one depth at a selected monitoring site.

**Site: LT\_1**

Organisation: SFS

Soil Layer(s) 0-10 cm 10-20 cm 20-30 cm

Add another property/procedure chart

Add Clear

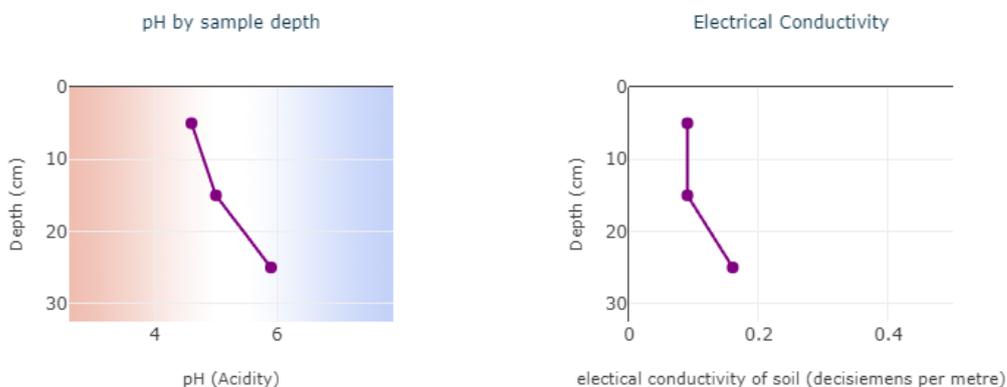


Figure 4.7: An example of charting two parameters by depth at a selected site.

### 4.2.3 Improved information model and system architecture

As illustrated in Figure 4.8, the VAS system is largely built around the CeRDI Observations System, the primary purpose of which is to collect and store observation and measurement data and publish this data for researchers, industry, government, and the public. It acts as a data aggregator that allows participants to login to a self-serve system to submit a dataset using a template that collects metadata and the soil data observations. Educational materials in the form of short 'how to' videos are provided as instructions. These data are then manually checked and mapped into the aggregator using the standard schema.

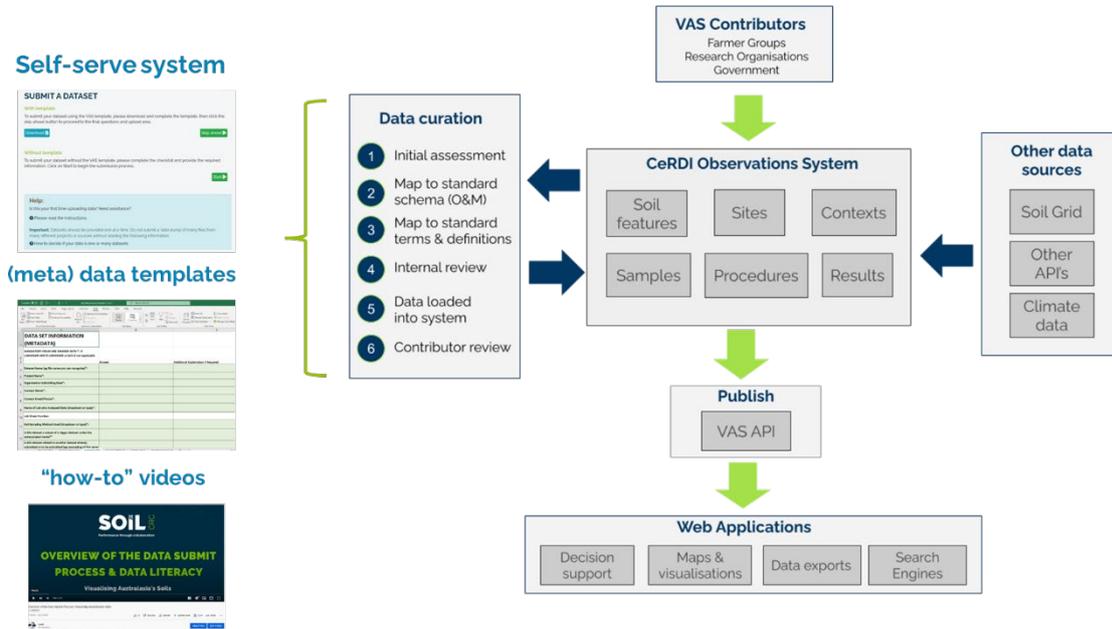


Figure 4.8: The VAS technical system overview.

As illustrated on the right-hand side of the figure, soil and ancillary data from other sources that interoperably serve standard datasets, such as ANSIS, New Zealand Soil Classification, Soil and Landscape Grid of Australia (SLGA), or the Bureau of Meteorology, can be interoperably consumed.

The data is published, subject to the data custodians' consent, via an API to the VAS portal, other systems, or users.

As described in the Phase 1 report (Dahlhaus et al. 2021), the observation system uses the ISO19156 and OGC Observations and Measurements (O&M) conceptual model to store field and laboratory environmental data in a domain independent structure. The observation model states that: 'An **Observation** is an action whose **result** is an estimate of the value of some **property** of the **feature-of-interest**, obtained using a specified **procedure**'.

In Phase 2 enhancements were made to the original observation system architecture to accommodate the complexity of the soil data received. A high-level view of the current model is illustrated in Figure 4.9 and the revised architecture is documented in Appendix D.

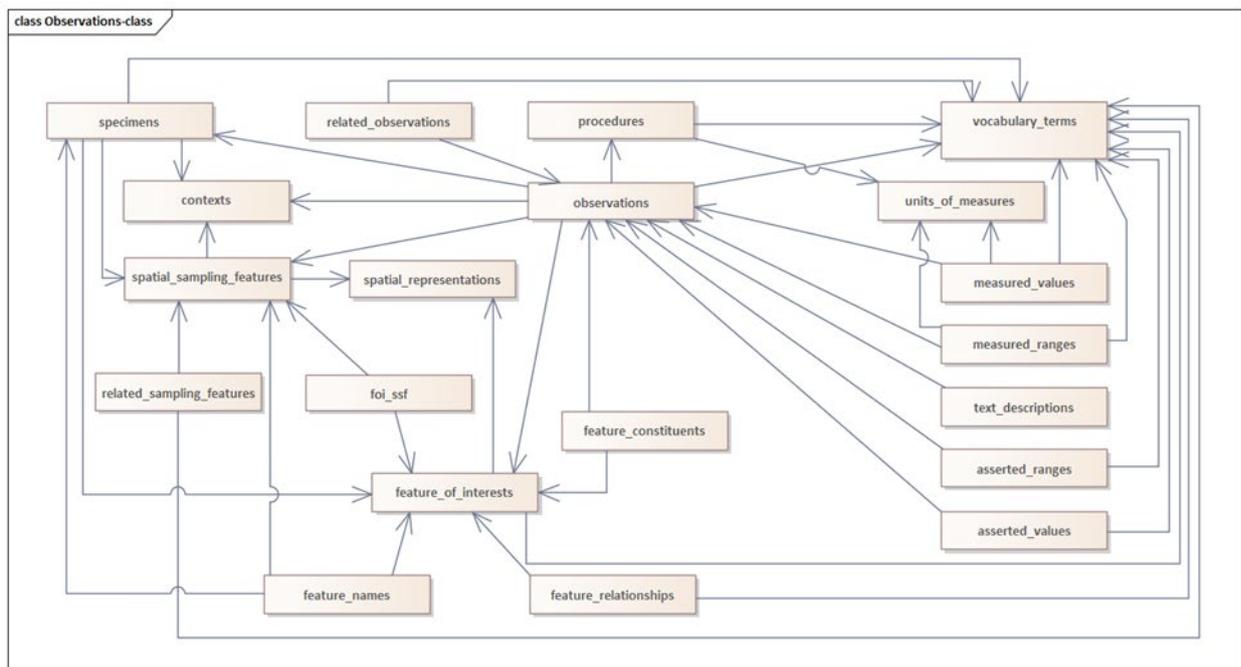


Figure 4.9: The systems architecture showing the database tables specifically relating to Observations.

#### 4.2.4 VAS application programming interface (API) development

Improved application programming interface (API) technology was developed and implemented in line with best practice set out by the World Wide Web Consortium (W3C) for publishing data on the web. This resulted in implementing an API based on Representational State Transfer (REST) architecture. The REST-full API uses persistent identifiers and data primarily delivered in JSON-LD (JavaScript Object Notation for Linked Data) format. These payloads were, in turn, based on a set of best practices for building JSON-LD APIs. To this end, the VAS API has attempted to reuse existing international standards, vocabularies, and common patterns when defining its data structures (i.e. the API Endpoints are based on the international Sensor, Observation, Sample, and Actuator [SOSA] ontology). The API and example implementations are described in Appendix E.

#### 4.2.5 Including sensor data

Since most of the participating farmer groups in the VAS project have soil moisture and temperature probes installed, methods were explored to include these data in the soil data federation. Sensor data are usually telemetered to a digital repository run by a third party (most often the probe installer), and then graphically displayed on the farmer group website via a member login.

A use-case to include SFS soil moisture and soil temperature probe data in the VAS portal was undertaken, building on their Probetrax portal<sup>5</sup> as mentioned in Section 3.3. In the Probetrax portal, SFS members hosting a probe can log into the portal and view their time-series soil moisture and temperature down profile as graphs, together with the rainfall, if a rain gauge is fitted. The sensor data are telemetered via the Adcon<sup>6</sup> network from which CeRDI sources the data and stores them in the Probetrax database.

<sup>5</sup> <https://probetrax.sfs.org.au/>

<sup>6</sup> <https://www.adcon.com/>

Displaying the SFS soil moisture probes in the VAS portal was achieved (Figure 4.10). Additional methods implemented include data from Agriculture Victoria’s soil moisture probe network in the VAS portal (Figure 4.11).

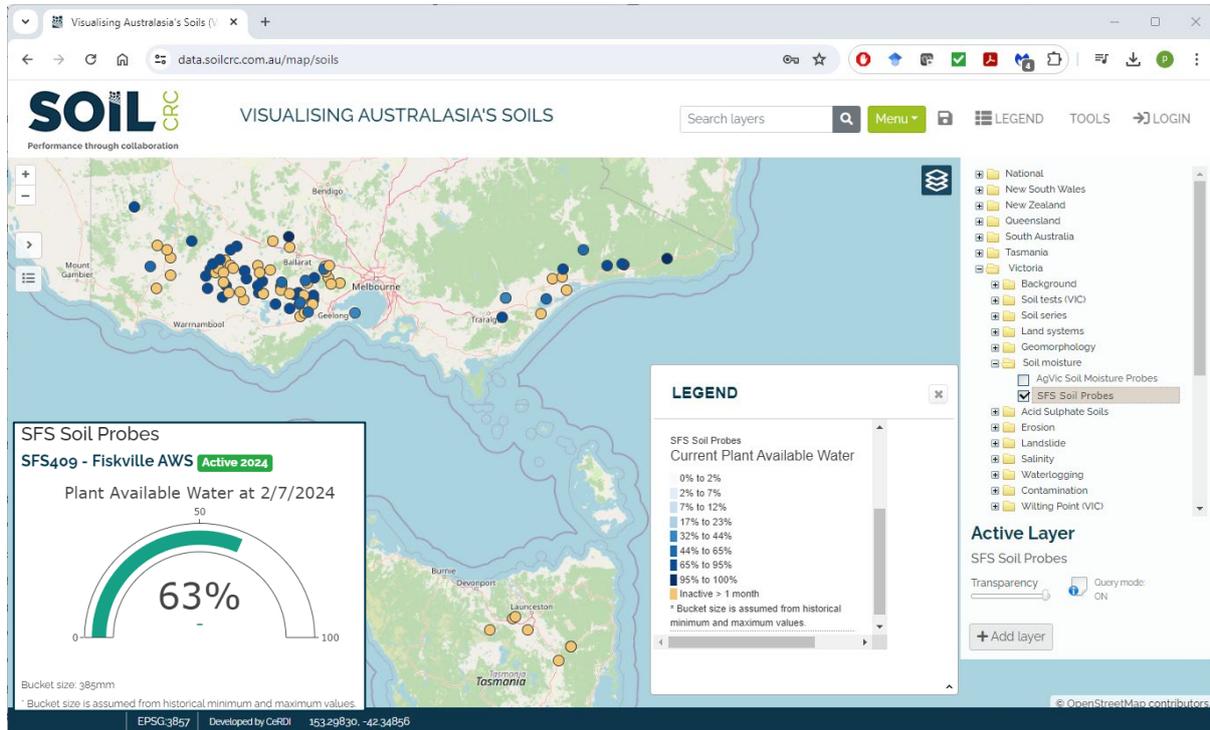


Figure 4.10: SFS soil moisture probes (public view: 2 July 2024). Insert shows the ‘speedo’ meter that is seen when clicking on a probe location.

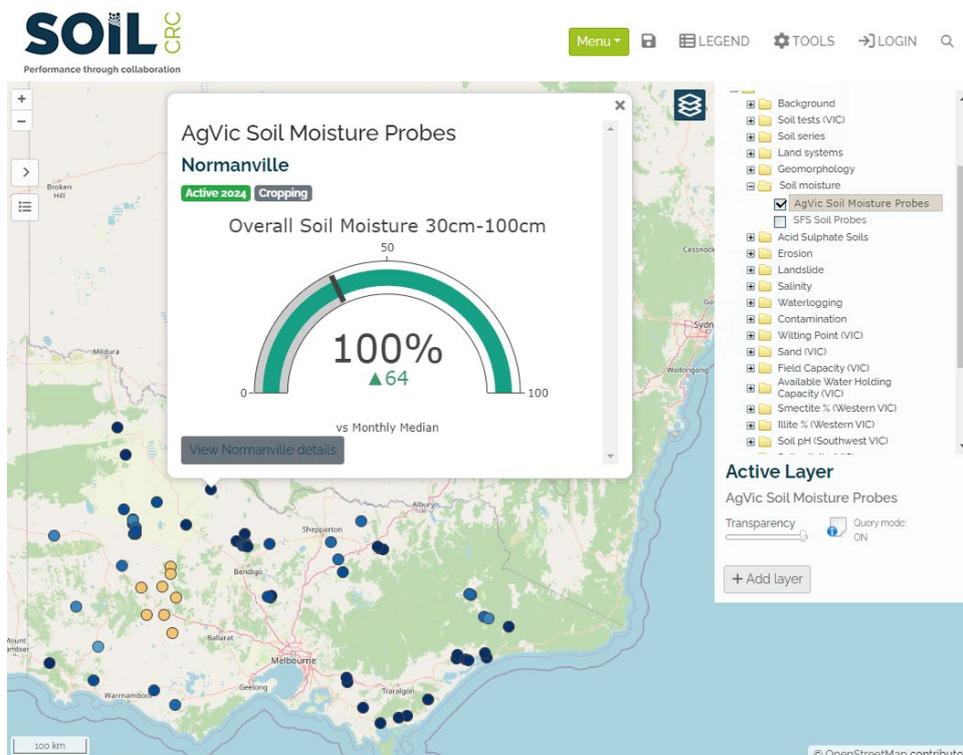


Figure 4.11: Agriculture Victoria’s soil moisture probes in the VAS portal (8 August 2024).

Research effort has continued throughout 2024 on integrating soil probes from more providers across Australia within VAS. The task has proven to be much more complex than envisaged, since the variety of data from disparate sensor types, makes, models, and third-party service providers has been challenging to standardise and harmonise. In addition, the conversion of the raw sensor readings to soil moisture values (i.e., the probe 'calibration') is not standardised and can be bespoke even for the same make and model installed by different service providers. Adding to the complexity is the fact that sensors can be moved from one location to another without being renamed or given a new identifier. These challenges are outlined in Appendix F.

A key objective is to be able to provide 'on demand' access for a user to request 'all the data', that is, include any sensor value that is reported, not just soil moisture. This request springs from researchers, especially PhD students, who are undertaking research projects that require access to time-series soil moisture and/or temperature data throughout a soil profile across a wide area. Among the many challenges this creates is the fact that these data for one provider (SFS) amounts to well over 600 million observations.

### 4.3 VAS SYSTEM USAGE

At the commencement of Phase 2 of the VAS project there were 19 farmer groups, and two catchment management authorities engaged in the project. As time progressed, seven of the groups withdrew for various reasons, such as staff turnover, time constraints, concerns around data security and perceived risk, and scepticism that the project would endure or provide value for the return on their time investment. The groups who remain in the project and have signed sub-contracts are listed in Table 4.1.

Table 4.1: Project partners participating in VAS Phase 2

Project partners	
Birchip Cropping Group	North Central Catchment Management Group
Burdekin Productivity Services	Riverine Plains Inc.
Central West Farming Systems	Southern Farming Systems
Farmlink	WA No-Till Farming Association
The Facey Group	West Midlands Group
The Liebe Group	Wheatbelt NRM
Mackillop Farm Management Group	Wimmera Catchment Management Authority

Not all of the listed partners have used the VAS data aggregator, with some yet to load any soil data. Approximately half of the partners have utilised the site on more than a couple of occasions and about a third are relatively regular users.

#### 4.3.1 Partner provisioned data

As of 31 July 2024, the VAS soil data aggregator holds contributions from 11 farmer groups, four catchment managers and two universities. This amounts to 55 datasets, representing 3,043 sites with 10,919 samples on which there are 218,375 observations (Figure 4.12 and Table 4.2). These data span the period from 1988 to 2024.

Two of the catchment managers and one university are not Soil CRC members but closely collaborate with project partners and have consented to contributing data for that reason (refer to the following section for more detail).

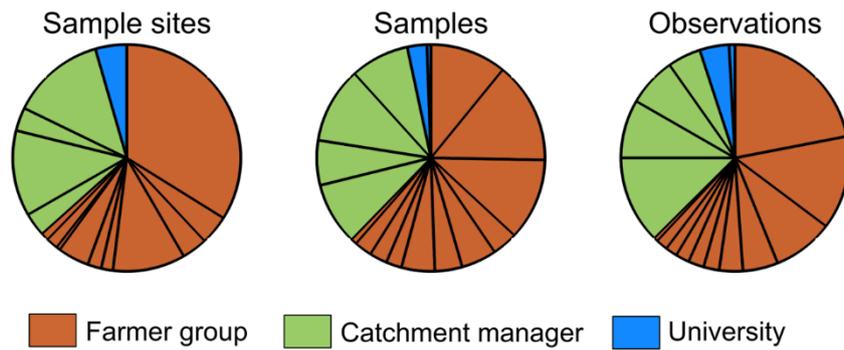


Figure 4.12: Sites, samples and observations by contributing group in the VAS soil data aggregator.

Table 4.2: Data counts in the VAS soil data aggregator.

Contributor	Datasets	Sites	Samples	Observations	Date range
Farmer group	2	1033	1188	47633	2014 - 2019
Farmer group	1	128	1592	28827	2017 - 2019
Farmer group	6	111	1273	18586	1988 - 2020
Farmer group	3	312	389	10866	2017 - 2021
Farmer group	2	53	558	7349	2011 - 2023
Farmer group	9	56	433	5067	2021 - 2024
Farmer group	1	134	522	4912	2019 - 2019
Farmer group	2	15	240	4035	2015 - 2018
Farmer group	2	48	264	3626	2017 - 2018
Farmer group	1	31	277	3574	2017 - 2020
Farmer group	1	1	90	1260	2022
Catchment manager	2	100	985	27276	2019 - 2024
Catchment manager	17	372	695	18370	2010 - 2021
Catchment manager	1	101	1173	15368	2014 - 2024
Catchment manager	1	405	915	10565	2018 - 2021
University	3	139	291	8845	2004 - 2008
University	1	4	34	2216	2013
<b>Total</b>	<b>55</b>	<b>3043</b>	<b>10919</b>	<b>218375</b>	<b>1988 - 2024</b>

The top ten properties measured are illustrated in Figure 4.13.

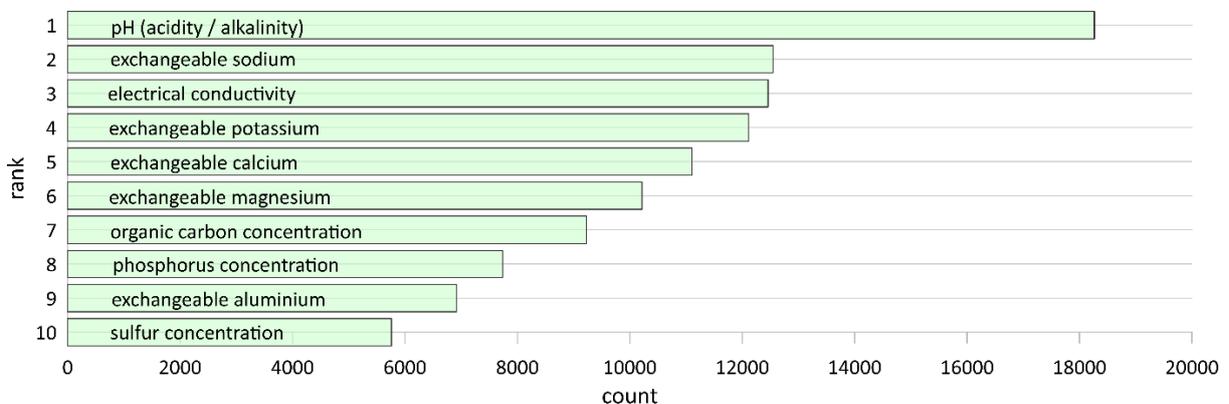


Figure 4.13: Top ten properties measured from the soil data in the VAS aggregator.

The geographic spread of these soil data is skewed to southeast Australian mainland (Figure 4.14). However, the largest contribution of samples and observations are from one farmer group in Queensland, and the second largest from a few trial sites in Western Australia.

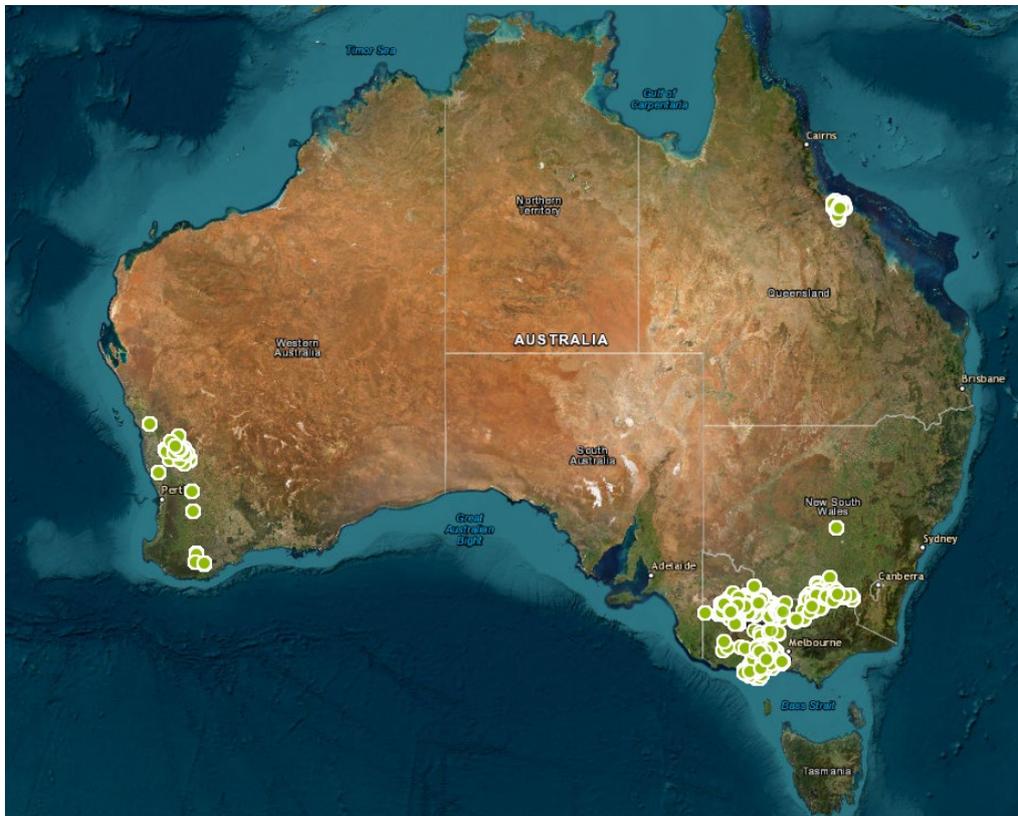


Figure 4.14 Geographic spread of the soil data stored in the VAS data aggregator.

### 4.3.2 Data sharing

The sharing of soil data is implied in both the project aims (i.e. maximise the reuse potential of the federated soil data) and the long-term objectives (i.e. motivate Australian farmers to make their data FAIR). However, the reality is that most data custodians still are reluctant to openly share data, as discussed in Section 4.1.3. While their concerns can be allayed to some extent by implementing appropriate social architecture tools, such as access controls and licensing arrangements, the decision was to proceed with four use-cases (or case studies) to explore the barriers and to demonstrate the soil data sharing options.

#### 4.3.2.1 Catchment Management Authority soil data

Although neither the Corangamite or Glenelg Hopkins CMA is a Soil CRC partner, their soil health monitoring data have been made publicly (openly) available in the VAS portal at the request of SFS (a Soil CRC member and VAS project industry partner).

The Corangamite CMA launched their soil health strategy in 2007 (Clarkson et al. 2007) which included establishing monitoring sites to check progress towards the resource condition targets in the strategy. To support the CMA's soil health initiatives, the Corangamite Soil Health Knowledge Base (SHKB) was launched as a collaborative research project between

the Corangamite CMA and CeRDI (Dahlhaus et al. 2018). The online portal, comprising an eLibrary and web-mapping interface, is still active<sup>7</sup> and provides online access to the soil data.

Soil health monitoring sites have subsequently expanded and the SHKB now displays the soil health monitoring sites for both the Corangamite CMA and the adjacent Glenelg Hopkins CMA. These comprise 101 sites in the Corangamite region which have been monitored (at set depths) on three occasions (February 2015; April 2018; May 2023) and 100 sites in the Glenelg Hopkins region that have been monitored on two occasions (February 2019; May 2023). The establishment of the sites and the collection of the data has been funded through Australian Government grants and the work was undertaken by SFS.

The access to these data is subject to a condition to not disclose the exact location of the soil monitoring sites to maintain landholder privacy (at the request of the CMAs). The method used was to randomly position the soil health monitoring site within a 2 km by 2 km polygon for the Corangamite CMA data and a 5 km x 5 km polygon for the Glenelg Hopkins CMA data (Figure 4.15). The public user has access to all the soil observations and can chart trends over time. The exact location coordinates are available to SFS via the login (i.e., hidden from public view).

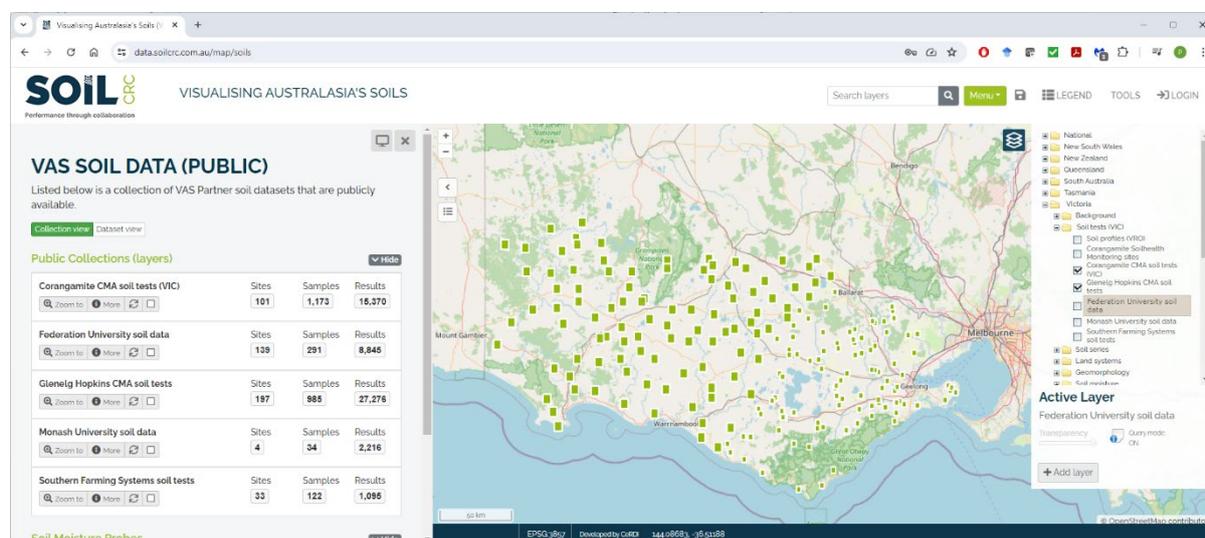


Figure 4.15: Screenshot of the CMA data in the public view of the VAS portal.

#### 4.3.2.2 University soil data

Selected soil data collected by Federation University (a Soil CRC member) and Monash University (not a Soil CRC member) has also been made publicly available in the VAS portal. The Federation University data currently comprises three soil datasets collected by Honours students for their dryland salinity research projects in 2004 and 2007, and a soil series mapping project in 2008. The Monash University dataset comprises acid sulphate soil data collected from four sites at Anglesea in 2013, as part of an Honours student project.

The only restrictions put on viewing or accessing data from these sites is to hide the landholder and property address details. Even though the Honours research data were collected on private property, permission was given at the time to publish these data in the student theses and subsequent reports for the research investors. The collection of the Federation University datasets was funded by various Australian Government grants and/or Victorian Government grants disbursed by the Corangamite CMA (i.e. public-sector funds).

<sup>7</sup> [soilhealth.ccmaknowledgebase.vic.gov.au](https://soilhealth.ccmaknowledgebase.vic.gov.au)

At the current time, not all these data can be shown, as the functionality to display the soil profile data and descriptions in a visual graphic is still in progress. An additional five datasets contributed by Federation University are currently being processed.

#### **4.3.2.3 Farmer group soil data**

While the VAS aggregator holds 30 soil datasets uploaded by farmer groups, only one been made publicly available to date. This is a dataset of 33 sites, across southwestern Victoria (Hamilton to St Leonards) collected by SFS for lime trials conducted in 2014. To protect landholder privacy, the exact locations of the sites are randomly placed within a 2 km by 2 km polygon, and users can access all the observations and generate charts of properties by depth.

These soil data have been made openly available because they were collected through a public funding grant that stipulated in the head agreement that the project data should be 'made available'.

#### **4.3.2.4 Farmer group soil probe data**

The SFS soil moisture probe data described in Section 4.2.5 is now publicly visible in the VAS portal (Figure 4.10). The public view only shows the current state of the summed moisture in the profile (i.e., the soil profile 'bucket') whereas the moisture and temperature content of the individual probes at different depths in the profile are only visible to SFS via login.

For data sharing, SFS have made the legacy probe data available for research use on request. This amounts to over 600 million observations, spanning the period from 2012 to 2024. The current database for these data is approximately 160Gb in size. The data request service is currently in the testing stage (31 July 2024). The service requires the requestor to complete a standard data sharing agreement form (Appendix G) which, once executed, supplies a download link to the requestor. If the requestor wishes to change the terms of the standard data sharing agreement, they must negotiate with SFS directly.

### **4.1 VAS SYSTEM DATA RESEARCH**

At the outset of the VAS Phase 2 project, one of the stated intentions was *'to demonstrate the value of harmonising digital soil data to create aggregated deidentified views, in a novel and pragmatic way, for example to benchmark areas, show trends over time, propose fit-for-purpose soil performance indicators, or identify significant research gaps. The aim is to encourage collaborative new projects based on the evidence drawn from data analytics supplied by the VAS'*. This research was planned as a collaboration with the University of Newcastle (Prof In-Young Yeo and Dr Indishe Senanayake) and CeRDI (A/Prof Peter Dahlhaus and Dr Nathan Robinson).

As the research planning progressed two limitations emerged:

- there was limited data of suitable quality in the VAS data aggregator
- obtaining the data custodians' permission to use the data for the research was a barrier.

As a result, a decision was made to demonstrate the potential using data sourced from the SLGA, NSW Landscape Maps (SEED database), Landsat 8 imagery, meteorological data (BoM), and a University of Newcastle dataset from two catchments in the Upper Hunter Valley. The resulting study demonstrated an effective framework for understanding and quantifying the complex relationships between soil properties and vegetation health. The research has been published in the journal *Soil Security* (Senanayake et al. 2024). A second manuscript on the research is currently under review with the journal *Science of the Total Environment* (Table 4.3). The intention is to explore this collaborative research further in VAS Phase 3 using some of the shared data (refer to Section 4.3.2) in the VAS data aggregator.

In addition to this collaborative research, there are other research outcomes that have been generated from the VAS project and from Soil CRC PhD projects in parallel to the VAS project (Table 4.3). The PhD project undertaken by Peter Weir has resulted in novel methods for paddock-scale delineation of soil moisture which are currently being programmed into applications that will be included in the VAS portal as part of the Phase 3 research.

It should be noted that there are several journal papers in preparation based on the research undertaken during VAS Phase 2. It is anticipated that at least three will be submitted by the end of 2024.

Table 4.3: Research outputs from VAS Phase 2 research<sup>8</sup>

Research outputs aligned to the VAS project
Senanayake, I.P., Yeo, I.Y., Robinson, N.J., Dahlhaus, P.G. & Hancock, G.R. (2024). Identification of high-performing soil groups in grazing lands using a multivariate analysis method. <i>Soil Security</i> , 16, 100163. <a href="https://doi.org/10.1016/j.soisec.2024.100163">https://doi.org/10.1016/j.soisec.2024.100163</a>
Senanayake, I.P., Yeo, I.Y., Hancock, G.R., Robinson, N.J., & Dahlhaus, P.G. (in review). Patterns in Greenery: Understanding Spatiotemporal Dynamics of Pasture Condition and Soil Influences. <i>Science of the Total Environment</i> (in review).
Dahlhaus P.G., Ollerenshaw A., Staines C., Corbett J. & Thompson H. (2023). Addressing the impacts of creating a soil data and dissemination system. <i>Our Soil: Spirit of a Nation. 2023 Soil Science Australia Conference, Darwin. Oral Abstracts, pp.39.</i>
Dahlhaus P., MacLeod A., Simons B., Neyland A., Wong M., Milne R., Gillett H. & Robinson N. (2023). The Visualising Australasia's Soils Project: towards a federation of public and private soil data. <i>Our Soil: Spirit of a Nation. 2023 Soil Science Australia Conference, Darwin. Oral Abstracts, pp.72.</i>
Cox S., Wong M., Gregory L. & Wilson P. (2023). Controlled vocabularies for improving soil data utility in a web-enabled world. <i>Our Soil: Spirit of a Nation. 2023 Soil Science Australia Conference, Darwin. Oral Abstracts.</i>
Dahlhaus P., MacLeod A., Ollerenshaw A., Robinson N., Wong M., Thompson H. & Doyle R. (2024). Creating a public-private soil data federation: the Visualising Australasia's Soils project. <i>Global Symposium on Soil Information and Data, September 25-28, 2004, Nanjing, China. (Oral presentation – online).</i> Food and Agriculture Organisation of the United Nations.
Published papers from Soil CRC PhD research loosely aligned with VAS research
Weir, P. & Dahlhaus, P. (2023). In search of pragmatic soil moisture mapping at the field scale: A review. <i>Smart Agricultural Technology</i> 6: 100330. <a href="https://doi.org/10.1016/j.atech.2023.100330">https://doi.org/10.1016/j.atech.2023.100330</a>
Weir, P. & Dahlhaus, P. (2024). Merging weather radar and rain gauges for dryland agriculture. <i>Journal of Southern Hemisphere Earth Systems Science</i> 74/2. <a href="https://doi.org/10.1071/ES23023">https://doi.org/10.1071/ES23023</a>
Senanayake, I.P., Pathira Arachchilage, K.R.L., Yeo, I.-Y., Khaki, M., Han, S.-C. & Dahlhaus, P.G. (2024). Spatial Downscaling of Satellite-Based Soil Moisture Products Using Machine Learning Techniques: A Review. <i>Remote Sensing</i> , 16, 2067. <a href="https://doi.org/10.3390/rs16122067">https://doi.org/10.3390/rs16122067</a>

<sup>8</sup> Published outputs only. Does not include the project Milestone Reports submitted to the Soil CRC.

## 5. DISCUSSION

The most obvious lesson from the research undertaken in VAS Phase 2 is that the challenges in creating and implementing a successful and enduring public-private soil data system are considerable. They include the social challenges of finding the value proposition for the project participants and end-users (i.e. the providers, consumers and prosumers of the soils data), the technical challenges of making it all work in an intuitive-to-use, seamless, and effortless manner, and creating a governance and business model that provides the project stakeholders with confidence and a vision that VAS will prosper beyond the lifetime of the Soil CRC.

Phase 2 of the VAS project had three stated aims:

Aim 1. Co-develop the functionality of the VAS online portal to demonstrate value for the participants (both farming groups and researchers).

Aim 2. Maximise the reuse potential of the federated soil data for research innovation and project participants' purposes through enabling data that is discoverable, accessible, and well described (semantically interoperable) to both humans and machines.

Aim 3. By 2024, transition to an agreed governance model that ensures that the future of the VAS data federation is supported by the members, and enduring.

The results of the research clarify that progress has been made towards meeting the first two aims, but not the third.

### 5.1 GOVERNANCE

Clearly, the governance of the VAS soil data federation remains the highest priority research yet to be resolved to ensure that the massive research effort to date is not wasted. While there is no doubt that the interruptions to VAS Phase 2 research due to the Covid-19 pandemic restrictions and the resultant impact on the project partners, VAS implementation team, and the collaborating research partners played a role, the governance aim is unlikely to have been met in this Phase. Many of the reasons are provided by the social science research which also shows the changing perspective of the project participants due to the evolution of the Australian soil data management ecosystem, and data management systems in general. The implementation of the National Soil Strategy and the associated initiatives announced in the 2021-22 Federal Budget, the implementation of ANSIS, the growth in use of AI tools (especially GPT), and publicity around data breaches and their consequences have all increased awareness of the value of data both as a commodity and as a security risk.

The governance of the VAS soil data federation is still being developed by weighing up the benefits and limitations of governance and business models such as data commons, collectives, co-operatives or corporations. Equally important is the governance and stewardship of the data that is shared within the system, and how the data sharing itself should be governed. The intention is to implement a governance structure that will ensure:

- trust in the VAS system (entity) and the way that it is run.
- an enduring data sharing community beyond the life of the Soil CRC.
- trust in the soil data and the community sharing these data.
- that data custodians remain in full control of sharing their data (i.e. they set the rules).
- that shared data is not used in a way that disadvantages or penalises the data custodians.
- that there are incentives to share and clear rewards for the data providers.

These challenges are common to global data sharing systems and often seen as barriers to their implementation. A significant component of the Phase 3 research will address the governance issue.

## 5.2 SOCIAL SCIENCE RESEARCH

The social science research identified similarities between the themes and ideas generated between the two interview rounds. Although there were mixed reports of VAS usage during the first round of interviews, the project continues to meet the broad and continued expectations of project partners. There remains widespread anticipation for VAS, particularly around soil data management. The practical benefits of VAS, in relation to soils data storage, access and sharing, was widely reported across all interviews. There was, however, a deeper understanding of the contributions of VAS outlined in the second round of interviews. This is evidenced by the discussions about the contribution for sharing soils data between organisations, and the potential for generating new insights and industry knowledge in the soil data available through VAS.

There remain some persistent challenges in fully realising VAS's potential as a soil data federation. Specifically, accommodating complex soil data in VAS remains a challenge. The soil data owned by project partners is intricate and these detailed data are currently unable to be fully integrated into VAS. Examples of this include complex research trials designs, additional contextual information, and additional category data from research such as treatment effects. Broader conceptual and industry-related challenges with VAS were also identified. Principally, uncertainties around data sharing and ambiguities associated with assigning data sharing permissions. These concerns are evident over time and were reported in both interview rounds (Ollerenshaw et al. 2024). Although some of these challenges are industry-inherent, insights aligned with sharing data and associated permissions, will be the focus of the third phase of VAS. Prioritising and accommodating complex and varied soils data within the VAS system remains a priority and an area for continued improvement. Streamlining the user interface to facilitate data uploads is another recommendation from users, for further developing and enhancing VAS.

Some noticeable differences in the experiences of VAS were evident in the interviews. During the second round of interviews, there were increased reports of participants entering soils data into VAS. They also articulated a comprehensive understanding of VAS for soils data management. In addition, fewer reported challenges with VAS were raised. One explanation for this is the 15-month time interval between the first and second round of interviews during VAS Phase 2. In the interim, VAS underwent substantial changes implemented in accordance with second stage developments to extend and enhance the capabilities of VAS. This included enhancing the system's technical capabilities, the application of social architecture tools, and researching use-cases to develop the data literacy of project partners and improve the soil data sharing experience for the private sector (Dahlhaus et al. 2024).

### 5.2.1 Considerations for the VAS project

The social science research clearly demonstrates that the project participants' expectations for future data sharing using VAS remains high, despite persistent challenges. Uncertainties associated with data sharing and addressing concerns around entering sensitive data, identified in the second round of interviews, must be prioritised. In particular, the continued promotion of VAS to project partners and facilitating soils data sharing is essential. To this end, the following areas must be considered in the next phase of VAS:

1. Uncertainty about sharing soil data and data sharing permissions: Actions to address this include developing a selection of easy to use and understandable resources (e.g. newsletter items, podcasts, YouTube videos, FAQ sheets) about soils data sharing in VAS. Developing flowcharts and clear and easy-to-understand guidelines about soil

data sharing and soil data permissions for project partners may reduce the current ambiguities reported by project partners.

2. Managing uncertainty about sharing farmer data in VAS: There is a need to develop resources and information about VAS for farmers. Fact sheets about VAS and explanation of the ways in which it safeguards the data and allows data custodians to restrict access to private data can outline and reinforce. A workshop for project partners in which examples and case studies are discussed to address and overcome data sharing barriers with their farmer-stakeholders may address current uncertainties. Project partners that have successfully worked with farmers to share their soil data in VAS could promote and showcasing their approaches to other project partners.
3. Ongoing and regular communication about the provision of data security in VAS and defining access to soil data within the system: This is recommended to ensure and reinforce to project partners and their stakeholders that VAS is a trusted and secure system for managing and sharing soil data. Continuous and consistent messaging about this, to project partners, is recommended. Such information could be included in fact sheets, podcasts, news items and FAQs, via the VAS newsletter.
4. Promote VAS and showcase the benefits of the system, providing content through the similarities and differences with the national soil data systems: Promoting the unique features of VAS to project partners and industry stakeholders is essential. Further outlining the contribution of VAS in comparison to other national soil data systems is essential. A news item in the VAS newsletter in May 2024 outlined the differences between VAS and other national systems. Continued discussion and promotion of VAS within the wider industry and soils data content will assist in establishing the value proposition for VAS.

These activities closely align with the proposed Phase 3 activities for VAS towards an enduring soil data federation. The value proposition of VAS is integral to its longevity, necessitating ongoing sharing of soil data by custodians. This will lead to broader industry benefits over time. Incentivising data sharing in VAS through an effortless and seamless approach is paramount to overcoming existing soil data barriers and facilitating uptake of the system (Dahlhaus et al. 2024). Defining the value proposition for the system relies on the development of use-cases that are informed by VAS users and project champions from each user/stakeholder group.

Understanding the impact of VAS in Phase 3 of the project will necessitate research that will identify and consolidate insights from all users of the portal. The perceptions and role of VAS will only be realised in this next phase of VAS, when it becomes a fully federated knowledge system of soil data delivered by soil data custodians that include farmer groups, catchment managers, university, government, and other research organisations. Research documenting the value proposition for key user groups will be integral to understanding the application of VAS during Phase 3.

### 5.3 VALUE PROPOSITIONS

Perhaps the headline takeaway from the social research is that if the VAS soil data federation can provide pragmatic value propositions through an intuitive-to-use web portal, then it will most likely continue to be supported by the data providers, consumers, and prosumers. It is apparent that access to a trusted, supported, web-based spatial soil data management system is considered very useful and highly desired by the participating farmer groups and catchment managers. This is reinforced by the common request to include the contextual data alongside the soil data (e.g., field histories, treatments, yields, climate records), so that their decision support tools can be auto populated with their latest data, and for the inclusion of an intuitive

toolkit of functions that makes their data relevant to their geographic location, and supports their stakeholder communities.

Since few of the project participants have the technical capability to interoperably serve their data according to the required standards, the soil data aggregator with the self-serve system to load their data into a secure cloud-based data infrastructure is welcomed. This is particularly because the design of the data aggregator addresses the challenge of taking varied data content and formats from a variety of data providers and makes it available to potential users in a standard format, with standard content, via a standard mechanism. That is, to make their data more FAIR, and thereby more valuable. While it is understood by the data custodians that better data lead to better decisions and can provide proof of soil stewardship for compliance with environmental standards and sustainable development goals, these potential rewards are yet to be demonstrated.

To demonstrate these benefits of FAIR soil data management, a strong focus in the next phase of the project is collaboration with participants willing to pilot use-cases and new initiatives, such as seamlessly feeding their soil data to existing tools (e.g. calculators, decision support) that add value to their data. These tools include applications already used by the participants, or 'off-the-shelf' tools developed by others and open sourced. In particular, one research priority for the next phase of the VAS project is to explore the use of AI tools that will allow farmers to converse with the federated FAIR data in a way that requires data sharing but respects data privacy. For example, using AI models based on transformer architecture that could potentially provide answers to frequently asked questions, such as: How much lime should be applied?; What is the soil carbon gap?; How much subsoil moisture will this paddock lose in the next three months?

### **5.3.1 Improving data stewardship and data literacy**

A significant limitation to implementing these potential value-adding functions is the generally poor level of data stewardship and data literacy among soil data custodians. The research to date shows that data and metadata management and stewardship practices are often immature and ad hoc, with data spread across a variety of file formats and repositories, even within a single organisation. Since metadata is key to making data FAIR, the metadata quality has been a barrier to data contribution and reuse. Almost all the participants' soil data uploaded to the soil data aggregator has required clarification of data ownership, licensing, project details, and soil test details to make it FAIR. Hence, educational tools to build data literacy within the participants and end-users has become a major priority in the project. These lessons are best developed as short videos, podcasts, and web-based tutorials.

Alongside the inadequacy of data management is the vexed issue of data variety, especially related to non-standard, poorly described, or unknown data sampling and analytical procedures. While this is a global problem being tackled by the soil data standards communities, the need for improved soil ontologies, controlled vocabularies, standardised procedures, and so on, will continue to limit the potential to maximise the value of soil data for some years to come. As more soil observation techniques evolve (e.g., soil biology and microbiology, proximal sensing, chemical determination techniques), the internationally agreed community standards for data interoperability, such as vocabularies and definitions, will take time to be developed.

## 6. CONCLUSION

The most rewarding conclusion to take from this research to date is the clear demonstration that Australasian soils data, sourced from both the public and private sectors, can be made FAIR, and shared subject to the access rules set by the data custodians. The research has developed and implemented a functional and useful soil data federation system, accessible via the soil data portal, that was co-developed with project participants to meet their needs and provision standardised data to end-users, subject to the data custodians' consent. A clear value proposition for the farmer groups and catchment managers is access to a trusted, supported, web-based spatial soil data management system that suits their purposes and is relevant to their location.

As Phase 3 of VAS commenced in August 2024, research will focus on establishing a governance framework, business case, and clear pragmatic value propositions for end-users in order to meet the goal of a creating self-sustaining and enduring Australasian soil data federation. The social science research remains a crucial component, although research examining the contribution and impact of VAS must shift from system-specific insights that have informed VAS developments to date, to a focus on the value propositions and industry-wide contributions for a federated soil data system. This will inform and document the industry-wide value associated with the development and uptake of an enduring portal for soil data federation, leading to industry-wide knowledge and insights.

The VAS project potentially provides a significant research legacy for the Soil CRC, although there remains considerable developmental work to ensure its future beyond June 2027 when the Soil CRC closes. High priorities include:

- Implementation of a governance framework for post-CRC management of the data federation
- Co-design, co-development, and implementation of use-cases that clearly demonstrate the value proposition for sharing soil data
- Inclusion of soil data collected or created in the projects funded by the Soil CRC
- Inclusion of soil data from participating research organisations, especially universities
- Interoperable connectedness with other data repositories of ancillary data, such as soil treatments, trial results, land management events, crop yields, and so on
- Co-design, co-development and implementation of seamless metadata harvesting to reduce the burden of data uploading and management via the self-serve system
- High-value outcomes for end-users, such as artificial intelligence systems to improve decision support.

The end vision is that VAS will be the legacy of an Australian soil data federation, established by the Soil CRC as an ongoing component of a national and international soil data federation. Arguably the most significant benefit from the VAS project is that it can empower Australasian farmers, agricultural industry, and researchers with the extensive data on which to make decisions, and ability to feed those data into existing and emerging tools to analyse and visualise those data. The benefits of digital agriculture have been widely lauded; for example, the economic modelling in the Precision to Decision project (Leonard et al. 2017) identified that it could lift the gross value of Australian agricultural production by \$20.3 billion. However, the VAS remains one of the very few projects that are actively trying to implement the required infrastructure to realise that benefit. Data is the fuel for the new decision support engines, such as machine learning and AI models that could benefit the Australasian agricultural industries. Good quality data that is sourced from trusted repositories and is provided in an international standard machine-readable format, with reliable metadata, will be required to model credible output.

## 7. RECOMMENDATIONS

Recommendations for the Phase 3 component are currently proposed as five work packages and the establishment of a VAS Community of Practice to guide the outcomes:

### 7.1 WORK PACKAGE 1 – THE VAS GOVERNANCE FRAMEWORK

(Estimated 20% of the project)

This research work package aims to develop and implement a governance framework for post-CRC management of the VAS system, that includes a business model.

A discussion document presented to the Steering Committee in 2020 (Dahlhaus 2019) outlined options for VAS system governance, and data governance and stewardship, with the preferred model being based on the Agricultural Research Federation (AgReFed). A decision on the VAS social architecture was deferred until there was a greater understanding of how the VAS system would work in practice. Since then, a greater understanding of international and national agricultural data sharing ecosystems has evolved. Potentially, VAS has a unique role in this evolving agricultural data sharing landscape as an enduring, self-sustaining, federation of both public and private sector soil research data. This work package will clarify how the VAS project could be governed, considering the ownership and use of the data in the system, the ownership and intellectual property of the technology, and how it connects with other data sharing initiatives. The implementation of the agreed governance and business model before the Phase 3 project ends will test the sustainability of VAS as an indispensable component of the Australian agricultural data ecosystem beyond the Soil CRC term.

### 7.2 WORK PACKAGE 2 – DATA SHARING

(Estimated 15% of the project)

A key to the enduring success of the VAS system will be including soil research data that has been collected through the Soil CRC's research projects. While interoperability with ANSIS provides links to public soil data, the VAS works with the partnering farmer groups to significantly enhance the value of their privately held data. Inclusion of sharable soil data from participating research organisations, especially universities, will also liberate quantities of high-quality data that may otherwise be hidden from future research projects, despite it being collected using public funds.

To date, sharing of soil data loaded into the VAS aggregator has been a relatively elusive research outcome. From the social science research undertaken, the reasons are complex, and include perceived risks that may be mitigated through: (a) adopting new technologies, such as encoding data licencing permissions; (b) implementing education modules to improve data literacy; and (c) use-cases that pragmatically demonstrate the value of data sharing.

### 7.3 WORK PACKAGE 3 – LINKING DATA

(Estimated 15% of the project)

A common request from VAS participants is to link their soil data to the related ancillary data, such as soil treatments, trial results, land management events, and crop yields, to improve the understanding of causation, and for input into decision support tools. The ancillary data is required to relate a project's soil research observations to the experimental actions (i.e., causation). Alternatively, their stored soil data can be linked to, and spatially and temporally matched with, ancillary data stored in other repositories, such as online farm trials or

environmental datasets. This may also extend to crowd-sourced data collected by communities (e.g., diseases, pests, environmental).

While the users' request is logical, the solution is challenging because of the limitless uniqueness of the ancillary data and the general paucity of related ontologies (e.g., information standards, controlled vocabularies). Nevertheless, as a minimum, linking to appropriate datasets through metadata catalogues is an obvious starting point.

## **7.4 WORK PACKAGE 4 – SEAMLESS METADATA HARVESTING**

(Estimated 20% of the project)

To enhance the knowledge around the various Soil CRC datasets, improved auto-data-capture capacity for soils information is required. The intention is to implement seamless metadata harvesting within the VAS self-serve system to reduce the current burden of data uploading as much as possible. These will be built using various tools such as AI systems within the current VAS soil data management system. A key requirement will be self-paced educational modules to build capacity in data literacy among some users that will explain the key components of metadata and its added value to their data.

## **7.5 WORK PACKAGE 5 – ENHANCED REPORTING**

(Estimated 30% of the project)

While 'better data leads to better decisions' remains a convincing sales pitch for the VAS system, in-built bespoke applications will be required to demonstrate value across a variety of uses. The value-proposition requires end-users to visualise their place-based data analytics using AI tools that add context and value to their soil data. Enhanced reporting of soil quality, function, and target values for a user's high-performance soils based on the analytics of their soil data, and within the context of their geography and farming system, may be the most valuable reason why the VAS portal would become an enduring, self-sustaining, soil research data federation into the future.

## **7.6 VAS COMMUNITY OF PRACTICE**

Since VAS Phase 3 focuses on pragmatic outcomes for the system end-users, a Community of Practice (VAS-CoP) will be established to guide, co-design, co-develop, and test the practicality of the applications to iterate the towards the most valuable research legacy. It is envisaged that the membership of the VAS-CoP will change throughout the project timeline, but initially would include Riverine Plains Inc., Burdekin Productivity Services, Western Australian No-Till Farmers Association, West Midlands Group, SFS, North Central CMA, and other stakeholders including CSIRO (ANSIS), Department of Agriculture Forests and Fisheries, Precision Ag, and the Regional Soil Coordinators. Discussions with these chosen organisations have indicated their strong support for co-development of applications based on their data.

## 8. ACKNOWLEDGEMENTS

This research is funded by the CRC for High Performance Soils and supported by the Cooperative Research Centres Program, an Australian Government initiative.

The authors acknowledge and thank the many individuals and organisations who have contributed significantly to the VAS project during Phase 2. In particular:

### The VAS Steering Committee Members:

- Soil CRC Program Coordinator – Richard Doyle, University of Tasmania
- Technical expert – Alistair Ritchie, Manaaki Whenua Landcare Research (New Zealand)
- Broadacre farming group – David Minkey, WA No Till Farmers Association
- Cane grower group – Rob Milla, Burdekin Productivity Services
- Mixed farming group – Jane McInnes, Riverine Plains Inc.
- Government/agency – Warwick Dougherty, NSW Department of Primary Industries (DPI), later replaced by Brian Jenkins, NSW Department of Climate Change, Energy, the Environment and Water (DCCEEW)
- Global initiatives – Peter Wilson, CSIRO Land and Water

### Soil CRC Research Support:

Michael Crawford, Jodi McLean, Lucy Weaver, Olivia Louis, Rhona Hammond, Felicity Harrop and Kathy Stokes

Past members of the CeRDI VAS team for periods during Phase 2: Bruce Simons, Rob Milne, Angela Neyland, Jennifer Corbett, Christianne Bahlo, Carolyn Staines, Dave Ryan, Cedar Martin, Derek Walters and Scott Limmer

CeRDI team members who have provided assistance: Birgita Hansen, Rick Pope, Craig Briody, Sudeera Abeywickrema, Francisco Ascui, Benedikt Fest, Gabi Ceregra, Jude Channon and Aakansha Chadha

### VAS Research Partners:

David Medyckyj-Scott, Alistair Ritchie, Brandon Whitehead, Manaaki Whenua Landcare Research (New Zealand); and Hanabeth Luke, Southern Cross University

PhD students who contributed: Basharat Ali, Rekha Attanayake, Peter Weir, Robert Clark, Kalani Pathira Arachchilage, Sepideh Jalayer

Advisors from other organisations: Troy Clarkson (DAFF), Kirsten Barlow (Precision Ag), Simon Cox and Paul Box (CSIRO), Cam Nicholson (Nicon Rural)

## 9. REFERENCES

- Ackoff, R.L. (1989). From data to wisdom. *Journal of applied systems analysis*. 16: 3-9. <http://faculty.ung.edu/kmelton/Documents/DataWisdom.pdf>
- Brown, C., Regan, Á. and Van Der Burg, S. (2023). Farming futures: Perspectives of Irish agricultural stakeholders on data sharing and data governance. *Agriculture and Human Values* 40/2: 565-580. <https://doi.org/10.1007/s10460-022-10357-8>.
- Clarkson, T., Crook, I. and Dahlhaus, P. (2007) Corangamite soil health strategy 2006–2012: a guide for investment to protect natural and built assets. Corangamite Catchment Management Authority, Colac, Australia. 144p. [http://www.ccma.vic.gov.au/soilhealth/csahs\\_report.htm](http://www.ccma.vic.gov.au/soilhealth/csahs_report.htm).
- Cox, S. and Wong, M. (2022). ANSIS Information Model (Version 1.0) <https://anzsoil.org/def/au/domain>
- Dahlhaus, P., MacLeod, A., Medyckyj-Scott, D., Simons, B., Bahlo, C., Sexton, A., Thompson, H., Wong, M., Robinson, N., Milne, R., Neyland, A., Gillett, H. and Ollerenshaw, A. (2021) Visualising Australasia's Soils: A Soil CRC Interoperable Spatial Knowledge System. Phase 1 (2019 – 2021). Project 2.3.001 Final Report. . Cooperative Research Centre for High Performance Soils (Soil CRC). Callaghan, NSW. 119p. [https://soilcra.com.au/wp-content/uploads/2022/04/FINAL\\_REPORT\\_2\\_3\\_001\\_DAHLHAUS\\_FINAL.pdf](https://soilcra.com.au/wp-content/uploads/2022/04/FINAL_REPORT_2_3_001_DAHLHAUS_FINAL.pdf)
- Dahlhaus, P., MacLeod, A., Thompson, H., Ollerenshaw, A., Robinson, N. and MacEwan, R. (2024). Implementing private sector data sharing. Visualising Australasia's Soils project. Milestone Report 13.1. Soil CRC Project 2.3.002. Callaghan, NSW, Cooperative Research Centre for High Performance Soils: 14.
- Dahlhaus, P.G. (2019) Visualising Australasia's Soils: A Soil CRC interoperable spatial knowledge system. Governance and Data Stewardship Guidelines. Discussion Paper. Soil CRC Project 2.3.001. Milestone 4 report. Cooperative Research Centre for High Performance Soils (Soil CRC), Callaghan, NSW. 22p. [https://soilcra.com.au/wp-content/uploads/private/2020/09/VAS-Governance-and-Data-Stewardship-Discussion-paper\\_27\\_02\\_20.pdf](https://soilcra.com.au/wp-content/uploads/private/2020/09/VAS-Governance-and-Data-Stewardship-Discussion-paper_27_02_20.pdf).
- Dahlhaus, P.G., Nicholson, C., Ryan, B., MacLeod, A. and Milne, R. (2018). Liberating soil data for profitable agriculture and catchment health in the Corangamite region, Australia. *New Zealand Journal of Agricultural Research*: 1-7. <https://doi.org/10.1080/00288233.2018.1432491>.
- DAWE. (2021a). Budget 2021-22. National Soil Strategy, Australian Government Department of Agriculture, Water and the Environment. . Factsheet 2. <https://www.awe.gov.au/sites/default/files/2021-05/national-soil-strategy-factsheet.pdf>
- DAWE. (2021b) National Soil Strategy. Department of Agriculture, Water and the Environment A. Government, Canberra, A.C.T. . 60p. <https://www.agriculture.gov.au/sites/default/files/documents/national-soil-strategy.pdf>.
- DAWE. (2023a). Outcomes of the Historical Soil Data Capture Payments Program Canberra, A.C.T. , Australian Government, Department of Agriculture, Water and the Environment: 2. <https://www.agriculture.gov.au/sites/default/files/documents/historic-data-factsheet.pdf>.
- DAWE. (2023b). Outcomes of the Pilot Soil Monitoring and Incentives Program. Canberra, A.C.T., Australian Government, Department of Agriculture, Water and the Environment 3. <https://www.agriculture.gov.au/sites/default/files/documents/psmip-fact-sheet.pdf>.

- FAO. (2018). 'Global Soil Partnership.' (web page) Retrieved 7 May 2018, from <http://www.fao.org/global-soil-partnership/en/>
- GODAN. (2023). 'GODAN. Global Open Data for Agriculture & Nutrition.' (web page) Retrieved 24/10/2023, from <https://www.godan.info/about-godan>
- ISRIC. (2021). 'ISRIC World Soil Information.' (web page) Retrieved 25 February 2021, from <https://www.isric.org/>.
- IUSS. (2020). 'IUSS Working Groups.' (web page) Retrieved 25 February 2021, from <https://www.iuss.org/organisation-people/organisation/working-groups/>.
- Kaur, J., Hazrati Fard, S.M., Amiri-Zarandi, M. and Dara, R. (2022). Protecting farmers' data privacy and confidentiality: Recommendations and considerations. *Frontiers in Sustainable Food Systems* 6: 903230. <https://doi.org/10.3389/fsufs.2022.903230>.
- Kitchin, R. (2021). *The Data Revolution: A critical analysis of big data, open data and data infrastructures*, SAGE Publications Ltd. <http://digital.casalini.it/9781529766776>.
- Leonard, E., (Ed), Rainbow, R., (Ed), Trindall, J., (Ed), Baker, I., Barry, S., Darragh, L., Darnell, R., George, A., Heath, R., Jakku, E., Laurie, A., Lamb, D., Llewellyn, R., Perrett, E., Sanderson, J., Skinner, A., Stollery, T., Wiseman, L., Wood, G. and Zhang, A. (2017). Accelerating Precision Agriculture to Decision Agriculture. Enabling digital agriculture in Australia. . Cotton Research and Development Corporation, Australia. 40p. <https://www.crdc.com.au/sites/default/files/CRD18001-001%20CRDC%20P2D%20Report%20low%20res.pdf>
- MacLeod, A., Gillett, H., Simons, B., Wong, M., Reynolds, A., Martin, C., Archer, R. and Dahlhaus, P. (2024). Visualising Australasia's Soils: Phase 2 System Architectures. Milestone Report 14.3. Soil CRC Project 2.3.002. Callaghan, NSW, Cooperative Research Centre for High Performance Soils: 51.
- NFF (2020). Farm Data Code. Canberra, Australia, National Farmers Federation: 8.
- NFF (2023). Farm Data Code. Edition 2. 2023. Barton, A.C.T. , National Farmers Federation. <https://nff.org.au/wp-content/uploads/2023/06/NFF-Australian-Farm-Data-Code-v2-May-2023.pdf>
- NHMRC (2018). Australian Code for Responsible Conduct of Research. Canberra, National Health and Medical Research Council, Australian Research Council and Universities Australia. Commonwealth of Australia. <http://www.nhmrc.gov.au/guidelines/publications/r41>.
- OGC. (2017). 'Agriculture DWG.' (web page). Retrieved 30 January 2017, from <http://www.opengeospatial.org/projects/groups/agriculturedwg>.
- Ollerenshaw, A., Staines, C., Thompson, H. and Dahlhaus, P. (2024). Visualising Australasia's Soils: Phase 2 (Final Report) Impact Research Report. Milestone Report 14.2. Soil CRC Project 2.3.002. Callaghan, NSW, Cooperative Research Centre for High Performance Soils: 35.
- Panagos, P., Van Liedekerke, M., Borrelli, P., Köninger, J., Ballabio, C., Orgiazzi, A., Lugato, E., Liakos, L., Hervas, J., Jones, A. and Montanarella, L. (2022). European Soil Data Centre 2.0: Soil data and knowledge in support of the EU policies. *European Journal of Soil Science* 73/6: e13315. <https://doi.org/10.1111/ejss.13315>.
- RDA. (2018). 'Agricultural Data Interest Group (IGAD).' (web page) Retrieved 7 May 2018, from <https://www.rd-alliance.org/groups/agriculture-data-interest-group-igad.html>
- Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of information science* 33/2: 163-180. <https://doi.org/10.1177/0165551506070706>.

Saldana, J.M. (2016). *The Coding Manual for Qualitative Researchers*. 3rd ed. London, England, SAGE Publications.

Senanayake, I.P., Yeo, I.Y., Robinson, N.J., Dahlhaus, P.G. and Hancock, G.R. (2024). Identification of high-performing soil groups in grazing lands using a multivariate analysis method. *Soil Security* 16: 100163. <https://doi.org/10.1016/j.soisec.2024.100163>.

Sexton, A. (2020) *Visualising Australasia's Soils: Social Engagement and Collaboration Learnings*. Soil CRC Research report. June 2020. Federation University Australia, Mt Helen, Vic. Cooperative Research Centre for High Performance Soils (Soil CRC), Newcastle, NSW. 24p. [https://soilcrc.com.au/wp-content/uploads/2020/12/VAS-Social-Engagement-Report-2020\\_FINAL.pdf](https://soilcrc.com.au/wp-content/uploads/2020/12/VAS-Social-Engagement-Report-2020_FINAL.pdf).

United Nations (2014). *A world that counts: mobilising the data revolution for sustainable development*, United Nations Secretary-General's Independent Expert Advisory Group on a Data Revolution for Sustainable Development 30. <https://www.undatarevolution.org/report/>

United Nations (2015). *Transforming our world: The 2030 agenda for sustainable development*. New York: United Nations, Department of Economic and Social Affairs. 1: 41. <https://sustainabledevelopment.un.org/post2015/transformingourworld/publication>

Zhang, A., Heath, R., McRobert, K., Llewellyn, R., Sanderson, J., Wiseman, L. and Rainbow, R. (2021). Who will benefit from big data? Farmers' perspective on willingness to share farm data. *Journal of Rural Studies* 88: 346-353. <https://doi.org/10.1016/j.jrurstud.2021.08.006>.

# APPENDIX A INTERVIEW QUESTIONS

## Interview questions (Existing project partners)

1. Please tell me about how you have been using VAS?
2. Have your expectations of VAS been met? If so, in what ways?  
If not: In what ways has it not met your expectations? Do you have suggestions for addressing your expectations?
3. Did meeting the Federation University team inform/expand your understanding and expectations for VAS over the coming three years? If so, in what ways?
4. In what ways do you anticipate the proposed extension to VAS will meet your organisation's needs and expectations?
5. a) Has your organisation's a) data access, b) data management, and c) data sharing practices changed since being involved in VAS? If so, in what ways?  
b) Do you anticipate additional changes in a) data access, b) data management, and c) data sharing practices during this second phase of VAS? If so, in what ways?
6. a) Has VAS encouraged or discouraged information sharing in the sector that you work? Please provide examples.  
b) In what ways do you see data sharing being facilitated/optimised during the extension of VAS?
7. a) Are there circumstances when your organisation is unable to share data? Please outline the circumstances/reasons why you/your organisation would be unable to share data.  
b) Do you anticipate this second phase of VAS will reduce current data sharing barriers?
8. Have you any recommendations for improving VAS in relation to how data is discovered, communicated, and synthesised?
9. Has VAS helped your decision making and/or changed your work practices? How and in what ways?
10. Do you have any insights about the governance associated with VAS? What has worked well? Are there any areas of governance for VAS that could be improved?
11. Do you have other comments or suggestions about VAS?
12. May we contact you in future for follow up research?

### **Interview questions (New project partners)**

1. Please tell why you/your organisation have become a partner with VAS?
2. Please explain what you know about VAS?
3. As a new partner, how and in what ways do you anticipate VAS will meet your organisation's needs and expectations? Which elements or features of VAS will be useful/beneficial to your organisation?
4. How will your organisation use VAS? Do you anticipate your current a) data access, b) data management and c) data sharing practices will change through participation in VAS? If so, in what ways?
5. a) In what ways do you see data sharing being facilitated/optimised through VAS? What opportunities do you see for sharing your own data through VAS?  
b) What are the anticipated benefits, for your organisation and the industry, in accessing data within the VAS partnership?
6. a) Are there circumstances when your organisation is unable to share data? Please outline the circumstances/reasons why you/your organisation would be unable to share data.  
b) Do you anticipate that your involvement in the VAS project could reduce these data sharing barriers?
7. Do you have other comments or suggestions about VAS?
8. May we contact you in future for follow up research?

## APPENDIX B DATA SUB-CONTRACTS

SUBCONTRACT TO PROVIDE SERVICES TO FEDERATION UNIVERSITY AUSTRALIA		
<b>University</b>	<b>FEDERATION UNIVERSITY AUSTRALIA</b> incorporated pursuant to the <i>Federation University Australia Act 2010</i> (ABN 51 818 692 256) of University Drive, Mt Helen, Victoria 3350	
	<table border="1"> <tr> <td> <b>Project Leader</b>            Assoc. Prof. Peter Dahlhaus            Principal Research Fellow,            Centre for eResearch and Digital Innovation (CeRDI)            PO Box 663, Ballarat VIC 3350  <a href="mailto:p.dahlhaus@federation.edu.au">p.dahlhaus@federation.edu.au</a>            Tel: +61 3 5327 9266         </td> <td> <b>University Representative</b>            [name]            Deputy Vice-Chancellor, Research and Innovation            PO Box 663, Ballarat VIC 3350  <a href="mailto:research.funding@federation.edu.au">research.funding@federation.edu.au</a>            Tel: +61 3 5122 6872         </td> </tr> </table>	<b>Project Leader</b> Assoc. Prof. Peter Dahlhaus Principal Research Fellow, Centre for eResearch and Digital Innovation (CeRDI) PO Box 663, Ballarat VIC 3350 <a href="mailto:p.dahlhaus@federation.edu.au">p.dahlhaus@federation.edu.au</a> Tel: +61 3 5327 9266
<b>Project Leader</b> Assoc. Prof. Peter Dahlhaus Principal Research Fellow, Centre for eResearch and Digital Innovation (CeRDI) PO Box 663, Ballarat VIC 3350 <a href="mailto:p.dahlhaus@federation.edu.au">p.dahlhaus@federation.edu.au</a> Tel: +61 3 5327 9266	<b>University Representative</b> [name] Deputy Vice-Chancellor, Research and Innovation PO Box 663, Ballarat VIC 3350 <a href="mailto:research.funding@federation.edu.au">research.funding@federation.edu.au</a> Tel: +61 3 5122 6872	
<b>Subcontractor</b>	<b>[organisation name]</b> <b>ABN:</b> <b>Contact:</b> [name] <b>Address:</b> [postal address] <b>Phone:</b> <b>Email:</b>	
Details of Services		
<b>Head Agreement</b>	The Project Agreement titled PJA 2.3.002 signed by the University and the CRC for High Performance Soils Limited, dated 08/11/2021.	
<b>Client</b>	CRC for High Performance Soils Limited (ABN 63 618 897 224)	
<b>Project</b>	2.3.002 - Visualising Australasia's Soils: Extending the soil data federation (VAS Phase 2)	
<b>Services</b>	<p>Contribute to the VAS Phase 2 project by carrying out activities across the following areas as further detailed in Schedule 1:</p> <ol style="list-style-type: none"> <li>Provide access to reliable soil and other on-farm data             <ul style="list-style-type: none"> <li>Adopt agreed data stewardship and governance practices so more soils data and related information can be discoverable through the VAS portal</li> <li>Regularly contribute soil test data via the VAS self-serve system to support the goal of co-creating and sustaining a soil knowledge system that is inherently useful for research, development, extension and adoption.</li> <li>Support the linking of local weather station and soil moisture network data via the VAS portal</li> </ul> </li> <li>Deliver extension activities to support VAS adoption             <ul style="list-style-type: none"> <li>Hold workshops and/or training sessions to support VAS awareness and adoption</li> <li>Gather feedback about VAS from a range of stakeholders (staff and members)</li> </ul> </li> </ol>	

	<ul style="list-style-type: none"> <li>Submit a short report to consolidate extension activities and findings (2-5 pages). Detail participant feedback, VAS self-service adoption progress and share recommendations on what could be improved</li> </ul> <p>3. Contribute to the delivery of tangible and attractive benefit to farming groups and other VAS participants</p> <ul style="list-style-type: none"> <li>Participate in the annual VAS partner workshop (face-to-face or online), contribute to and share VAS updates via member and other communication channels</li> <li>Assist in the co-design of 'use cases' that will provide specific value to VAS partners and farmer members</li> <li>Contribute to defining specific functionality that would get Farming Systems Groups and members using the VAS system.</li> </ul>			
<b>Start Date</b>	Upon Subcontract execution	<b>Completion Date</b>	Upon final payment to the Subcontractor	
<b>Fees</b>	Subject to performance of the Subcontracted Services and Deliverables by the Subcontractor to the satisfaction of the University's Project Leader, and payment by the Client under the Head Agreement, the University will pay the following fees:			
	<b>Payment No.</b>	<b>Milestone</b>	<b>Milestone due date</b>	<b>Fees payable on completion of Milestone (Excluding GST)</b>
	1	Acceptance of the activity plan and execution of subcontract	Execution of subcontract	\$ 5,000
	2	Progress report and updated project plan	[date]	\$ 5,000
	3	Final report	[date]	\$ 5,000
<b>Total Project Fee (excluding GST)</b>			\$15,000	
<b>Payment Terms</b>	Within 30 days of the date of the invoice			
<b>Specified Personnel</b>	<b>The University's Chief Investigator</b>		<b>The University's key personnel</b>	
	Assoc Prof Peter Dahlhaus		Assoc Prof Helen Thompson [administrator name]	
	<b>The Subcontractor's Chief Investigator</b>		<b>The Subcontractor's key personnel</b>	

**EXECUTED AS AN AGREEMENT**

**SIGNED** for and on behalf of  
**FEDERATION UNIVERSITY AUSTRALIA**  
in the presence of:

)  
)  
)

.....  
Signature of authorised person

.....  
Signature of Witness

.....  
Title

.....  
Name of Witness  
(block letters)

.....  
Name of authorised person  
(block letters)

.....  
Date

**SIGNED** for and on behalf of  
**INSERT SUBCONTRACTOR NAME**  
in the presence of:

)  
)  
)

.....  
Signature of authorised person

.....  
Signature of Witness

.....  
Title

.....  
Name of Witness  
(block letters)

.....  
Name of authorised person  
(block letters)

.....  
Date

## BACKGROUND

- A The University entered into the Head Agreement with respect to the Project.
- B The University wishes to enter into this Subcontract Agreement with the Subcontractor for delivery of a portion of the Project.
- C The Subcontractor has agreed to undertake the Subcontracted Services and provide the Subcontractor's contributions in accordance with the terms and conditions of this Agreement.

## OPERATIVE TERMS

### 1. Definitions and Interpretation

---

#### 1.1. Definitions

In this Agreement

**Agreement** means this Subcontract Agreement, including any schedules or other attachments.

**Background IP** means material, and the Intellectual Property rights in such material created independently of the Services or in existence at the Commencement Date or in the case of the University, consists of the University's generated Intellectual Property under the Head Agreement (whether before or after the date of this Agreement) and made available by that party for the purpose of performing the Services.

**Client** means the Client specified on the cover page on whose behalf the University requires the Subcontractor to perform the Services further to the Head Agreement.

**Confidential Information** means information of either party which is identified as being confidential at the time of disclosure to the other party, or which a party knows or ought to know is confidential to the other party, and in the case of the University includes the Project Intellectual Property which is regarded as confidential to it.

**Contract Material** means all material forming part of the Deliverables or brought or required to be brought into existence as part of or for the purposes of performing the Services.

**Force Majeure Event** affecting a person means anything outside that Party's reasonable control including fire, storm, flood, earthquake, explosion, war, invasion, rebellion, sabotage, pandemic, epidemic and act or omission (including laws, regulations, policy changes, disapprovals or failures to approve) of any third person (including, but not limited to, customers, governments or government agencies).

**Head Agreement** means the agreement between the University and the Client, specified on the cover page in relation to the Project, as outlined in Schedule 1 to this Agreement.

**Intellectual Property** means without limitation all intellectual property protectable by statute and includes all patentable inventions, copyright, trademarks, registrable designs, registrable plant breeder's rights and circuit layouts.

**Project IP** means IP arising from the Project or the Services, including such IP in the Deliverables and Contract Material.

**Services** means those activities to be performed by the Subcontractor outlined on the cover page.

**Subcontractor** means the subcontractor specified on the cover page.

**University** means Federation University Australia.

#### 1.2. Interpretation

In this Agreement

- a) references to a party means the University or the Subcontractor and references to the parties mean both the University and the Subcontractor
- b) Single words include plural words, and vice versa.
- c) Words of inclusion are not intended to be words of limitation.
- d) words defined in the Head Agreement have the same meaning if used in this Agreement; and
- e) a reference to a clause or schedule is to a clause or schedule of this Agreement, unless stated otherwise.

### 2. Subcontractor's Obligations

---

- 2.1. The Subcontractor acknowledges that although the Subcontractor is not a party to the Head Agreement, the provisions of the Head Agreement must be read with this Agreement.
- 2.2. Notwithstanding anything in this Agreement, the Subcontractor agrees to all reasonable steps to enable the University to meet its obligations under the Head Agreement.
- 2.3. The Subcontractor must take all reasonable steps to assist the University to minimise delay in completing the Services.
- 2.4. The Subcontractor warrants that:
  - a) the provision of the Services will be carried out with all due care and skill in accordance with all applicable standards, principles and practices;
  - b) the provision of the services will not infringe any right of any third party (including any intellectual property right) or any Laws;
  - c) all representations made by the Subcontractor in or in connection with the Services were and remain accurate and the

Subcontractor has and will maintain at all relevant times the quality assurance arrangements set out in the Head Agreement or as otherwise specified by the University.

- 2.5. The Subcontractor will deliver the Deliverables to the University by the Delivery Date.
- 2.6. The Subcontractor will liaise with the University regarding progress and provide progress reports as such time and in the manner required by the University to enable the University to meet its reporting obligations under the Head Agreement.
- 2.7. The Subcontractor will not do anything that would place the University in breach of its obligations under the Head Agreement.

### 3. Fees

---

#### 3.1. Payment of Fees

Subject to performance of the Services in accordance with this Agreement and Clause 3.2(a), the University must pay the Fees to the Subcontractor in accordance with the cover page.

#### 3.2. Head Agreement

If the Client fails to pay the amounts contemplated by the Head Agreement to the University at the times contemplated by the Head Agreement then the obligation on the University to pay the Fees under clause **Error! Reference source not found.** is deferred until such time as those amounts have been paid.

For clarity, clause 0 does not apply to any failure of the Client to pay the amounts contemplated by the Head Agreement if that failure arises from a breach by the University of the Head Agreement that is not contributed to by the acts or omissions of the Subcontractor.

The Subcontractor will:

use the Fees solely for the purpose of performing the Subcontracted Services and providing the Deliverables in accordance with this Agreement, and not for any other purpose; and

return to the University, promptly upon request, any part of the Fees that remains unexpended or uncommitted for the purpose of performing the Services as at the date of expiry or termination of this Agreement.

#### 3.3. Invoicing

- (a) The Subcontractor must provide an invoice to the University for the Fees at the times specified in the Schedule.
- (b) Notwithstanding Clause 3.1 the University is not obliged to pay the Fees until 30 days after receipt of a valid invoice which contains sufficient reasonable and specific details to enable the work to which the invoice applies to be identified.

### 4. Project IP and Contract Material

---

- 4.1. All Project IP will vest in the Client as per the Head Agreement. The Subcontractor assigns all Project IP created or developed in the performance of the Services accordingly and agrees to do all things reasonably necessary to give effect to such ownership and assignment (including without limitation, ensuring its employees do the same).
- 4.2. The parties agree that ownership of Background IP will remain with the party who contributed the Background IP.
- 4.3. Where Background IP is made available by the Subcontractor for the performance of the Services, the Subcontractor does so in accordance with the Head Agreement (updating the project IP register with the Client), and further grants to the University a licence to use the Background IP for the performance of the Project.

### 5. Confidentiality and Publication

---

- 5.1. Each party agrees to keep Confidential Information disclosed by the other party secret, secure and confidential, and must not disclose, communicate, or otherwise make known to any person any part of the Confidential Information without the prior written consent of the disclosing party, which will not be unreasonably withheld.
- 5.2. A party receiving Confidential Information under this Agreement will be relieved of its obligations contained in clause 5.1 in respect to any Confidential Information which:
  - (i) is required to be disclosed by law;
  - (ii) was in the possession of the Subcontractor prior to the disclosure; or
  - (iii) becomes part of the public domain otherwise than by a breach of this Agreement, or
  - (iv) was received in good faith from a person who is not a party to this Agreement, and who did not receive the Confidential Information from the Subcontractor or any person in respect to whom the Subcontractor can trace the provision of the Confidential Information originating with it.
- 5.3. The Subcontractor must not publish or disclose information in relation to the Project or Service, including Confidential Information or the Project IP, except as permitted under the Head Agreement.
- 5.4. Each party must comply with the Privacy and Data Protection Act 2014 (Vic) and the reasonable directions of the other party in relation to the handling of any personal information which that party holds or has held and used in connection with this Agreement.
- 5.5. Clauses 5.1 to 5.4 will survive the termination or expiration of this Agreement.

### 6. Indemnity

---

6.1. Each Party (“Indemnitor”) irrevocably and unconditionally indemnifies and agrees to keep indemnified each other Party and their respective officers, employees, contractors and consultants (“Indemnitee”) from and against any and all liability, loss, harm, damage, costs or expense, including legal fees, which the Indemnitee may suffer or sustain as a result of a negligent or unlawful act or omission of, or breach of this Agreement by, the Indemnitor and its respective officers, employees, contractors and consultants in connection with the Project except to the extent that the liability, loss, harm, damage, cost or expense arises from the Indemnitee’s breach of this Agreement, negligent or unlawful acts or omissions or wilful misconduct or bad faith.

6.2. The amount payable by the Subcontractor under this Clause 6 will be limited to the amount the Subcontractor will receive under this Agreement.

## 7. Liability

7.1. Notwithstanding any other provision of this Agreement:

- i. The University’s total liability to the Subcontractor for all claims (other than claims for personal injury or death or property damage including third party property damage) arising out of, or in connection with this Agreement is limited to the Contract Fee; and
- ii. the Subcontractor’s total liability to the University for all claims arising out of, or in connection with this Agreement is limited to the greater of:
  - a) the Contract Fee; and
  - b) any amounts the Subcontractor is able to recover under any insurances taken out under this Agreement.

7.2. The Subcontractor’s total liability to the University for claims for property damage (excluding claims for third party property damage) arising out of, or in connection with this Agreement is limited to any amounts the Subcontractor is able to recover under any insurances taken out under this Agreement or which it would have been able to recover if the insurances had been taken out, or which it would have been able to recover under the insurances, in either case, but for any act, failure or omissions of the Subcontractor.

## 8. Term and Termination

8.1. This Agreement will commence on the Start Date and will expire on the End Date unless terminated earlier in accordance with Clauses **Error! Reference source not found.**(i) to 8.3(iii).

8.2. The University may terminate this Agreement immediately by notice to the other party if:

- (i) the Subcontractor commits a material breach of this Agreement and does not remedy that breach within thirty days of that breach being notified to the Subcontractor;
- (ii) an Insolvency Event occurs with respect to the Subcontractor;
- (iii) the Head Agreement expires or is

terminated for any reason;

8.3. The Subcontractor acknowledges and agrees that if the Client changes the scope of the Services as contemplated by the Head Agreement, then the University may give notice to the Subcontractor notifying the Subcontractor of the change and any respective changes to the scope of the Services and the Fees in the same manner.

8.4. Following termination or expiration of this Agreement, the parties agree that:

- (a) accrued rights and remedies of a party will not be affected;
- (b) each party must return the Confidential Information of the other party in that party’s possession or control, subject to statutory record keeping requirements the party is subject to;
- (c) the Subcontractor will deliver up to the University all Project IP and any licensed Background IP (if it has not already done so); and
- (d) clauses 4, 5, 6, 7 and 13.1 and any other provisions intended by the parties to survive termination or expiration, except to the extent the context requires otherwise, will survive termination or expiration of this Agreement and will not be affected by such termination or expiration.

## 9. Force Majeure Event

9.1. If a Force Majeure Event affecting a party precludes that party (**Precluded Party**) partially or wholly from complying with its obligations (except its payment obligations) under this Agreement, then as soon as reasonably practicable after that Force Majeure Event arises, the Precluded Party must notify the other party of:

- (i) the Force Majeure Event;
- (ii) which obligations the Precluded Party is precluded from performing (Affected Obligations);
- (iii) the extent to which the Force Majeure Event precludes the Precluded Party from performing the Affected Obligations (Precluded Extent);
- (iv) the expected duration of the delay arising directly out of the Force Majeure Event;
- (v) the Precluded Party’s obligation to perform the Affected Obligations will, to the Precluded Extent, be suspended for the duration of the actual delay arising directly out of the Force Majeure Event (Actual Delay); and
- (vi) the other party’s obligations to perform any obligations dependent on the Affected Obligations will be suspended until the Precluded Party resumes performance.

9.2. If the Actual Delay continues for more than 30 Business Days, the other Party may terminate this Agreement immediately by giving notice to the Precluded Party.

## **10. Dispute Resolution**

---

- 10.1.** If a dispute (“Dispute”) arises out of, or relates to, this Agreement (including any Dispute as to breach or termination of this Agreement or as to any claim in tort, in equity or pursuant to any statute) a Party to this Agreement may not commence any court or arbitration proceedings relating to the Dispute, except where the Party seeks urgent interlocutory and/or declaratory relief, unless it has complied with this Clause 9.
- 10.2.** A Party to this Agreement claiming that a Dispute has arisen under or in relation to this Agreement must give written notice to the other Party to this Agreement specifying the nature of the Dispute.
- 10.3.** On receipt of a notice of Dispute the Parties to this Agreement must endeavour to resolve the Dispute expeditiously using a dispute resolution procedure such as mediation, expert evaluation or determination or similar procedures agreed by them.
- 10.4.** If the Parties do not agree, within 7 days of receipt of the notice (or such further period as agreed in writing by them), as to:
- i. The dispute resolution procedures to be adopted;
  - ii. The timetable for all steps in those procedures; and
  - iii. The selection and compensation of the independent person required for such procedure,
- then the Parties shall:
- iv. Mediate the Dispute in accordance with the guidelines for conduct of mediations published by the Law Institute of Victoria; and
  - v. Execute a mediation agreement in a form approved by the Law Institute of Victoria and on request of either Party the President of the Law Institute of Victoria or the President’s nominee will select the mediator and determine the mediator’s remuneration.
- 10.5.** Notwithstanding the existence of a Dispute, both Parties will continue to perform their respective obligations under this Agreement.

## **11. Notices**

---

- 11.1.** Any notice to be given by either Party will be in writing and forwarded to the other Party by hand, post or electronic transmission at the address for notices set out in the cover page (or as last notified by the recipient in writing). A notice will be taken to be delivered or served as follows:
- i) in the case of delivery by hand, when delivered;
  - ii) in the case of delivery by post, two business days after the date of posting; and
  - iii) in the case of electronic mail, on successful transmission to the recipient’s email address and acknowledgment of receipt by the sender.

## **12. Insurance**

---

- 12.1.** The Subcontractor must maintain the same insurances as required under the Head Agreement and must, upon request, provide the University with evidence that it holds such insurances.

## **13. GST**

---

- 13.1.** Terms used in this Clause 13 have the meaning given to them in the A New Tax System (Goods and Services Tax) Act 1999 (Cth), or any other act relating to the imposition or admiration of a goods and services tax.
- 13.2.** Any consideration specified as payable by one party to the other party under this agreement is expressed as a GST exclusive amount, unless the amount is defined to be inclusive of GST
- 13.3.** If a taxable supply made under this agreement is varied, in addition to the amount of consideration required to be refunded or paid in addition, the amount of GST paid or payable in respect of the variation must be refunded or paid by the relevant party as the case requires
- 13.4.** A party’s right to payment under this clause is subject to a Tax Invoice or Adjustment Note as the case may be and which complies with the GST Act being delivered to the party liable to pay for that supply at or before the time of payment or variation.
- 13.5.** A Tax Invoice or Adjustment Note as required under Clause 3 includes any document or record treated by the Commissioner of Taxation as enabling the claiming of an input tax credit for which an entitlement arises under this agreement
- 13.6.** If this agreement requires a party to pay for, reimburse or contribute to any expense, loss or outgoing of another party, and that other party can obtain an input tax credit on an acquisition associated with the expense, loss or outgoing, the amount required to be paid, reimbursed or contributed by the first party will be the amount of the expense, loss or outgoing reduced by the amount of that input tax credit
- 13.7.** The rate of GST payable under this agreement is the amount payable under the A New Tax System (Goods and Services Tax) Act 1999 (Cth), any change in GST rates will amend the rate and amount of GST payable under this agreement

## **14. General**

---

### **14.1. Further assurances**

Each party must do all things necessary or desirable to give effect to the provisions of this Agreement including by signing all documents and performing all acts.

### **14.2. Entire agreement**

This Agreement contains the entire agreement of the parties; and supersedes all prior representations, conduct and agreements that may have been made by the University about the Services

### **14.3. Costs**

Each party is responsible for its own costs of entering into and performing this Agreement.

**14.4. Governing Law**

This Agreement is governed by the laws of the State of Victoria and the Parties agree to irrevocably submit to the jurisdiction of the courts of the State of Victoria.

**14.5. Severability**

To the extent that any portion of this Agreement is void or otherwise unenforceable then that portion will be severed and this Agreement will be construed as if the severable portion had never existed.

**14.6. No agency**

This Agreement does not create a relationship of employment, trust, agency or partnership between the parties.

**14.7. Waiver**

A right under this Agreement may only be waived in writing signed by the party granting the waiver, and is effective only to the extent specifically set out in the waiver.

**14.8. Discretion**

Unless otherwise expressly contemplated, where a provision of this Agreement contemplates that a party may exercise its discretion then that party is entitled to exercise that discretion absolutely, with or without conditions and without being required to act reasonably or give reasons.

**14.9. No assignment or subcontracting**

The Subcontractor must not assign or subcontract its rights or obligations to any third party.

## SCHEDULE 1.

### 1. Provision access to reliable soil and other on-farm data

Activity	Description	Output/quantity (if applicable)	Timeline
Contribute soil test data	Provide soil test result data from grower group managed research trial sites using the VAS portal data upload interface.	Approx. number of locations, number of datasets and approximate number of results	Upload results annually (mid year) for period of the agreement
Link regional soil moisture sensor data with VAS	Work with probe system provider(s) and CeRDI to facilitate access to soil moisture sensor network data for display on VAS portal.	12 soil probe locations	Quarter 3 of Year 1

### 2. Deliver extension activities to support VAS adoption

Activity	Description	Output/quantity	Timeline
Awareness raising	Prepare and publish news items on group website, newsletters and annual report.	Two website news articles, two newsletter articles, one report article	Annually for period of agreement
Training and feedback	Work with VAS team to organise and run VAS demonstration workshop for members/farmers. Collect feedback on VAS portal functionality and suggestions for improvement.	Two 2hr workshop events. Brief report on workshop with summary of collected feedback.	Quarter 2 of Year 1

### 3. Contribute to the delivery of tangible and attractive benefit to farming groups and other VAS participants

Activity	Description	Output/quantity	Timeline
VAS workshop	Participate in annual VAS workshop.	One workshop event	Annually
Co-develop VAS use case	Create focus group and develop an example of how VAS portal functionality or tools could support research or decision making.	Documented use case	Quarter 2 of Year 1
VAS platform testing	Contribute to testing development versions of the VAS system and provide feedback on system functionality and performance	1-3 days Testing feedback summary report	Quarter 2 of Year 1

## **CRC for HIGH PERFORMANCE SOILS PROJECT AGREEMENT (PJA 2.3.002)**

# APPENDIX C DATA INPUT TEMPLATES

Self-serve system data input templates in an Excel spreadsheet comprising six worksheets. Worksheets 3, 4 & 5 require data entry.

[Excel worksheet 1: READ ME FIRST]

## DATA SUBMISSION INSTRUCTIONS

- Before submitting data you must consider if this is a single dataset with shared dataset information. This means that it is all from a single project, data provided by a single laboratory and have shared data ownership/organisational details. These are the things outlined in the **DATASET INFO** sheet which should be filled out first. If you have a dataset where there is varying dataset information, this ideally should be split into separate sheets and submitted separately.
- Fill out the **DATASET INFO** sheet first. Mandatory fields are marked with a \*. Some have drop down options. If you are unclear what is being asked for, check the 'What Headings Mean' sheet.
- Soil data goes in the **SOIL DATA RESULTS** sheet. Mandatory Fields are marked with a \*. See the top row for data examples in these columns. Some fields will have drop down values. We cannot accept data that is missing mandatory fields. It is strongly encouraged to provide answers to all non-mandatory fields you have details for as this will be of benefit to the organisation provisioning the data. For example lab barcodes may enable you to investigate suspect data values in the future with the laboratories. It is best not to assume something is not important.
- The **What headings mean** sheet provides additional information to help you understand the meaning of heading terms.
- A fundamental benefit of VAS is the ability to standardise soil data to allow comparison between datasets for data providers, which is possible by specifying details on the exact lab procedure/methods. Please provide as much information as possible about the lab results you are including. Many labs provide method codes which are helpful and/or have interpretation guides or other documentation listing what methods they have used. See examples of **IDEAL to UNSUITABLE SOIL DATA METHOD INFO** in the table to the right. Sometimes its useful to include calculations in datasets that are in addition to what was provided by the lab. For these to be included, it is necessary to make it clear in the Method row how the values were calculated.

SEE EXAMPLES OF IDEAL TO UNSUITALE SOIL DATA METHOD INFO ----->

RATING	IDEAL. Rayment and Lyons Method Codes (aka the Greenbook) are ideal. Many labs provide interpretation guides with this info or it may be already listed on the lab sheets.	Good. Often lab sheets provide lab specific method codes. If you are not able to find the Rayment and Lyons method code or its even a in house lab method then providing their codes along with as much details of the method (e.g. the fact it used 1M Ammonium Acetate) is helpful.	OK. In this instance no method codes were known but the lab sheet had information about the method (1M Ammonium Acetate).	Bare Minimum. No details can be provided about the method. But at least it can be said its an exchangeable Mg result (vs another kind of Magnesium measurement of which there are many).	NOT SUITABLE. This data does not meet minimum standards. It is not clear what kind of magnesium measurement it is. Additionally no units of measure have been provided which is mandatory.
Lab Test/Observation	Exchangeable Mg	Exchangeable Mg	Exchangeable Mg	Mg	Mg
Method	Rayment and Lions Method 15D3	1M Ammonium Acetate APAL method code TMs-015D	1M Ammonium Acetate	Exchangeable Mg. Unknown method.	-
Unit of Measure	mg/kg	mg/kg	mg/kg	mg/kg	-

- The submission of **SPATIAL DATA** is a requirement for submission to the VAS platform. This can be in the form of Latitude/Longitude or other GPS readings, or as a GIS file (Shapefile, Google Earth KLM, etc). It is a requirement that this spatial data is linked consistently with a **site id\*** that is also used in the **SOIL DATA RESULTS** sheet. While there is a column in the **SOIL DATA RESULTS** to give these sites a name (e.g. a paddock name), it is necessary to link them with a numerical Site ID number which can be created just for this dataset to prevent disconnection of spatial information due to misspellings, etc. If you don't already use site numbers simply adding sequential numbers starting at 1 is usually the easiest.

- If there are any original lab data sheets available for a dataset it is strongly encouraged to attach them with this data template along with any additional supplementary information such as reports, photos, etc along with details about what they are and how they relate to the dataset.

[Excel worksheet 2: What headings mean]

DATSET INFO FIELDS	MANDATORY?	INFORMATION
<b>Dataset Name (e.g. file name you can recognise)*:</b>	y	A name that will make it easy for you to identify the dataset (e.g. filename of the original datasheet) should you need to refer back to it in the future.
<b>Project Name*:</b>	y	Name of the project that this dataset is part of.
<b>Organisation Submitting Data*:</b>	y	Name of the organisation submitting the dataset with this template
<b>Contact Name*:</b>	y	Name of a person at organisation who could answer questions about this dataset if necessary
<b>Contact Email/Phone*:</b>	y	Best contact details for that person
<b>Name of Lab who Analysed Data*:</b>	n	Name of soil lab used. Dropdown menu or write your own. If unknown select unknown. If dataset is not from a soil lab select N/A
<b>Lab Sheet Number:</b>	n	If there is a number provided by a lab on the original sheet for the whole dataset provide it here. Note this is different to sample ID numbers, its for the whole dataset not one sample.
<b>Soil Sampling Method Used*:</b>	y	Select from dropdown or write your own.
<b>Is this dataset a subset of a bigger dataset under the same project name?*</b>	y	Yes/No/Unsure dropdown. Use this where the data is only one part of a dataset that has been split into multiple sheets (e.g. due to using different labs). Its essential all subsets of the data are given the same project name for linkage purposes.
<b>Is this dataset related to another dataset already submitted or to be submitted (e.g. resampling of the same sites)*:</b>	y	YES/NO/UNSURE dropdown. This relates to when a dataset is related to another dataset that is submitted or to be submitted to VAS. But it is a separate dataset (e.g. different project names) and is related in some other way that must be specified in the next question
<b>How is the other datasets related?</b>	if Yes is selected in above question	Fill this in if you are aware the dataset you are submitting is related to another separate dataset. You are using this to explain in what way is it related. For example is it resampling the same sites at a different time? Explain this as best as you can.
<b>What is the name of this related dataset:</b>	n	Name of the related dataset - this should be the Dataset name ideally rather than just the project name as projects may have multiple datasets.
<b>What is the relationship between this dataset and the above mentioned dataset?:</b>	Y if there is a related dataset mentioned above.	Try and explain how this dataset is related to the other data set(s). For example are the same exact site locations being tested, it is in the same area but different sites?
<b>Spatial Information about sample locations has been provided in the following format*:</b>	y	Dropdown list or add your own.
<b>GPS data projection for this spatial data is*:</b>	y	Dropdown list or add your own. If you don't know 'unsure' is an option
<b>How accurate is the Spatial Information in relation to the real world location of the soil samples?*</b>	y	Drop down list. Do the GPS locations for the site relate to the exact location of the soil sample? Is the location a centroid of the location of multiple locations used to make a composite sample? Is the accuracy only to the level of the centre of the paddock that was sampled?

Do you have an original Lab sheet you can submit with this dataset?*	y	drop down or type Yes/No
What Organisation Owns this Dataset?*	y	Name of data owner organisation
Name of Funding Organisation*:	y	Name of funding organisation. Select unknown if not known.
Name of additional Contributing Organisations (if any):		Any additional organisations that have contributed to this dataset.
Role of additional Organisations listed:	Y if there is one or more additional organisations listed	Dropdown list. To understand the meaning of these dropdown options see organisational role definitions listed at the bottom of this sheet.

Do you have any additional documents you can submit that relate to this dataset (e.g. reports, photos)?*	y	Yes/No dropdown
--	---	-----------------

site id*	y	<b>The term 'site' refers to the location where a soil sample is taken. Provide an ID number for the location. If you do not have one already give each site location a sequential number. This is important for linking spatial data.</b>
site name*	y	Name of the site
additional name (if helpful)	n	If there are any additional names of the site you would like included
details of additional name	n	If an additional name of a site has been provided what kind of name is this? E.g. is it a paddock name etc
site land use	n	The use of the land under the official ALUM classification scheme. Dropdown provided. Details on ALUM are here: <a href="https://www.awe.gov.au/abares/aclump/land-use/alum-classification">https://www.awe.gov.au/abares/aclump/land-use/alum-classification</a>
site description	n	Any additional description of the site that is helpful
sampling date*	y	Date the soil was collected in the field. Use Australian date format (e.g. 01/12/2022). If day or month is not known choose the first of month or first of January of the year
date soil analysed	n	Date the soil was analysed by the lab. Sometimes provided on lab sheets.
soil type	n	If the soil that is analysed has an official or unofficial name give it here
soil type standard used	n	Choose from the dropdown what standard was used to name the soil or select no standard if it is an unofficial name which has meaning to the data provider only.
soil sample number	n	Soil sample number. Either designated by soil sample collector or lab
soil sample number type	n	What was the sample number actually labelled on the lab sheet? Put that heading here (e.g. Lab Sample Number, Barcode number etc)
additional sample number	n	If a sample has more than one number (e.g. a sample number and a barcode number) you can put the addition number here.
additional sample number type	n	What is this additional sample number labelled as on the lab sheet? E.g. Barcode number
upper depth (cm)*	y	A numeric measurement
lower depth (cm)*	y	A numeric measurement

<b>Lab Test/Observation*</b>	y	The type of soil test result the data underneath is. E.g. Exchangeable magnesium
<b>Method*</b>	y	The method used by the lab (the how a calculation was made for calculated results). See the information in the READ ME FIRST sheet for clear examples
<b>Unit of Measure*</b>	y	Unit of measure (UOM) of the results (e.g. mg/kg). This is mandatory. For results like pH or ratios that don't have an obvious UOM select # from the dropdown. You may also type other options if needed

#### Definition of Organisation Roles

author	party who authored the resource'
co Author	party who jointly authors the resource'
collaborator	party who assists with the generation of the resource other than the principal investigator'
contributor	party contributing to the resource'
custodian	party that accepts accountability and responsibility for the resource and ensures appropriate care and maintenance of the resource'
distributor	party who distributes the resource'
editor	party who reviewed or modified the resource to improve the content'
funder	party providing monetary support for the resource'
mediator	a class of entity that mediates access to the resource and for whom the resource is intended or useful'

originator	party who created the resource'
owner	party that owns the resource'
point Of Contact	party who can be contacted for acquiring knowledge about or acquisition of the resource'
principal Investigator	key party responsible for gathering information and conducting research'
processor	party who has processed the data in a manner such that the resource has been modified'
publisher	party who published the resource'
resource Provider	party that supplies the resource'
rights Holder	party owning or managing rights over the resource'
sponsor	party who speaks for the resource'
stakeholder	party who has an interest in the resource or the use of the resource'
user	party who uses the resource'

## DATA SET INFORMATION (METADATA)

MANDATORY FIELDS ARE MARKED WITH \*. IF UNKNOWN  
WRITE UNKNOWN or N/A if not applicable

	Answer	Additional Explanation if Required
<b>Dataset Name (eg file name you can recognise)*:</b>		
<b>Project Name*:</b>		
<b>Organisation Submitting Data*:</b>		
<b>Contact Name*:</b>		
<b>Contact Email/Phone*:</b>		
<b>Name of Lab who Analysed Data (dropdown or type)*:</b>		
<b>Lab Sheet Number:</b>		
<b>Soil Sampling Method Used (dropdown or type)*:</b>		
<b>Is this dataset a subset of a bigger dataset under the same project name?*</b>		
<b>Is this dataset related to another dataset already submitted or to be submitted (eg resampling of the same sites)*:</b>		
<b>What is the name of this related dataset:</b>		
<b>What is the relationship between this dataset and the above mentioned dataset?:</b>		
<b>Spatial Information about sample locations has been provided in the following format*:</b>		
<b>GPS data projection for this spatial data is*:</b>		

<b>How accurate is the Spatial Information in relation to the real world location of the soil samples?*</b>		
<b>Do you have an original Lab sheet you can submit with this dataset?*</b>		
<b>What Organisation Owns this Dataset?*</b>		
<b>Name of Funding Organisation*:</b>		
<b>Name of additional Contributing Organisations (if any):</b>		
<b>Role of additional Organisations listed:</b>		
<b>Do you have any additional documents you can submit that relate to this dataset (eg reports, photos)?*</b>		
<b>Provide any additional information about this dataset you think is relevant here. Eg. Details about the project it is part of, soil sampling strategy, research design etc where this is not provided in an attached report.</b>		

[Excel worksheet 4: SOIL DATA RESULTS]

site id*	site name*	additional name (if helpful)	details of additional name	site land use (ALUM)	site description	sampling date*	date soil analysed	soil type	soil type standard used	soil sample number	soil sample number type	additional sample number	additional sample number type	upper depth (cm)*	lower depth (cm)*	SOIL DATA RESULTS Lab Test/Observation* Method* Unit of Measure*	SOIL DATA EXAMPLE Exchangeable Mg Rayment and Lions Method 15D3 meq/hg	CUT AND PASTE SOIL DATA RESULTS HERE below grey example
1	JP45	Joe's Paddock	Paddock Name	Cereals	Wheat paddock	1/12/2022	30/12/2022	Chromosol	Australian Soil Classification	4565	Lab sample number	23412341234	Barcode number	0	10		8.12	
2	JP45	Joe's Paddock	Paddock Name	Grazing native vegetation	Mixed grazing	1/12/2022	30/12/2022	Black Soil	No standard. Field observation	4566	Lab sample number	2342342323	Barcode number	0	10		8.56	

A	B	C	D	E	F	G	H	I
site id*	site name*	additional name (if helpful)	details of additional name	site land use (ALUM)	site description	sampling date*	date soil analysed	soil type
1	JP45	Joe's Paddock	Paddock Name	Cereals	Wheat paddock	1/12/2022	30/12/2022	Chromosol
2	JP45	Joe's Paddock	Paddock Name	Grazing native vegetation	Mixed grazing	1/12/2022	30/12/2022	Black Soil

J	K	L	M	N	O	P
soil type standard used	soil sample number	soil sample number type	additional sample number	additional sample number type	upper depth (cm)*	lower depth (cm)*
Australian Soil Classification	4565	Lab sample number	23412341234	Barcode number	0	10
No standard. Field observation	4566	Lab sample number	2342342323	Barcode number	0	10

Q	R	S	T	U
SOIL DATA RESULTS	SOIL DATA EXAMPLE	CUT AND PASTE SOIL DATA RESULTS HERE below grey example		
Lab Test/Observation*	Exchangeable Mg			
Method*	Rayment and Lions Method 15D3			
Unit of Measure*	meq/hg			
	8.12			
	8.56			

**[Excel worksheet 5: SPATIAL DATA]**

SITE ID*	SITE NAME	lat/northing	long/eastings	elevation (m)
1	JP45	-34.07448	117.4219	125

[Excel worksheet 6: dropdown menus (don't touch)]

Units of measure	ALUC Landuse	Labs	Soil Sampling Methods	yes/no/unsure	Spatial formats	Projections	yes/no	Organisaton roles
mg/kg	<b>Nature conservation</b>	APAL	Coring	yes	lat longs	WGS84	yes	author
%	Strict nature reserves	ALS	Soil Pit	no	easting northings	Other (please type)	no	co Author
meq/hg	Wilderness area	AgVita Analytical	Shovel Sample	unsure	Shapefile	Unsure		collaborator
cmol/kg	National park	BioScience	Auger		Google Earth KLM			contributor
kg/ha	Natural feature protection	CERAR	Unknown					custodian
dS/m	Habitat/species management area	ChemCentre						distributor
pH	Protected landscape	CSBP						editor
ppm	Other conserved area	CSIRO						funder
#	<b>Managed resource protection</b>	DPI						mediator
	Biodiversity	EAL						originator
	Surface water supply	Nutrient Advantage						owner
	Groundwater	University of Adelaide						point Of Contact
	Landscape	N/A						principal Investigator
	Traditional Indigenous uses	Unknown						processor
	<b>Other minimal use</b>							publisher
	Defence land - natural areas							resource Provider
	Stock route							rights Holder
	Residual native cover							sponsor
	Rehabilitation							stakeholder
	<b>Grazing native vegetation</b>							user
	<b>Production native forests</b>							
	Wood production forestry							
	Other forest production							
	<b>Plantation forests</b>							
	Hardwood plantation forestry							
	Softwood plantation forestry							
	Other forest plantation							
	Environmental forest plantation							
	<b>Grazing modified pastures</b>							

	Native/exotic pasture mosaic							
	Woody fodder plants							
	Pasture legumes							
	Pasture legume/grass mixtures							
	Sown grasses							
	<b>Cropping</b>							
	Cereals							
	Beverage and spice crops							
	Hay and silage							
	Oilseeds							
	Sugar							
	Cotton							
	Alkaloid poppies							
	Pulses							
	<b>Perennial horticulture</b>							
	Tree fruits							
	Olives							
	Tree nuts							
	Vine fruits							
	Shrub berries and fruits							
	Perennial flowers and bulbs							
	Perennial vegetables and herbs							
	Citrus							
	Grapes							
	<b>Seasonal horticulture</b>							
	Seasonal fruits							
	Seasonal flowers and bulbs							
	Seasonal vegetables and herbs							
	<b>Land in transition</b>							
	Degraded land							
	Abandoned land							
	Land under rehabilitation							

No defined use								
Abandoned perennial horticulture								
<b>Irrigated plantation forests</b>								
Irrigated hardwood plantation forestry								
Irrigated softwood plantation forestry								
Irrigated other forest plantation								
Irrigated environmental forest plantation								
<b>Grazing irrigated modified pastures</b>								
Irrigated woody fodder plants								
Irrigated pasture legumes								
Irrigated legume/grass mixtures								
Irrigated sown grasses								
<b>Irrigated cropping</b>								
Irrigated cereals								
Irrigated beverage and spice crops								
Irrigated hay and silage								
Irrigated oilseeds								
Irrigated sugar								
Irrigated cotton								
Irrigated alkaloid poppies								
Irrigated pulses								
Irrigated rice								
<b>Irrigated perennial horticulture</b>								
Irrigated tree fruits								
Irrigated olives								
Irrigated tree nuts								
Irrigated vine fruits								
Irrigated shrub berries and fruits								
Irrigated perennial flowers and bulbs								
Irrigated perennial vegetables and herbs								
Irrigated citrus								
Irrigated grapes								

	<b>Irrigated seasonal horticulture</b>							
	Irrigated seasonal fruits							
	Irrigated seasonal flowers and bulbs							
	Irrigated seasonal vegetables and herbs							
	Irrigated turf farming							
	<b>Irrigated land in transition</b>							
	Degraded irrigated land							
	Abandoned irrigated land							
	Irrigated land under rehabilitation							
	No defined use - irrigation							
	Abandoned irrigated perennial horticulture							
	<b>Intensive horticulture</b>							
	Production nurseries							
	Shadehouses							
	Glasshouses							
	Glasshouses - hydroponic							
	Abandoned intensive horticulture							
	<b>Intensive animal production</b>							
	Dairy sheds and yards							
	Feedlots							
	Poultry farms							
	Piggeries							
	Aquaculture							
	Horse studs							
	Saleyards/stockyards							
	Abandoned intensive animal production							
	<b>Manufacturing and industrial</b>							
	General purpose factory							
	Food processing factory							
	Major industrial complex							
	Bulk grain storage							
	Abattoirs							

	Oil refinery							
	Sawmill							
	Abandoned manufacturing and industrial							
	<b>Residential and farm infrastructure</b>							
	Urban residential							
	Rural residential with agriculture							
	Rural residential without agriculture							
	Remote communities							
	Farm buildings/infrastructure							
	<b>Services</b>							
	Commercial services							
	Public services							
	Recreation and culture							
	Defence facilities - urban							
	Research facilities							
	<b>Utilities</b>							
	Fuel powered electricity generation							
	Hydro electricity generation							
	Wind electricity generation							
	Solar electricity generation							
	Electricity substations and transmission							
	Gas treatment, storage and transmission							
	Water extraction and transmission							
	<b>Transport and communication</b>							
	Airports/aerodromes							
	Roads							
	Railways							
	Ports and water transport							
	Navigation and communication							
	<b>Mining</b>							
	Mines							
	Quarries							

Tailings								
Extractive Industry not in use								
<b>Waste treatment and disposal</b>								
Effluent pond								
Landfill								
Solid garbage								
Incinerators								
Sewage/sewerage								
<b>Lake</b>								
Lake - conservation								
Lake - production								
Lake - intensive use								
Lake - saline								
<b>Reservoir/dam</b>								
Reservoir								
Water storage - intensive use/farm dams								
Evaporation basin								
<b>River</b>								
River - conservation								
River - production								
River - intensive use								
<b>Channel/aqueduct</b>								
Supply channel/aqueduct								
Drainage channel/aqueduct								
Stormwater								
<b>Marsh/wetland</b>								
Marsh/wetland - conservation								
Marsh/wetland - production								
Marsh/wetland - intensive use								
Marsh/wetland - saline								
<b>Estuary/coastal waters</b>								
Estuary/coastal waters - conservation								

	Estuary/coastal waters - production							
	Estuary/coastal waters - intensive use							

# APPENDIX D DATA AGGEGATOR

## BACKGROUND

The usual approach when establishing a database is to design it to meet specific user requirements via use cases, mission statements or such like. Similarly, when designing a data structure to exchange domain specific data, a domain specific model is usually created. Domain data exchange examples include geology (GeoSciML), mining and exploration (EarthResourceML), groundwater (GroundWaterML2), soils (ANZSoilML), and water and energy supply and consumption (WESCML). In these database designs and models all the features/objects that interest the domain are specified, along with the properties that describe these features, and the relationships and the relationship properties that exist between these features. In a relational database these are hypostasised as tables and columns.

This 'hard-type' approach to design presents a number of challenges. These may include finding the right people with the skills and resources, specifying scope given most domains have overlap with other domains (e.g. Boreholes). In addition, the resulting domain model if it accurately reflects the real world, is usually complicated, making it difficult to use by both domain experts and non-domain users.

An alternative is to specify the domain-specific features and relationships between them, but 'soft-type' the properties. That is, populate the features (classes) with attribute:value pairs. An example of this design is the soil/terrain model SoTerML. This approach is more readily extendable, but requires a vocabulary of appropriate attributes for each class. It moves the problem from the data structure (schema) design to specifying vocabulary or look-up tables.

VAS took this last approach, but to wrap it up in a generic 'Observation' framework. So instead of specifying what all the properties any feature, such as a Soil or Soil Horizon or Landform, may have, VAS collects all the observations, measurements and interpretations that have been made on these features. The domain-specific feature types that may be stored in the database are not specified as tables, but rather as a 'feature-type' property that comes from a vocabulary listing. This provides the flexibility of adding new feature-types without any database schema changes. It also means that once the O&M model schema is understood, the data structure is understood for all the domains where observations and measurements are made.

The 'Observation framework' chosen was based on Observation & Measurements (O&M) version 2 (see "CeRDI Observations System Overview" below), although it is also considered to be compatible with the more recent O&M version 3, and the W3C derivative, 'Sensor-Observation-Sampling-Actuator' (SOSA). Instead of identifying domain specific features and properties, O&M specifies what features and properties are common to all Observations. For example, one of these O&M features is the "FeatureOfInterest" (Fol). The Fol can be any thing about which observations are made, and its type identified via its 'featureType' property, which can come from a vocabulary list. An O&M property is the "ObservableProperty", the Fol property that is being observed or measured. Again these ObservableProperties don't need to be specified in the database design, but can come from a vocabulary list.

How do we know which "ObservableProperties" can be associated with which "featureTypes"? This is normally the remit of the domain experts during the database/exchange language design phase. However, VAS has taken a different (and new and previously untried?) approach.

All the observations that have been stored in VAS have been made by experts in their domain. So for example, all the soil layer laboratory results or the soil profile descriptions, have been made

either by soil scientists, or on observed properties specified by soil scientists. So all the Observable Properties that have been associated with the same feature type (such as a 'Soil Layer'), represent a collection of the attributes of that feature. What about if observers use an inappropriate Observable Property, such as 'happiness quotient' for a Soil Horizon? There is nothing built-in to VAS preventing this. However, it is unlikely since all the observations and measurements are provided by domain experts. Even if 'wrong' observed properties are entered it is unlikely to influence the data usage, as it is unlikely to be discovered as part of a search, and if found the expert user can choose to treat it as bad data if it doesn't fit their concept of an appropriate Soil Horizon property.

Importantly, this list of Observable Properties is not unique to a specific domain: for example 'density' or 'Nitrogen concentration' or 'pH' may be Observable Properties of features from many domains. So a single list of Observable Properties can be maintained, rather than a list specific to each domain.

Further, the observation result values that have been assigned to these attributes identify that attribute's appropriate vocabulary. For example, terms such as 'Loam', or 'Loamy sand' specify the vocabulary terms that can be associated with the soil texture observed property, but these terms would not have been associated with other observed properties, such as coarse fragment shape.

VAS captures all the results, observable properties, and procedures that have been made for each of the soil feature types (such as Soil Body, Soil Layer, Soil Horizon, Soil Profile). If VAS captures sufficient breadth of Observations (including Measurements and Interpretations) then it is possible to identify all those properties usually associated with each feature, the procedures used for generating the result values for each property, and the range of appropriate results for each procedure for each observable property. As the number and range of observations increases, this model should approach a model generated from the traditional domain-expert process.

# CERDI OBSERVATIONS SYSTEM OVERVIEW

The primary purpose of the observations system is to collect and store observation and measurement data and publish this data for researchers, industry, government and the public. It uses the ISO19156<sup>9</sup> and Open Geospatial Consortium (OGC)<sup>10</sup> Observations and Measurements (O&M) conceptual model to store field and laboratory environmental data in a domain independent structure. The observation model states:

“An **Observation** is an action whose **result** is an estimate of the value of some **property** of the **feature-of-interest**, obtained using a specified **procedure**” (SJD Cox)

The key insights are

1. to separate
  - the *observation act* from
  - the *procedure*
    - which may be used for other observationsand
  - the *feature-of-interest*
    - which has many properties, the values of each of which may be estimated more than once, at different times or using different procedures
2. and to recognise that the outcome of an Observation is a *result*
  - the value of which constitutes an estimate of a value of a property
    - which may be a value or range of values if a measurement, or a term, a term range or a description if an assertion

In addition to standardising the data structure, the system makes use of existing domain-specific controlled vocabularies and ontologies to standardise the semantic content.

The system is a custom development by CeRDI, primarily using PHP and PostgreSQL, services are provided via PostGrest.

---

<sup>9</sup> Cox, Simon Jonathan David (2011). "[ISO 19156:2011 Geographic information – Observations and measurements](#)". International Organization for Standardization. [doi:10.13140/2.1.1142.3042](#)

<sup>10</sup> "[OGC Abstract Specification Topic 20: Observations and measurements](#)". 2010. Retrieved 2010-11-22.

# VAS CONCEPTUAL OVERVIEW

The VAS technical system is illustrated below (Figure D.1).

The VAS soil research data is aggregated via the CeRDI Observations System as described in the previous section. The data aggregator allows participants to login to a self-serve system to submit a dataset using a template that collects metadata and the soil data observations. Educational materials in the form of short “how to” videos are provided as instructions. These data are then manually checked and mapped into the aggregator using the standard schema.

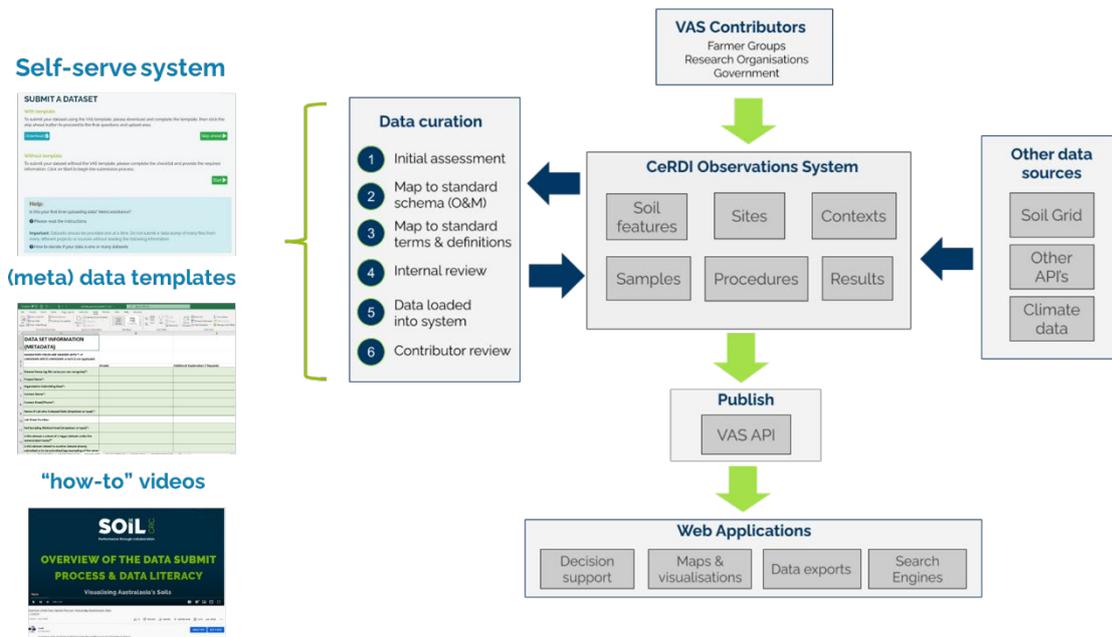


Figure D.1: The VAS technical system overview.

Soil and ancillary data from other sources that interoperably serve standard data sets, such as ANSIS, New Zealand Soil Classification, Soil and Landscape Grid of Australia (SLGA), or the Bureau of Meteorology, can be interoperably consumed.

The data is published, subject to the data custodians consent, via and API to the VAS portal, other systems, or users.

Future work will include improving the self-serve system to include a data management system to give the data custodians more control, and make the input metadata mapping as seamless as possible.



is used for (e.g. "length") is required. The unit of measure identifier provides a link to an external ontology maintained by an appropriate authority.

## OBSERVATIONS

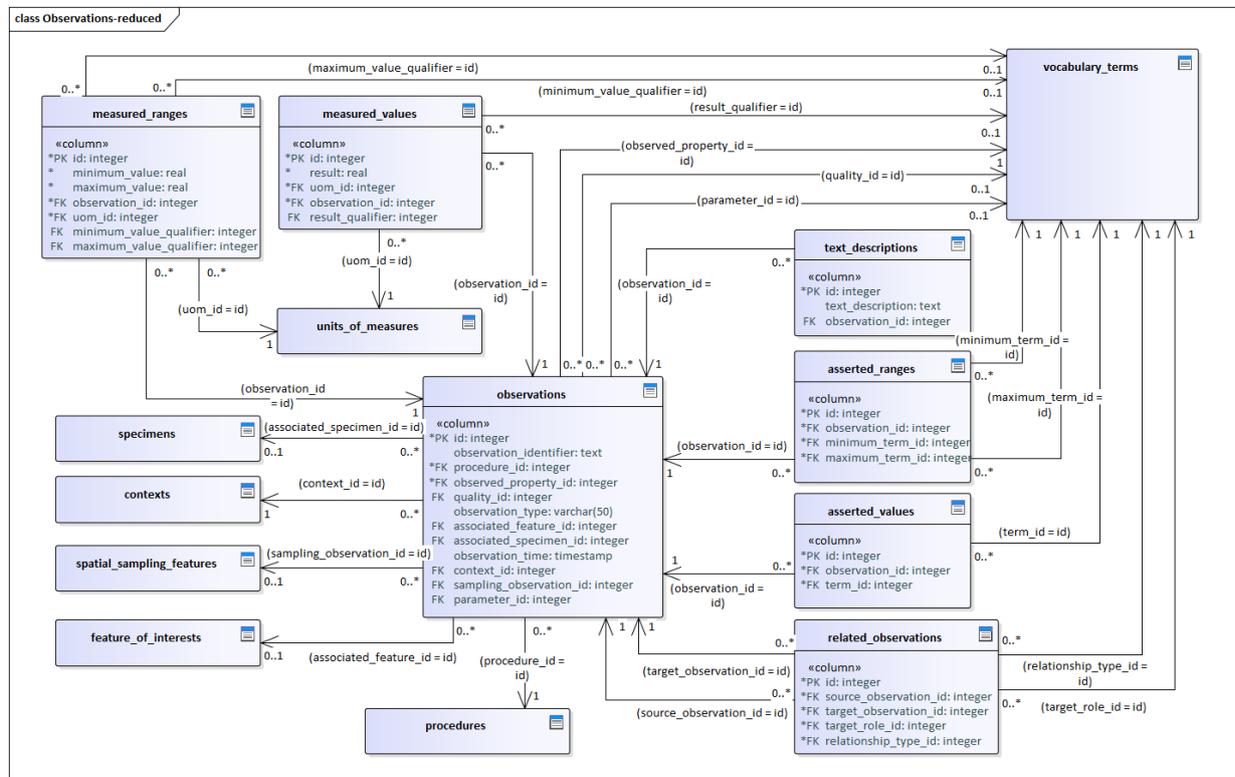


Figure D.3: "observations" and the results tables. Columns of related tables and associated relationships are not shown

### 'observations' table description:

The **observations** table records information about individual observations (measurement or assertions) made on the soil feature.

1. The type of Observation (e.g. "Laboratory Measurement", "Field Observation"). This could potentially be expanded to include "Application Rate", "Crop Yield", etc.
2. The result of the Observation could be a numerical range, single numerical result, a range of terms or a single term. This could be expanded to include raster, time series, image results.

COLUMN NAME	DATA TYPE	NOT NULL	COMMENTS
id	integer	True	The database id
observation_identifier	text	False	The persistent identifier of the individual Observation (URI)
procedure_id	integer	True	The foreign key to the procedures table identifying the procedure used to generate the observation

observed_property_id	integer	True	Foreign key to the vocabulary_terms table identifying the observed property being measured (e.g. "K concentration", "pH").
quality_id	integer	False	Foreign key to the vocabulary_terms table to describe the quality of the observation.
observation_type	varchar(50)	False	The type of Observation (e.g. "Laboratory Measurement", "Field Observation"). This could be expanded to include "Application Rate", "Crop Yield", etc. Future upgrade to a Foreign key to the vocabulary_terms table
associated_feature_id	integer	False	Foreign key to the soil feature the observations were made on (the real world FeatureOfInterest such as a Soil Body, Soil Horizon, Soil Profile, Soil Layer).
associated_specimen_id	integer	False	The foreign key to the (Soil) Specimen that the observation was made on.
observation_time	timestamp	False	The time and date of the observation. The O&M model allows for additional time properties, such as the time the observation was applicable, the time the observation was made, the time the observation is valid for. These variations have not been included here.
context_id	integer	False	A foreign key to the contexts table specifying the context in which the observation was made. Usually a job or set of tasks undertaken by a single observer/machine as part of a broader project. Links Job, Project, Client information
sampling_observation_id	integer	False	A foreign key to the feature that was used to sample the real world environmental feature such as SoilProfile, or Specimen that the observation was made on. SpatialSamplingFeatures may be "Plot", "Site", "Station", "Paddock", etc.

parameter_id	integer	False	<p>Describes an arbitrary event-specific parameter. This might be an environmental parameter, an instrument setting or input, or an event-specific sampling parameter that is not tightly bound to either the feature-of-interest or to the observation procedure.</p> <p>In some contexts the <i>Observation::procedure</i> is a generic or standard procedure, rather than an event-specific process. In this context, parameters bound to the observation act, such as instrument settings, calibrations or inputs, local position, detection limits, asset identifier, operator, may augment the description of a standard procedure.</p> <p>EXAMPLE The fraction of the soil ('coarse fraction', 'fine fraction', 'whole soil') from which the particle size proportion was taken.</p>
--------------	---------	-------	---

#### 'related\_observations' table description:

The **related\_observations** table records information about relationships between individual observations.

In addition to identifying observations that were made as part of a project for example, this table could be used to relate previous application or yield results to current results.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	The database id
source_observation_id	integer	True	The foreign key to the observation that is the source in the relationship
target_observation_id	integer	True	The foreign key to the Observation that is the target in the relationship
target_role_id	integer	True	The foreign key to the vocabulary_terms table that specifies the role that the target Observation plays in the relationship (e.g. "previous application rate", "related observation", "alternative result").
relationship_type_id	integer	True	The foreign key to the vocabulary_terms table that specifies the type of relationship between the two Observations

The results of an observation may be one of five types:

A numeric value resulting from a measurement (**measured\_values**);

A numeric range resulting from a set of measurements (**measured\_ranges**);

A term value resulting from an asserted value (**asserted\_values**);

A range of term values resulting from a set of asserted values (**asserted\_ranges**);

A description (**text\_descriptions**).

Depending on the type of result, these are stored in separate tables.

*'measured\_values' table description:*

Table to store empirical results from laboratories, field measuring devices etc. where the result is a single numeric value.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	The database id
result	real	True	The numerical result of the Observation
uom_id	integer	True	The Unit Of Measure for the result. This is a foreign key to the units_of_measures Table.
observation_id	integer	True	The Observation that the result relates to.
result_qualifier	integer	False	Foreign key to the vocabulary_terms table. Qualifies the result. Examples are '<', '<=', '>', '>=', '~', 'below detection limit', 'above detection limit'. Also to capture the gml:nilReasonTypes ('unknown', 'withheld', 'inapplicable', 'missing', 'template').

*'measured\_ranges' table description:*

Table to store empirical results from laboratories, field measuring devices etc. where the result is a single numeric value.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	The database id for the measured range
minimum_value	real	True	The smallest numerical value in the range
maximum_value	real	True	The largest numerical value in the range
observation_id	integer	True	The foreign key to the Observation which created this result
uom_id	integer	True	A foreign key to the units_of_measures table for the Unit Of Measure

minimum_value_qualifier	integer	False	Foreign key to the vocabulary_terms table. Qualifies the minimum value result. Examples are '<', '<=', '>', '>=', '~', 'below detection limit', 'above detection limit'. Also to capture the gml:nilReasonTypes ('unknown', 'withheld', 'inapplicable', 'missing', 'template').
maximum_value_qualifier	integer	False	Foreign key to the vocabulary_terms table. Qualifies the maximum value result. Examples are '<', '<=', '>', '>=', '~', 'below detection limit', 'above detection limit'. Also to capture the gml:nilReasonTypes ('unknown', 'withheld', 'inapplicable', 'missing', 'template').

*'asserted\_values' table description:*

Table to store the result of an observation that is a term from a vocabulary.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	The database id
observation_id	integer	True	The foreign key to the Observation that generated the term result
term_id	integer	True	The foreign key to the vocabulary_terms table

*'asserted\_ranges' table description:*

Table to store the result of an observation where the results are a lower term and an upper term

COLUMN NAME	DATATYPE	NOT NULL	DESCRIPTION
id	integer	True	The database id
observation_id	integer	True	The foreign key to the Observation that generated the term ranges
minimum_term_id	integer	True	The foreign key to the vocabulary_terms table for the minimum term (e.g. "poorly drained").
maximum_term_id	integer	True	The foreign key to the vocabulary_terms table for the maximum term (e.g. "well-drained").

*'text\_descriptions' table description:*

Table to store the result of an observation where the result is a free text description.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	database id
text_description	text	False	The free text description result
observation_id	integer	False	The foreign key to the Observation that generated the free text description.

## PROCEDURES

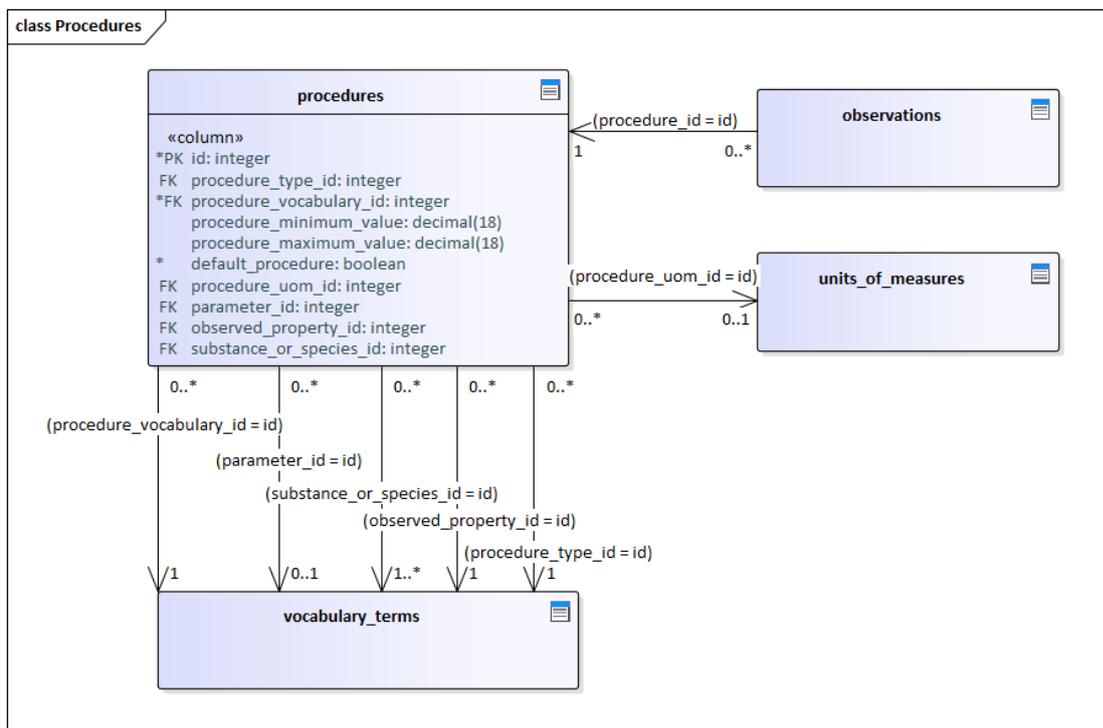


Figure D.4: “procedures” table. Columns of related tables and associated relationships are not shown

### ‘procedures’ table description:

This table records information about the procedures or methods used to record the observation. A challenge with recording procedures is that they sometimes need to be treated as a series of activities or steps, sometimes as a term with additional procedure specific information, and sometimes just as a term from a vocabulary. The former can cater for these variations, but comes with added complexity, particularly during data entry. The current database design caters for the second approach by accommodating additional information, such as whether the procedure is the default method, what result unit of measure the procedure uses, and what the data ranges.

The purpose of the 'procedures' table is to group the properties of units of measure, observed property, minimum and maximum values and substance or species that apply to a certain procedure. From an observation perspective the 'procedures' table only provides a link via the procedure\_id to the vocabulary term that identifies the procedure. The Units of Measure for the actual observation are stored against the result. The observed property for the observation is stored against the observation.

The downside of this is that there is no guarantee that the UoM and observed property associated with the selected procedure will match the UoM stored against the result or the observed property stored against the observation.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	The database id
procedure_type_id	integer	False	A foreign key to the vocabulary_terms table specifying the type of procedure ("measurement", "interpretation", "assertion", "calculation").
procedure_vocabulary_id	integer	True	A foreign key to the vocabulary_terms table that identifies the procedure used (e.g. "pH using 1:5 HCl")
procedure_minimum_value	decimal(18)	False	The minimum allowable value for the procedure if the result is expected to be a measurement.
procedure_maximum_value	decimal(18)	False	The maximum allowable value for the procedure if the result is expected to be a measurement.
default_procedure	boolean	True	Whether the Procedure is the default procedure for this Observed Property
procedure_uom_id	integer	False	A foreign key to the units_of_measures table corresponding to the unit of measure for the procedure
parameter_id	integer	False	A foreign key to the vocabulary_terms table to specify a procedure specific parameter term
observed_property_id	integer	False	A foreign key to the vocabulary_terms table that identifies the observed property being measured by the procedure (e.g. "soil texture")
substance_or_species_id	integer	False	A foreign key to the vocabulary_terms table of the term that identifies the object, substance or species that the property is related to (e.g. "Potassium", "Ammonia").

## SPECIMENS

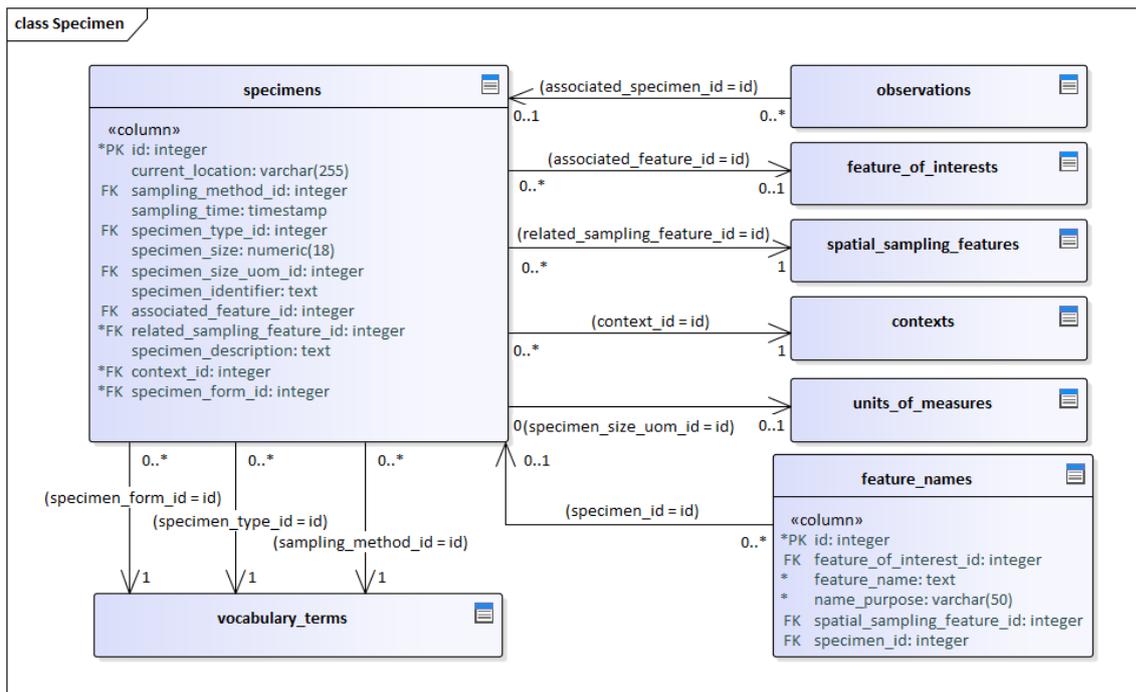


Figure D.5: “specimens” table. Columns of related tables and associated relationships are not shown

### ‘specimens’ table description:

The Specimen table captures information relating to Specimens (= Soil Samples). These may be temporary field specimens that have been discarded or ones that have been archived. It is analogous to O&M SF\_Specimen.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	The database id of the Specimen
current_location	varchar(255)	False	The current storage facility and/or shelf and/or container of the Specimen
sampling_method_id	integer	False	A foreign key to the term in the vocabulary_terms table that identifies the sampling method. In O&M this is an SF_Process, i.e. the sampling procedure consists of a series of steps (to describe for instance combining samples from 10 sites along a transect), not just a single term method. This may be required in future developments.
sampling_time	timestamp	False	The date/time that the specimen was obtained in the field.

specimen_type_id	integer	False	A foreign key to the term in the vocabulary_terms table that identifies the specimen type.
specimen_size	numeric(18)	False	The size of the actual specimen.
specimen_size_uom_id	integer	False	A link to the term in the units_of_measures table that identifies the Unit of Measure of the specimen size.
specimen_identifier	text	False	The persistent identifier (URI) of the Specimen
associated_feature_id	integer	False	A foreign key to the associated features_of_interests table (i.e. the Soil Feature)
related_sampling_feature_id	integer	True	A foreign key to the spatial_sampling_features table to the feature that sampled the Specimen.
specimen_label	varchar(255)	False	The name or label of the Specimen.
specimen_description	text	False	A text description or comment associated with the Specimen.
context_id	integer	True	A foreign key to the context in which the specimen was taken. Usually a job or set of tasks undertaken by a single observer/machine as part of a broader project. Links Job, Project, Client information
specimen_form_id	integer	True	A foreign key to the vocabulary_terms table specifying the basic form of the specimen. e.g. "polished section", "core", "pulp", "solution" Corresponds to SF_Specimen/specimenType

## SAMPLING FEATURES

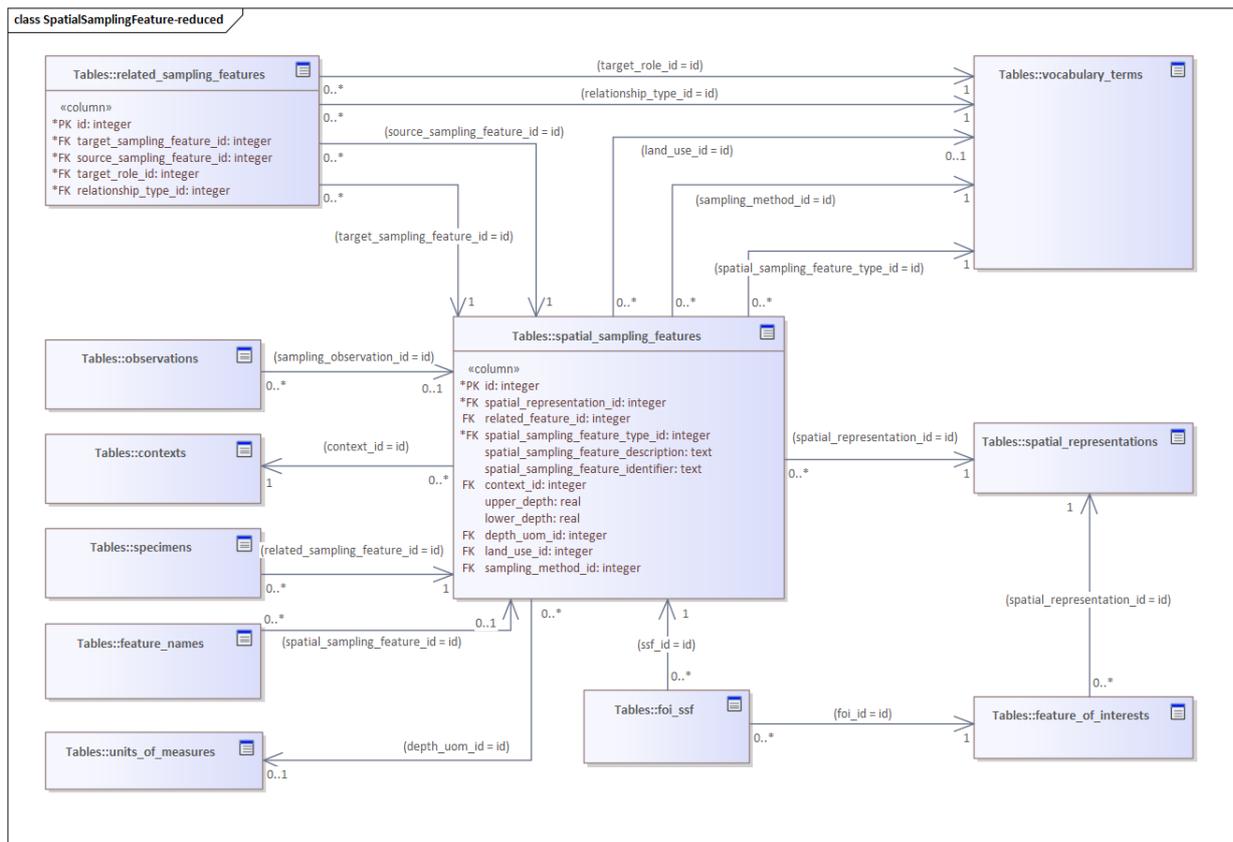


Figure D.6: “spatial\_sampling\_features”, “related\_sampling\_features” and “spatial\_representations” tables.

Columns of related tables and associated relationships are not shown. Together the “spatial\_sampling\_features” and “spatial\_representations” tables describe the 3-dimensional geometry

### ‘spatial\_sampling\_features’ table description:

The spatial\_sampling\_features and spatial\_representations tables store properties analogous to those of the O&M SF\_SpatialSamplingFeature. A spatial sampling feature is used when property observations are made on a geospatial feature. This sampling feature may be in one, two or three spatial dimensions. Properties observed on sampling features may be time-dependent, but the temporal axis does not generally contribute to the classification of sampling feature classes. Sampling feature identity is usually less time-dependent than is the property value.

The SF\_SpatialSamplingFeature is a type of SF\_SamplingFeature. A SF\_SamplingFeature is intended to sample some feature of interest in an application domain. They are artefacts of an observational strategy, and have no significant function outside of their role in the observation process. The physical characteristics of the features themselves are of little interest, except perhaps to the manager of a sampling campaign.

SpatialSamplingFeatures may be "Plot", "Site", "Station", "Paddock", etc. These types are managed through the vocabulary\_terms table. A “station” is essentially an identifiable locality where a sensor system or procedure may be deployed and an observation made. In the context of the observation model, it connotes the “world in the vicinity of the station”, so the observed properties relate to the physical medium at the station, and not to any physical artefact such as a mooring, buoy, benchmark, monument, well, etc. A transient sampling

feature, such as a ships-track or flight-line, may be identified and described, but is unlikely to be revisited exactly.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	The database id
spatial_representation_id	integer	True	A link to the SpatialRepresentation of the SpatialSamplingFeature
related_feature_id	integer	True	The real world feature (e.g. for a SoilFeature the Soil Body, Soil Horizon, Soil Layer, Soil Profile) that the SpatialSamplingFeature samples. (Deprecated 19/10/2022 and replaced with foi_ssf table to allow multiple cardinalities)
spatial_sampling_feature_type_id	integer	True	Identifies the type of SpatialSamplingfeature (e.g. Pit, Plot, Site, Paddock, Borehole, etc). This is a link to a term from a vocabulary in the VocabularyTerm table.
spatial_sampling_feature_description	text	False	A text description or comment related to the SpatialSamplingFeature
context_id	integer	False	A foreign key to the context in which the specimen was taken. Usually a job or set of tasks undertaken by a single observer/machine as part of a broader project. Links Job, Project, Client information
spatial_sampling_feature_identifier	text	False	The persistent identifier (URI) of the SpatialSamplingFeature
upper_depth	real	False	The distance from the surface to the top of the sampling feature
lower_depth	real	False	The distance from the surface to the bottom of the sampling feature

depth_uom_id	integer	False	Identifies the units_of_measures unit that contains the Unit of Measure for the depth values.
land_use_id	integer	False	Identifies the vocabulary term that corresponds to the land use of the spatial sampling feature.
sampling_method_id	integer	False	A link to the term in the VocabularyTerm table that identifies the sampling method. This should be a SF_Process, i.e. the sampling procedure consists of a series of steps (to describe for instance combining samples from 10 sites along a transect), not just a single term method. This may be required in future developments.

#### ‘spatial\_representations’ table description:

Spatial representation (geometry) of the feature of interest. The depth range and associated unit of measure for true 3D information is stored with the associated spatial\_sampling\_feature due to database geometry handling considerations.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	The database id
geometry	geometry	False	FK to the Spatial geometry of the feature
longitude	real	False	
latitude	real	False	
easting	real	False	
northing	real	False	
mapzone	integer	False	
elevation	real	False	
datum	varchar(50)	False	
spatial_id	varchar(50)	False	A non database, non-unique historical id for the GIS geometry.

#### ‘related\_sampling\_features’ table description:

A table that identifies the relationships between the Spatial Sampling Features and the roles each feature plays in the relationship.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	The database id of the relationship
target_sampling_feature_id	integer	True	The target SpatialSamplingFeature
source_sampling_feature_id	integer	True	The source SpatialSamplingFeature
target_role_id	integer	True	A link to a term in the vocabulary_terms table that identifies the role the target SpatialSamplingFeature plays in the relationship.
relationship_type_id	integer	True	A link to a term in the vocabulary_terms table that identifies the type of relationship between the SpatialSamplingFeatures.

#### **'foi\_ssf' table description:**

A table that identifies the relationships between the Spatial Sampling Features and the Features of Interest.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	The database id of the foi_ssf
foi_id	integer	True	Link to the Feature Of Interests table.
ssf_id	integer	True	Link to the Spatial Sampling Features Table.



lower_depth	float	False	Furthest distance from the Earth's surface to the feature.
depth_uom_id	integer	False	Identifies the VocabularyTerm that contains the Unit of Measure for the depth values.

#### **'feature\_relationships' table description:**

A table that identifies the relationships between the Features of Interests and the roles each feature plays in the relationship. Captures relationships between features such as Soil Horizon to Soil Horizon, Soil Horizon to Soil Body, Soil Horizon to Soil Profile and Soil Layer to Soil Body relationships.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	The database id of the relationship
target_role_id	integer	True	Identifies the relationship role the target plays from VocabularyTerm table.
relationship_type_id	integer	True	Identifies the relationship type in VocabularyTerm table
source_id	integer	True	Identifies the source soil feature in SoilFeature table
target_id	integer	True	Identifies the target soil feature in SoilFeature table

#### **'feature\_constituents' table description:**

A table that identifies the constituent parts of a feature and the proportion that each part comprises.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	The database id of the constituent
constituent_feature_id	integer	True	Foreign Key to the constituent feature of interest. Is part of a composite primary key with the feature_of_interest_id - allows the PK index to group record for the same feature together.
feature_of_interests_id	integer	True	Foreign Key to the undivided feature of interest. Is part of a composite primary key with the constituent_interest_id - allows the PK index to group record for the same feature together.

observation_id	integer	True	The observation that provides the proportion ('asserted values', 'measured ranges', 'measured values')
----------------	---------	------	--

## PART 2: CONTEXTS (CONTEXTS, PROJECTS, ORGANISATIONS, COLLECTIONS)

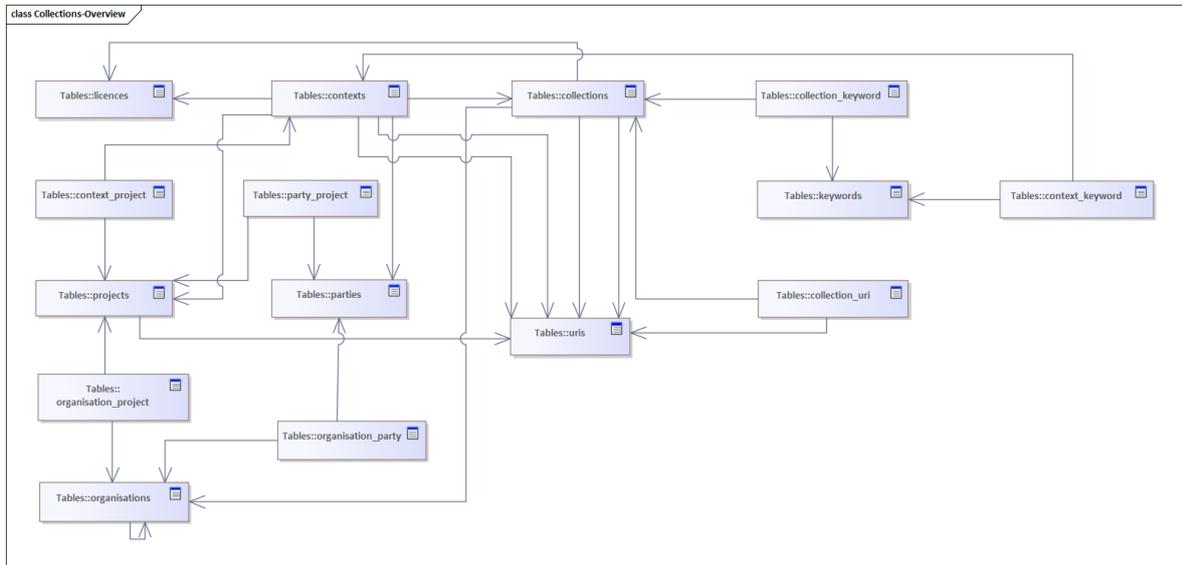


Figure D.8: Database tables that relate to data contexts

The 'Contexts' component stores information about the context of the observational data, its 'metadata'. This includes job information, project names, organisations and their roles, and users (parties) and their roles, along with licensing information, associated keywords or URI links to further information. The term 'parties' has been used to define people who have a specific relationship with the dataset. A separate 'users' table defines members that have access to a VAS login and is a system table not defined here.



obfuscation_type	varchar(255)	False	If not False, specify the type of obfuscation to be applied, e.g. "5km_bbox"
obfuscation_distance_dd	float	False	The distance of obfuscation in decimal degrees
obfuscation_rounding	integer	False	The amount of decimal rounding for location randomisation
public	boolean	False	Is this context intended to be fully publicly available?
state	varchar(50)	False	The Australian state abbreviation most appropriate this dataset
description	text	False	A description of the dataset
originating_source_uri_id	integer	False	Link to the URI containing the original source if relevant
further_information_uri_id	integer	False	Link to the URI containing further information about dataset if relevant
rights	text	False	Any specific rights applicable to the dataset

#### **'projects' table description:**

The Project that the observations and/or sampling were part of. Projects are a collection of activities, jobs or tasks that together have a common goal. Although multiple organisations may be associated with a Project, usually only a single organisation has the role "client".

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	The database id
project_name	varchar(255)	False	The name of the Project
commodity_id	integer	True	The FK to the id in the vocabulary_terms table that corresponds to the commodity that the project is primarily interested in, such as "soil", "groundwater".
short_name	varchar(100)	False	An abbreviated name for the project
description	text	False	A description of the project
further_information_uri_id	integer	False	Link to the URI containing further information about project if relevant.

#### **'parties' table description:**

An individual user's contact details. Multiple users may have multiple roles in a Project.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
-------------	----------	----------	----------

id	integer	True	The database id of the party
first_name	varchar(50)	True	First name of the party
last_name	varchar(50)	True	The parties last name
email	varchar(255)	False	Email address of the party
phone	varchar(20)	False	Phone number of the party
prefix	varchar(10)	False	Prefix or title for the party
user_id	integer	False	Link to user record if relevant (user also has a login to VAS)

### ‘organisations’ table description:

Individual organisation details.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	Database id of the organisation
organisation_name	varchar(255)	True	Name of the organisation
organisation_address	varchar(255)	False	Address of the Organisation
organisation_contact_id	integer	False	Link to the individual contact for the organisation’s information
organisation_web_page	text	False	The Web home page of the organisation
organisation_type_id	integer	False	A FK to the vocabulary_terms table that corresponds to the term that describes the type of organisation, such as “government”, “research”, “grower group”.
organisation_suburb	varchar(100)	False	Town or suburb of the address of the organisation
organisation_state	varchar(50)	False	State or Territory of the address of the organisation
organisation_postcode	varchar(50)	False	Postcode of the address of the organisation. Needs to be able to handle international postcodes
organisation_country	varchar(100)	False	Country of the address of the organisation
parent_organisation_id	integer	False	The id of the parent organisation, if present.

### ‘organisation\_project’ table description:

Multiple organisations may be associated with a project and multiple projects per organisation.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
-------------	----------	----------	----------

id	integer	True	Database id
organisation_id	integer	True	FK to the Organisation Table
project_id	integer	True	FK to the Project table
organisation_role_id	integer	True	FK to the vocabulary_terms table for the term that corresponds to the role the organisation plays in the Project. A project, such as 'Precision Agriculture soil tests' may have multiple organisations such as 'Farm A', 'Farm B' playing the role 'contributor' with the organisation 'Precision Agriculture' playing the role 'client'. In different projects the same organisations may play different roles.

#### **‘organisation\_party’ table description:**

Multiple organisations may be associated with a party and multiple parties per organisation.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	Database id
organisation_id	integer	True	FK to the Organisation Table.
party_id	integer	True	FK to the Party Table.
party_role_id	integer	True	FK to the vocabulary_terms table for the term that corresponds to the role the party plays in the organisation. An organisation, such as 'Precision Agriculture' may have multiple parties such as 'John Smith' playing the role 'contributor' with the party 'Jane Smith' playing the role 'point of contact'. In different organisations the same party may play different roles.

#### **‘context\_project’ table description:**

Multiple projects may be associated with multiple contexts per project.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	Database id
context_id	integer	True	FK to the Contexts Table.
project_id	integer	True	FK to the Project table.

#### **‘party\_project’ table description:**

Multiple parties may be associated with a project and multiple projects per party.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	Database id
party_id	integer	True	FK to the Party Table
project_id	integer	True	FK to the Project table
party_role_id	integer	True	FK to the vocabulary_terms table for the term that corresponds to the role the party plays in the Project. A project, such as 'Precision Agriculture soil tests' may have multiple parties such as 'John Smith', 'Jane Smith' playing the role 'contributor'. In different projects the same party may play different roles.

### ‘collections’ table description:

A collection can be used to group a number of contexts together into a related set of data.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	The database id of the collection
provider_code	varchar	True	Provider code used to restrict access to datasets
name	varchar	True	The name of the collection
short_name	varchar	False	A short version of the collection name if relevant
description	text	False	A description of the collection
originating_source_uri_id	integer	False	Link to the URI containing original source information if relevant
further_information_source_id	integer	False	Link to the URI containing further information about dataset if relevant
rights	text	False	Any specific rights applicable to the collection
licence_id	integer	False	FK to licence if relevant.
organisation_id	integer	True	FK to the organisation that the collection belongs to

### ‘keywords’ table description:

Keywords can be used to store terms that are relevant to the dataset.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	Database id
name	varchar(255)	True	The name of the keyword

vocabulary_term_id	integer	False	FK to the vocabulary term if relevant.
--------------------	---------	-------	--

### **‘licences’ table description:**

Details on the licence applicable to the context.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	Database id
scheme	varchar(255)	True	The licence scheme
uri	text	True	The URI for the licence scheme (e.g. ‘GPL (General Public License)’)

### **‘uris’ table description:**

Table to store any collection or dataset related URI's, this could be an external link to a report or to an organisations website.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	Database id
uri	text	True	The URI link
label	varchar(255)	False	A label for the URI

### **‘collection\_keyword’ table description:**

Multiple keywords may be associated with a collection and multiple collections per keyword.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	Database id
collection_id	integer	True	FK to the Collection Table.
keyword_id	integer	True	FK to the Keyword Table.

### **‘collection\_uri’ table description:**

Multiple URIs may be associated with a collection and multiple collections per URI.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	Database id
collection_id	integer	True	FK to the Collection Table.
uri_id	integer	True	FK to the URI Table.

### **‘context\_keyword’ table description:**

Multiple keywords may be associated with multiple contexts per keyword.



### ‘vocabulary\_terms’ table description:

A table to store required terms. These may be terms from traditional look-up tables such as terms associated with 'laboratory methods' (such as "pH using 1.5 HCl"), or with 'drainage' (such as "well-drained", "poorly drained"). They may also be the terms that are traditionally the look-up table names and appear as database columns, such as "laboratory method", "pH", "P concentration" etc.

Importantly, in addition to a set of terms the table indicates what schema the term comes from (what the look-up table name would be that the term came from if from a look-up table), a description, and a persistent identifier to an external resource that contains more information (such as images, other relationships, conversion values, alternative labels or languages etc.).

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	The database column id
vocabulary_collection	varchar(255)	False	The label of the collection that the vocabulary member belongs to (e.g. "drainage", "laboratory method", "observed property")
collection_identifier	text	False	Unique persistent id (URI) of the collection the vocabulary term belongs to
member_label	varchar(255)	False	Preferred label of the vocabulary member (e.g. "poorly drained"). This is what is normally stored in the look-up table.
member_identifier	text	True	Unique external persistent identifier (URI) of the vocabulary term
member_notation	varchar(50)	False	Code/abbreviation for the vocabulary member
member_description	text	False	Description of the vocabulary term
broader_term_id	integer	False	The id of the broader term in a vocabulary hierarchy, if present
member_alias	varchar(255)	False	An alternative member label
date_modified	date	False	Recorded database date of modification of vocabulary term
date_deprecated	date	False	Date this record was deprecated (if applicable)
replace_with_member_identifier	text	False	Member identifier this term has been replaced by (if applicable)

## UNITS OF MEASURE

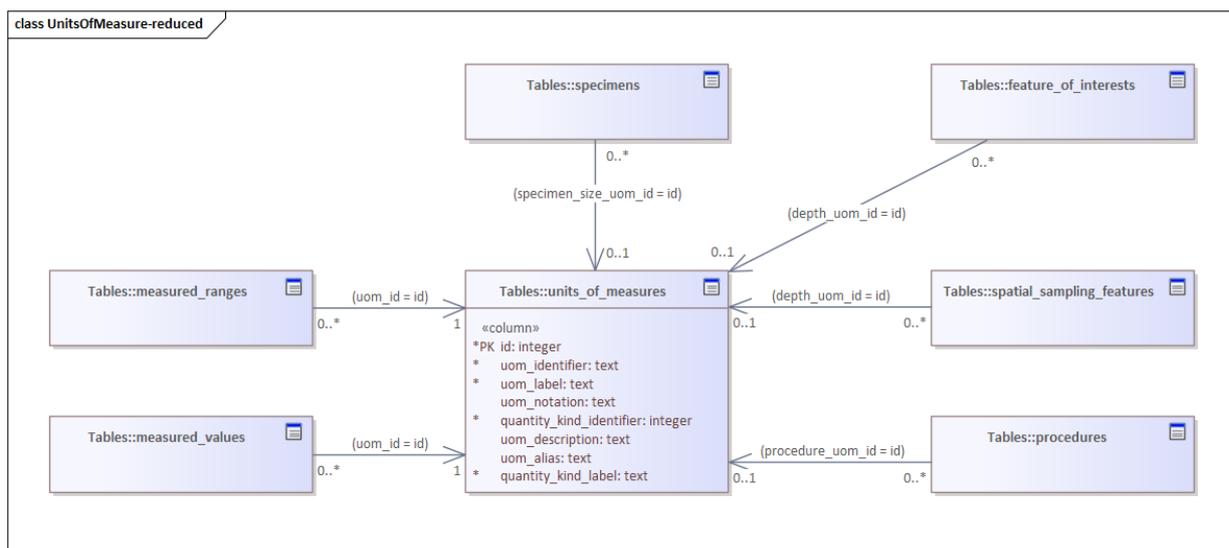


Figure D.11: The ‘units\_of\_measures’ table and the associated tables that use it. Columns of related tables and associated relationships are not shown.

### ‘units\_of\_measures’ table description:

The units of measure is based on the Quantities, Units, Dimensions and Types Ontology (QUDT). It allows specifying the label (e.g. “metre”), its abbreviation (e.g. “m”), alternative labels (e.g. “meter”), and the kind of measure the unit of measure relates to (e.g. “Length”). The table also caters for the identifiers (URIs) for the unit of measure and its quantity kind.

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	Database column id
uom_identifier	text	True	URI of the unit of measure term. From Linked Data Registry = rdf:about
uom_label	text	True	The text label of the UoM term. From Linked Data Registry = rdfs:label
uom_notation	text	False	The abbreviation for the UoM term. From Linked Data Registry = qudt:abbreviation or qudt:symbol or skos:notation

quantity_kind_identifier	integer	True	<p>The URI for the type of quantity that the UoM is (e.g. 'metre' is a 'Length' 'QuantityKind')  From Linked Data Registry = <a href="https://qudt.org/schema/qudt/QuantityKind">https://qudt.org/schema/qudt/QuantityKind</a>  Because QuantityKinds are hierarchical a Unit of Measure may have multiple QuantityKinds:  e.g. '<b>Atoms per Litre</b>' is a unit of measure for QuantityKinds of '<b>Concentration</b>' as well as its specialization '<b>Amount of Substance Per Unit Volume</b>'.  The database design only allows for a Unit of Measure to be assigned to a single QuantityKind, preferably the most specialized.</p>
uom_description	text	False	<p>Text to describe the UoM.  From Linked Data Registry = qudt:description or dct:description or skos:definition</p>
uom_alias	text	False	<p>An alternative label or spelling of the UoM term (e.g. 'metre' rather than 'meter')  From Linked Data Registry = skos:altLabel</p>
quantity_kind_label	text	True	<p>The label of the QuantityKind (e.g. "Length").  From Linked Data Registry = <a href="https://www.w3.org/2000/01/rdf-schema#label">https://www.w3.org/2000/01/rdf-schema#label</a></p>

## FEATURE NAMES

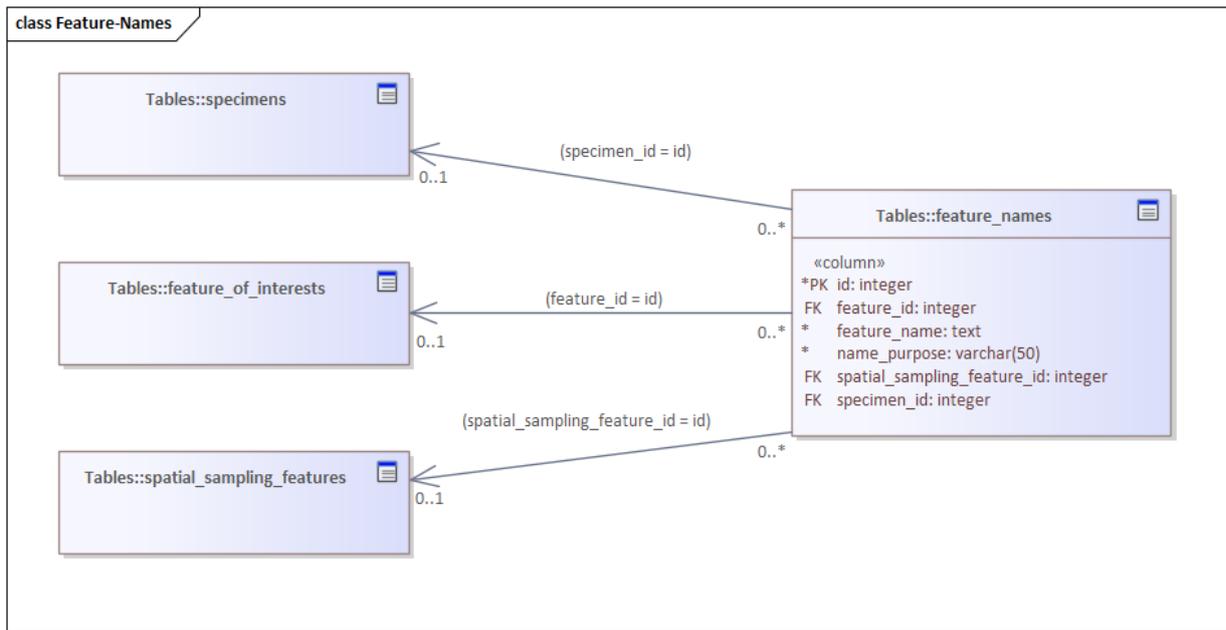


Figure D.12: The `feature_names`' table allows specifying multiple names (identity) to individual features, whether they are features of interest, specimens or spatial sampling features. Columns of related tables and associated relationships are not shown.

### 'feature\_names' table description:

Allows for distinguishing multiple names of any single feature

COLUMN NAME	DATATYPE	NOT NULL	COMMENTS
id	integer	True	The database column id
feature_id	integer	False	The identity of the feature of interest that the name applies to
feature_name	text	True	The name of the feature (e.g. "Lismore Grey Clays" and/or "33", or "Burt's Paddock", or "Specimen Barcode 222222", "Borehole M124" etc.)
name_purpose	varchar(50)	True	Allows distinguishing the reason for multiple names
spatial_sampling_feature_id	integer	False	The identity of the spatial sampling feature that the name applies to
specimen_id	integer	False	The identity of the specimen that the name applies to

## APPENDIX E VAS SYSTEM API

As shown in Figure 4.1 of this report, soil data stored in the VAS aggregator and other data consumed from external sources is published from the VAS system via an application programming interface (API). The VAS API strives to implement the best practice set out by the World Wide Web Consortium (W3C) for publishing data on the web.

- Data on the Web Best Practices <https://www.w3.org/TR/dwbp/>
- Best Practices for Publishing Linked Data <https://www.w3.org/TR/ld-bp/>
- Spatial Data on the Web Best Practices <https://www.w3.org/TR/sdw-bp/>

To achieve these goals, the project chose to implement a REST-full API, using persistent identifiers and data primarily delivered in JSON-LD format. These payloads were in turn, based on a set of best practices for Building JSON-LD APIs <https://json-ld.org/spec/latest/json-ld-api-best-practices/>

JSON-LD is an ideal data format for programming environments familiar with the pervasive JSON format, being fully backwards compatible with JSON libraries and parsers. It adds interoperability by providing the ability to bind JSON elements, attributes and values to formal data structures, standards and vocabularies. To this end, the VAS API has attempted to reuse existing international standards, vocabularies and common patterns when defining its data structures.

In particular, the Endpoints described here have been based on the Sensor, Observation, Sample, and Actuator (SOSA) ontology <https://www.w3.org/TR/vocab-ssn/> but should not be considered fully compliant with that specification. Other standards and vocabularies utilised include:

- Schema.org [<https://schema.org/>]
- Quantities, Units, Dimensions, and Types (QUDT) [<https://www.qudt.org/>]
- Provenance ontology (PROV) [<https://www.w3.org/TR/prov-dm/>]
- Simple Knowledge Organization System (SKOS) [<https://www.w3.org/TR/skos-reference/>]
- GeoSPARQL [<https://www.ogc.org/standard/geosparql/>]
- ANZSoilML & ANSIS controlled vocabularies for soil data

## API SPECIFICATION

An OpenAPI specification document describing the VAS API is available at <https://app.swaggerhub.com/apis-docs/FedUniCeRDI/vas-soils-api/1.0#> but an abbreviated version is presented here as an introduction.

Base URI: <https://id.cerdi.edu.au>

Main Endpoints:

Name	URI
<b>Feature</b>	<code>/ {provider} /data/sosa/feature</code> <code>/ {provider} /data/sosa/feature/ {id}</code>  Examples: <a href="#">/ccma/data/sosa/feature</a> <a href="#">/ccma/data/sosa/feature/ccma.soil.feature.7174</a>
<b>Observation</b>	<code>/ {provider} /data/sosa/observation</code> <code>/ {provider} /data/sosa/observation/ {id}</code>  Examples: <a href="#">/ccma/data/sosa/observation</a> <a href="#">/ccma/data/sosa/observation/ccma.soil.observation.10840</a>
<b>Dataset</b>	<code>/ {provider} /data/schema/dataset</code> <code>/ {provider} /data/schema/dataset/ {contextJobNumber}</code>  Examples: <a href="#">/ccma/data/schema/dataset</a> <a href="#">/ccma/data/schema/dataset/000001</a>
<b>Observed Properties</b>	<code>/query/observedproperties</code>
<b>Procedures</b>	<code>/query/procedures</code>
<b>PropertiesProcedures</b>	<code>/query/propertyprocedures</code> <code>/query/propertyprocedures/ {provider}</code>

## Example response for a single observation

```
◦ {
◦   "@graph": {
◦     "@id":
"https://id.cerdi.edu.au/ccma/data/sosa/observation/ccma.soil.observation.10840",
◦     "@type": "Observation",
◦     "hasFeatureOfInterest": {
◦       "@id":
"https://id.cerdi.edu.au/ccma/data/sosa/sample/ccma.soil.sample.1659.1659"
◦     },
◦     "hasResult": {
◦       "numericValue": 5.2,
◦       "unit": {
◦         "@id": "https://qudt.org/vocab/unit/PH",
◦         "abbreviation": "pH",
◦         "label": "Acidity"
◦       }
◦     },
◦     "hasSimpleResult": 5.2,
◦     "hasUltimateFeatureOfInterest": {
◦       "@id": "https://id.cerdi.edu.au/ccma/data/sosa/feature/ccma.soil.feature.7174"
◦     },
◦     "inDataset":
"https://id.cerdi.edu.au/ccma/data/schema/dataset/ccma.schema.dataset.000001",
◦     "name": "ccma.soil.observation.10840",
◦     "observedProperty": {
◦       "@id": "http://environment.data.gov.au/def/property/PH",
◦       "@type": "Concept",
◦       "description": "negative logarithm of hydrogen ion concentration in ph units",
◦       "label": "pH",
◦       "name": "pH",
◦       "notation": null
◦     },
◦     "phenomenonTime": {
◦       "@type": "http://www.w3.org/2001/XMLSchema#dateTimeStamp",
◦       "@value": "2018-05-15T00:00:00"
◦     },
◦     "provider": "ccma",
◦     "resultTime": {
◦       "@type": "http://www.w3.org/2001/XMLSchema#dateTimeStamp",
◦       "@value": "2018-05-15T00:00:00"
◦     },
◦     "soDepth": {
◦       "soMaximumValue": {
◦         "@value": 20
◦       },
◦       "soMinimumValue": {
◦         "@value": 10
◦       },
◦       "unit": {
◦         "@id": "https://qudt.org/vocab/unit/CentiM",
◦         "abbreviation": "cm"
◦       }
◦     },
◦     "ufoi_id": 7174,
◦     "usedProcedure": {
◦       "@id": "http://anzsoil.org/def/au/scm/4B4",
◦       "@type": "Concept",
◦       "description": "pH of 1:5 soil/0.01 M calcium chloride extract - following
method 4A1 (with stirring during measurement)",
◦       "label": "pH of 1:5 soil/0.01 M calcium chloride extract - 4B4",
◦       "name": "pH of 1:5 soil/0.01 M calcium chloride extract - following method 4A1
(with stirring during measurement)",
◦       "notation": "4B4"
◦     }
◦   }
◦ }
```

## Other example requests

In the case of the Corangamite Catchment Management Authority (CCMA) dataset, a suite of soil chemistry analyses - 'observations' - is available for each of the features of type *SoilLayer*. These observations can be retrieved using the query syntax below, i.e. return all the observations that relate to these two features. In this case these two features represent 0-10cm (soil layer depth) at a specific location in 2014 and 0-10cm at the same location in 2018.

```
https://id.cerdi.edu.au/ccma/data/sosa/observation/?or=(
  hasUltimateFeatureOfInterest-->"@id".eq.https://id.cerdi.../feature/ccma.soil.feature.6868,
  hasUltimateFeatureOfInterest-->"@id".eq.https://id.cerdi.../feature/ccma.soil.feature.14
)
```

To go the other way (from observation up to feature) is also possible. For example, return all the observations that have Observed Property = Bulk Density.

```
https://id.cerdi.edu.au/ccma/data/sosa/observation/?
observedProperty-->"@id".eq.http://...cerdi.edu.au/property/bulk_density_soil
```

The response to the request above is a standard SOSA pattern, where the value URI of `hasUltimateFeatureOfInterest` for each observation would need to be resolved to get more information about the feature (including spatial information).

There is a convenience version of the observation end-point which includes the *f\_geom* in each observation for use in spatial visualisation by property value.

```
https://id.cerdi.edu.au/ccma/data/sosa/observationsearch/?
observedProperty-->"@id".=eq.http://placeholder.cerdi.edu.au/property/bulk_density_soil
```

## Observations matching specific observedProperty and usedProcedure

```
https://id.cerdi.edu.au/ccma/data/sosa/observation/?or=(
  observedProperty-->"@id".eq. http://environment.../def/property/phosphorus_concentration,
  usedProcedure-->"@id".eq.http://www.anzsoil.org/def/au/scma/9B)

```

As the API is based on PostgREST middleware, it should support all the features of a typical ODATA [https://www.odata.org/documentation/] API. This includes attribute selection, pagination, ordering etc. Most of the example queries documented at the URL below should apply.

[https://postgrest.org/en/stable/references/api/tables\\_views.html](https://postgrest.org/en/stable/references/api/tables_views.html)

For example, the observations end point example for bulk density with a limit and attribute selection added.

```
https://id.cerdi.edu.au/ccma/data/sosa/observationsearch/?observedProperty
-->"@id".=eq.http://placeholder.cerdi.edu.au/property/bulk_density_soil
&limit=20
&select=provider,f_geom,hasResult
```

## AUTHENTICATION

Since the majority of datasets are non-public at the present time, access is provided to authorised datasets via an API key. An authentication header with a valid API Key is required

for many of the REST endpoints when requesting non-public data. The Authorization field in the HTTP header is used to pass the API key (or token).

# APPENDIX F SOIL SENSOR DATA: ARCHITECTURE DEVELOPMENT

VAS Phase 2 aspired to include data collected by soil sensors available via the VAS portal. Since most of the participating farmer groups in the VAS project have soil moisture and temperature probes installed, methods were explored to include these data in the soil data federation. Probe data are usually telemetered to a digital repository run by a third party and then graphically displayed on the farmer group website, via a member login.

An initial use-case to include Southern Farming Systems (SFS) probe data in the VAS portal was undertaken. The methods built on a previous project undertaken by CeRDI for SFS, which established a portal known as Probetrax<sup>11</sup> in which the public can view the current plant available water (PAW) as a percentage value (%) at each probe site as a coloured dot. In the Probetrax portal, SFS members hosting a probe can login to the portal and view their time-series soil moisture and temperature down profile as graphs, together with the rainfall, if a rain gauge was fitted.

Displaying the SFS soil moisture probes in the VAS portal was achieved (Figure F.1). Additional methods were implemented include data from Agriculture Victoria's soil moisture probe network in the VAS portal (Figure F.2).

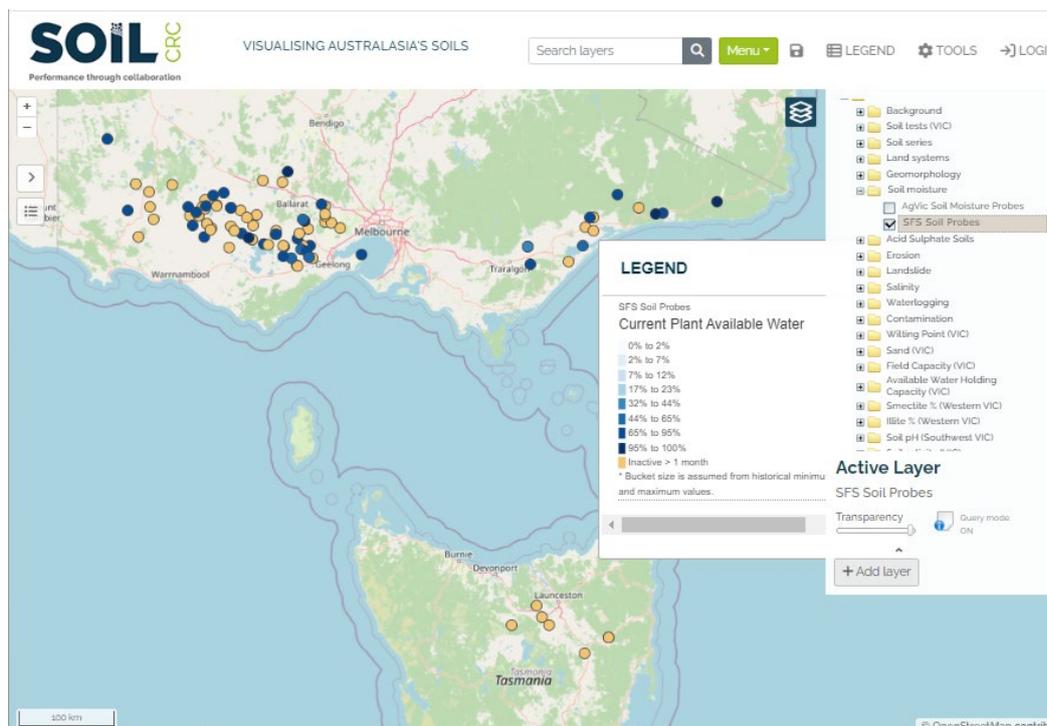


Figure F.1 SFS soil moisture probes displayed in the VAS portal (8/8/2024).

<sup>11</sup> <https://probetrax.sfs.org.au/>

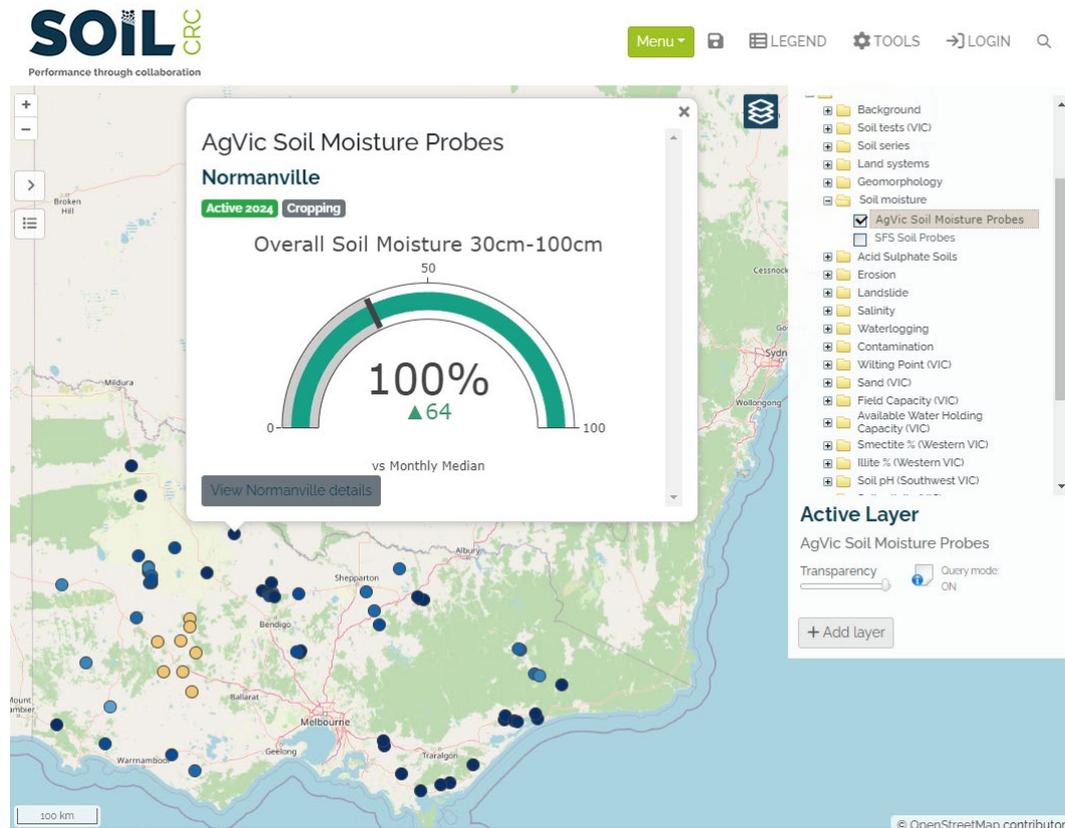


Figure F.2 Agriculture Victoria’s soil moisture probes in the VAS portal (8/8/2024).

Research effort has continued throughout 2024 on integrating soil probes from more providers across Australia within VAS. A key objective is to be able to provide “on demand” access for a user to request “all the data”, i.e. include any sensor value that is reported, not just soil moisture. This request springs from researchers, especially PhD students, who are undertaking research projects that require access to time-series soil moisture and/or temperature throughout a soil profile across a wide area. Among the many challenges this creates is the fact that these data for one provider (SFS) amounts to well over 600 million observations.

A new VAS soil moisture probe interface was developed in early 2024 to allow access to “Verified” (Soil Moisture) sensors and “Unverified” sensors (All other sensors), as well as selecting an appropriate data frequency (Figure F.3). However this interface upgrade has remained on hold pending other VAS upgrades and a resolution to the probe data issues.

After lengthy development and testing, the technical team made the decision to abandon the existing SensorThings based Frost service, initially deployed for the VAS architecture, as it has proven to be a major barrier in this objective. While being OGC compliant and achieving some FAIR data objectives, the lack of aggregate data queries in the standard makes the end use very developer-unfriendly, requiring thousands of API calls to build the dataset.

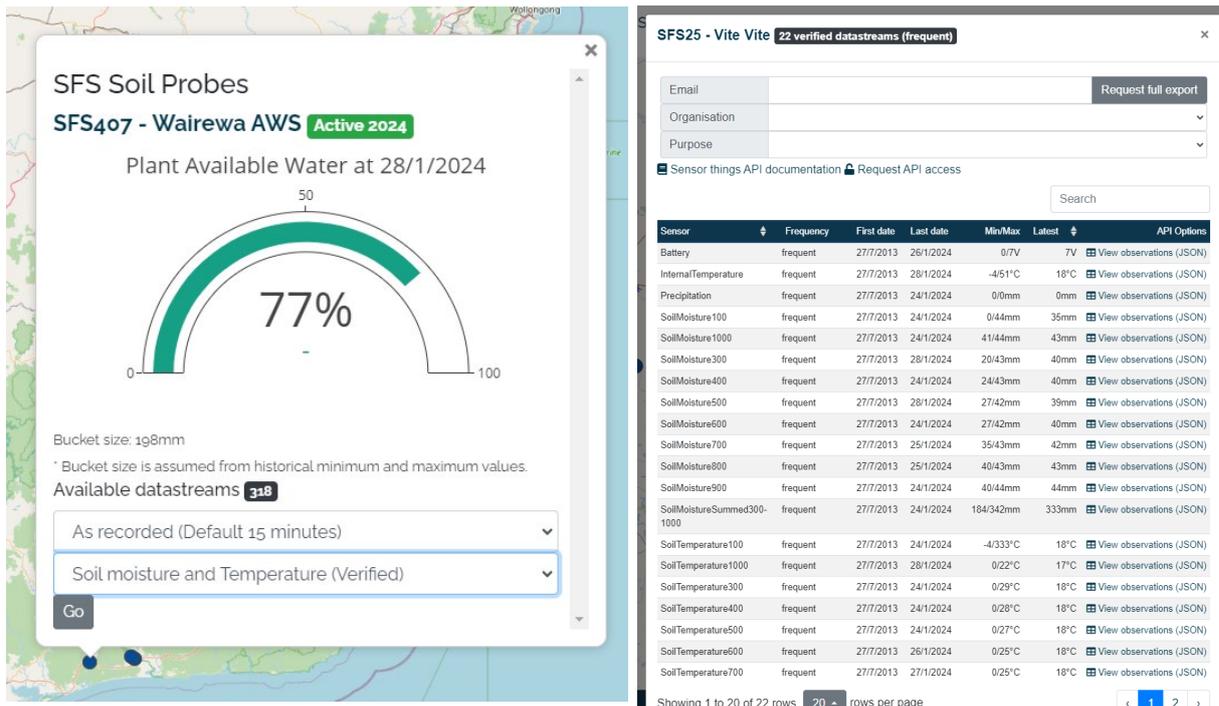


Figure F.3 An example of the interface using verified data.

The background processing for importing multi-year backlogs of data via SensorThings Frost Server is very lengthy (Weeks/Days), and much difficulty was encountered in verifying the imports had completed successfully, although it is acknowledged that some of these difficulties may be particular to CeRDI's combination of infrastructure and SensorThings, and not specific to the Frost Server itself. Nevertheless, the process introduced a significant risk, since if a central Frost SensorThings service was used for all providers, and a data issue was uncovered, it may take months to wipe and re-ingest the data for all providers. On the flip side, a dedicated Frost SensorThings service for each provider bought other issues when it comes to the goal of providing all the data together, as well as for maintainability.

A secondary goal in the VAS soil data federation is to have a central system ready to connect other soil probe data providers across Australasia. This development has been hampered by the fact that no project participants have yet provided API access to their probe data in 2023/2024. At present only API connections from SFS data stream provider, Adcon, and Mackillop Farm Management Group (MFMG) data stream provider, Wildeye, have been tested. Given the number of issues uncovered from these two experiments, it is expected that many more issues will be uncovered as more providers come onboard.

As a result of the above research, a decision was made to develop a new technical architecture and system to overcome these challenges.

## CERDI MULTI PROVIDER PROBE COLLECTOR

In 2024 development has started on a new probe collector platform for CeRDI's purposes. This collector

- Replaces existing standalone import scripts for Wildeye and Adcon data stream providers.
- Replaces custom frameworks and legacy code with a modern Laravel framework.
- Utilises TimescaleDB for performance and efficient live data aggregates.
- Provide a straightforward token-based API service to request data instead of a SensorThings datastream service.

### API Routes

The API routes in development are as follows:

#### Create Sanctum Token

*User Required*

POST /api/login

**NOTE:** Username and login must be created via Artisan Tinker until a registration process exists.

#### Get Device List (All providers)

*Token Required*

GET /api/devices/list/devices/list

#### Get Sensor List

*Token Required*

GET /{provider}/sensors/{device}/list

#### Get Observed Properties

*Token Required*

GET /observed-properties/list

#### Get Units of Measure

*Token Required*

GET /unit-of-measure/list

#### Get Observations

*Token Required*

Paginated at 5000 per page

Optional parameters:

start\_date YYYY-MM-DD 00:00:00

end\_date YYYY-MM-DD 00:00:00

## By Sensor

### *Token Required*

GET `/{provider}/observations/sensor/{sensor}/list`

## By Device

### *Token Required*

GET `/{provider}/observations/device/{device}/list`

## By Frequency

### *Token Required*

`/{provider}/observations/frequency/hourly/list` `/{provider}/observations/frequency/daily/list`  
`/{provider}/observations/frequency/weekly/list`  
`/{provider}/observations/frequency/monthly/list`

## Export Observations

### Generate CSV Request

#### *Token Required*

`/{provider}/observations/export`

### Download generated CSV\*

#### *Token Required*

`/{provider}/observations/export/{download_number}`

\*Maximum of 500k observations returned

## Progress and research challenges

The development of this API has raised many research challenges with soil probe data not previously considered by the VAS technical team. Aside from VAS, the Multi Provider Probe Collector service is required for other significant CeRDI research projects such as a Future Drought Fund Small Farm Dams Project, and the SFS My Farm Dashboard project. Some of the same issues have been encountered working on probe data for these projects, and the learnings and solutions are transferrable VAS. These issues include:

### ***Inactive probe issues***

- There are a growing number of probes from some project participants that are no longer operational.
- The upcoming impact of the Telstra 3G Mobile Service shutdown will also make this issue larger unless replacement probe SIM cards are funded.
- The number of inactive probes can result in misleading maps. That is, the map may display what appears to be the current state of soil moisture in a region,

but inactive probes may be showing the last operational value (i.e. stale data).

### ***Ingesting data without sufficient context***

The original SensorThings API was mapped carefully to observed properties and manual descriptions of the soil moisture probe sensors so that the data made sense to the end user. Only sensors with an established definition and vocabulary were imported into the system. However, when ingesting all API data without such a vetting process, the quality of information is mixed and context is lacking, issues include:

- Inconsistent probe naming conventions (Both across providers and within the same provider)
- Misleading sensor configurations, for example:
  - “10cm soil moisture” is actually “20cm soil moisture” at a certain location, as the 10cm probe was moved to a different location, and a 10-90cm probe was installed 20cm deep.
  - An extension called “NW 1” is actually “Rainfall since 9am at NW1”, but the name is only in the parent node.
  - Sensors close together that have different soil and crop types (Lucerne and Veldt)
  - A user preference to display one of those probes, and not the other.
- Exposing data that should not be exposed.
  - CeRDI’s API access can be more elevated than required. Without careful checking, partners not participating in the VAS project may be exposed.
  - There have been instances where probe data was included via the API, but this data was never intended for the project partner, since the probe provider was merely testing a different brand of sensor in parallel, which was not made available via the partner’s dashboard.
- Unannounced changes in naming convention or hierarchy
  - This can mean probe data ceases to import, as the corresponding identifier is no longer valid. This potentially allows for CeRDI to ingest data it is not permitted to use if the identifier is changed to one thought to be permitted.
- New probes can be included without warning.
  - In many cases the system would work as intended when a new probe was installed.
  - In one case, at the request of the project partner, new probes were included in the API feed from a local council by the data stream provider. This automatically and prematurely fed into the live site before any notification had been received. This occurred before permission to use the data was secured, and before any review of the incoming data.

- Mis-representing sensor data.
  - The VAS project partners typically access their data via a dashboard configured by their data stream provider (e.g. Adcon, Wildeye, or Outpost Central). This dashboard is carefully curated by the data stream provider and only contains the information the provider thinks is useful and accurate for the project partner. It may not include all sensors, and may display particular sensors in a different way.
  - The partner dashboard typically includes clear descriptions and units, but this information is not supplied via the API.
  - The data stream provider may have their own bespoke logic around how summed soil moisture is calculated at different locations. This logic is not communicated via the API, and will automatically attempt to sum all sensors from 300mm to 1000mm as summed soil moisture.
  - As a result, access via the API has no context of what has been curated by the data stream provider and can present all the data in a “dumb” manner. While technically functioning correctly, the values presented to the user in the VAS system can be very different to those they viewed in their own dashboard, eroding trust in the data sharing.

# APPENDIX G VAS DATA TRANSFER AGREEMENT

## Agreement Details

<b>University</b>	<b>Federation University Australia</b> ABN 51 818 692 256 of University Drive, Mt Helen, Victoria 3350, Australia	
<b>Recipient Organisation</b>	[Insert name of Recipient Organisation] ABN [Insert ABN details] of [Insert address]	
<b>Key Researcher</b>	[Insert name and title] [Insert address] [Insert phone number] [Insert email]	
<b>Commencement Date</b>		
<b>Completion Date</b>		
<b>Dataset description</b>	Overview:	<i>Delete this note once overview included: Insert a description of the data that is being licenced under this Agreement. The description should be detailed enough that individuals a few years down the track will know exactly what has been licenced.</i>
	Owner:	
	Custodian:	
	Project:	
	Size:	
	Date:	
<b>Purpose</b>	The Data is to be used only for research purposes. Please indicate from the options below how you intend to use the Data:	
	<input type="checkbox"/> Academic research. <i>Please specify:</i>	
	<input type="checkbox"/> Government research. <i>Please specify:</i>	
	<input type="checkbox"/> Industry research. <i>Please specify:</i>	
	<input type="checkbox"/> Individual research. <i>Please specify:</i>	
	<input type="checkbox"/> Other. <i>Please specify:</i>	
<b>Location</b>	[Insert the location where the Recipient Organisation and Key Researcher will use and stored]	

<b>Fee</b>	Nil
<b>University address for notices</b>	Visualising Australasia's Soils Data Administration Centre for eResearch and Digital Innovation P.O. Box 691, Ballarat, Vic 3353 vas.support@federation.edu.au
<b>Recipient Organisation address for notices</b>	[Insert contact person name] [Insert organisation name] [Insert postal address] [Insert telephone number] [Insert email address]

# Execution

## Executed as an Agreement

Executed by an authorised person of **Federation University Australia** ABN 51 818 692 256 in the presence of:

.....  
Signature of authorised person

.....  
Signature of witness

.....  
Name of authorised person

.....  
Name of witness

Date: .....

Executed by an authorised person of the [Data owner or custodian] in the presence of:

.....  
Signature of authorised person

.....  
Signature of witness

.....  
Name of authorised person

.....  
Name of witness

Date: .....

Executed by an authorised person of the [Insert organisation name] ABN [Insert ABN] in the presence of:

.....  
Signature of authorised person

.....  
Signature of witness

.....  
Name of authorised person

.....  
Name of witness

Date: .....

Read and acknowledged by

.....  
Signature of Key Researcher

.....  
Name of Key Researcher

.....  
Date:

## Background

- A. The Recipient Organisation has requested the supply of the Data for the Purpose.
- B. The University has agreed to supply the Recipient Organisation with Data on the terms set out in this Agreement.

## Operative Part

### 1. Provision of Data

---

#### Ownership of Data

The Recipient Organisation acknowledges and agrees that the Data:

- (a) have been developed, acquired or curated by the University; and
- (b) are of considerable value to research and potential commercial applications.

#### Use of Data

The Recipient Organisation agrees that it will not, without the prior written consent of the University:

- (c) sell, loan or otherwise provide or transfer any Data to any third party;
- (d) use the Data or Derivatives for any purpose other than the Purpose; or
- (e) use or store the Data or Derivatives in any place other than at the Location and under the Key Researcher's direct supervision,

and the Recipient Organisation agrees to promptly refer to the University any request for access to Data from a third party.

#### Title

All rights, title and interest in the Data and Derivatives (including all Intellectual Property rights subsisting in them) remain with the Data Owner and University at all times.

#### Compliance

- (f) The Recipient Organisation must ensure Data and Derivatives are stored, transported, used and disposed of in accordance with all relevant Laws, codes of practice, ethical principles and any reasonable direction of the University.
- (g) The Recipient Organisation must not use the Data or Derivatives in any research trials, or for diagnostic purposes, involving human subjects without the prior written consent of the University.

#### Not exclusive

- (h) Any rights granted to the Recipient Organisation to use the Data under this Agreement are non-exclusive.
- (i) The University may make the Data available to third parties without consultation with, or consent from, the Recipient Organisation.

### 2. No warranty

---

- (a) To the extent permitted by Law, all express and implied warranties and conditions relating to the Data are excluded or, where such an exclusion is prohibited by Law, liability under any such implied

conditions and warranties is limited to the extent permitted by Law.

- (b) The University gives no express warranty that:
  - (i) the use of the Data or Derivatives will not infringe the Intellectual Property or other rights of any third party;
  - (ii) the Data are fit for purpose; or
  - (iii) the Data are of a particular quality or possess particular characteristics.

### 3. Acknowledgement

---

The Recipient Organisation acknowledges that:

- (a) the Data is experimental in nature and that the speculative nature of scientific research is such that it is unreasonable to expect the University to give any assurances as to the performance of the Data; and
- (b) the University's awareness of the Purpose of the project and the role of the Data does not constitute an endorsement of the project or advice on the project.

### 4. Provision of Information

---

#### Reports

- (a) The Recipient Organisation must keep and maintain accurate and reasonably detailed records in connection with use, storage, transport and disposal of the Data.
- (b) The Recipient Organisation agrees to provide the University with a report setting out such things as the Results of the research, Derivatives and Data IP produced, upon request from the University.

### 5. Intellectual Property

---

#### Ownership

- (a) All Data and IP subsisting in or in relation to them are in the custodianship of the University.
- (b) The University grants the Recipient Organisation a non-exclusive, non-transferable, fee-free licence to use the Data for the Purpose only.
- (c) All intellectual Property created by or on behalf of the Recipient Organisation as a result of using the Data other than in accordance with this Agreement will vest in and, by this Agreement, is assigned to, the University upon its creation.
- (d) The Recipient Organisation must:
  - (i) sign all documents and do all things necessary to perfect and record the University's ownership rights under this clause 5; and
  - (ii) not directly or indirectly engage in any conduct that might impair or prevent the protection of Data IP or challenge the University's ownership of it.

#### Commercialisation

- (e) If the Recipient Organisation wishes to commercialise or have commercialised any Results or Data IP, or otherwise deal in the Data or Derivatives for any commercial purpose, it must

first enter into an appropriate licence agreement with the University and data owner.

- (f) The University agrees to negotiate in good faith a licence agreement on terms acceptable to all parties.
- (g) Neither party is under any obligation to enter into a licence agreement on any specific terms, or at all.

## 6. Confidential Information and Privacy

---

### General obligations

The Recipient Organisation must:

- (a) hold any Confidential Information in strict confidence and not disclose, or cause or permit the disclosure of, the Confidential Information, except as permitted under this Agreement or with the prior written consent of the University;
- (b) limit access to the Key Researcher;
- (c) keep any Confidential Information secure and protected from any use, disclosure or access which is inconsistent with this Agreement;
- (d) ensure all Key Researcher with access to any Confidential Information agree and be bound to keep the Confidential Information confidential and not to use the Confidential Information other than for the Purpose; and
- (e) promptly notify the University if the Recipient Organisation suspects, or becomes aware of, any unauthorised use, storage, copying or disclosure of any Confidential Information.

### Use of Confidential Information

The Recipient Organisation must:

- (f) only use Confidential Information for the Purpose;
- (g) not perform any analysis, synthesis, reformation, decomposition, disassembly or reverse engineering of any Confidential Information without the written consent of the University; and
- (h) not make use of any Confidential Information to the commercial, financial or competitive disadvantage of the University.

### Return of Confidential Information

- (i) If requested by the University at any time, and upon the expiry or termination of this Agreement, the Recipient Organisation must immediately return to the University, or destroy, delete and erase as the University directs, all original documents and copies that are or contain Confidential Information or that reproduce, are based on, utilise or relate to Confidential Information.
- (j) The return, destruction, deletion or erasure of Confidential Information does not relieve the Recipient Organisation from its other obligations under this Agreement.

### Privacy

- (k) Each Party must comply with its obligations under all applicable laws in relation to the collection, storage, use and disclosure of any Personal Information or Health Information (as defined in any Applicable Privacy Laws) which it obtains during the conduct of the Project or to which it becomes privy as a result of this Agreement.

## 7. Publication

---

### Consent to publication

The Recipient Organisation will not use the name or logo of the University in any publication without the University's consent, which will not be unreasonably withheld.

### Procedure for publication

- (a) Any restriction on a publication (including a student thesis) will be limited to the restrictions reasonably necessary to protect the University's Confidential Information, Data IP and Derivatives.
- (b) The Recipient Organisation must acknowledge the contribution of the University in any publication.

## 8. Insurance

---

The Recipient Organisation warrants that it has, and will maintain at its cost, appropriate insurance coverage in respect of the loss of, or damage to, the Data, on terms consistent with prudent risk management.

## 9. Indemnities

---

### General indemnity

The Recipient Organisation indemnifies, and agrees to keep indemnified, the University against any Loss suffered or incurred by the University arising out of the Recipient Organisation's use, handling and disposal of the Data and Derivatives or a breach of this Agreement by the Recipient Organisation.

### Injury or damage

The Recipient Organisation indemnifies, and agrees to keep indemnified, the University against any Claim which may be made against the University by any person, in respect of:

- (a) injury or damage to any property; and
- (b) personal injury to, or death of, any person,

that is caused by or contributed to by an act or omission of the Recipient Organisation.

### Release

The Recipient Organisation releases the University, and the officers and employees of the University, from any liability, Loss or damage incurred or suffered by:

- (c) the Recipient Organisation; or
- (d) an employee, sub-contractor, officer, customer, supplier or agent of the Recipient Organisation,

arising out of any act or omission of the Recipient Organisation in connection with this Agreement.

### Benefit of indemnity

The University holds the benefit of this clause 9 on its own behalf and on behalf of University Personnel.

## 10. Termination

---

### Termination on notice

Either party may terminate this Agreement at any time by giving the other party 30 days' written notice of termination.

### Immediate termination by the University

The University may immediately terminate this Agreement by giving written notice to the Recipient Organisation if:

- (a) the Recipient Organisation breaches a material term of this Agreement, which is not capable of remedy;
- (b) the Recipient Organisation commits a breach of any of the provisions of this Agreement which is capable of being remedied to the University's satisfaction, but which is not remedied within 14 days after written notice of that breach has been given to the Recipient Organisation; or
- (c) the Recipient Organisation suffers an Insolvency Event.

### Immediate termination by the Recipient Organisation

The Recipient Organisation may terminate this Agreement immediately by giving written notice to the University if:

- (d) the University suffers an Insolvency Event; or
- (e) the University commits a breach of any of the provisions of this Agreement, and such breach is not remedied within 14 days after written notice of that breach has been given to the University.

### Reservation of rights and remedies

Termination of this Agreement for any reason will not prejudice any rights or remedies already accrued to the University under, or in respect of, any breach of this Agreement.

### Survival

Clauses 5, 6, 7, 8 and 9 survive expiry or termination of this Agreement, survive expiry or termination of this Agreement and are enforceable at any time at law or in equity.

## 11. General

---

### Notices

Any notice under this Agreement must be in writing, and may be delivered to a party at the addresses and numbers set out in the Details of this Agreement.

### Further assurances

Each party must promptly do all further acts required by Law or reasonably requested by the other party to give effect to this Agreement or any obligation under it.

### Assignment

- (a) The Recipient Organisation must not, without the prior written consent of the University and in accordance with this Agreement, assign, novate, subcontract or delegate its responsibility to provide the Services.
- (b) Neither party may, without the other party's consent, assign or novate this Agreement.

### Waiver

- (c) Waiver of any right arising from a breach of this Agreement must be in writing and executed by the party granting the waiver.

- (d) A failure to exercise, a delay in exercising or a partial exercise of a right created under or arising from a breach of this Agreement does not result in a waiver of that right.

### Relationship between the parties

This Agreement is not intended to create a partnership, joint venture or agency relationship between the parties.

### Severability

- (e) A provision of, or the application of a provision of, this Agreement which is void, illegal or unenforceable in any jurisdiction does not affect the validity, legality or enforceability of that provision in any other jurisdiction, or the remaining provisions in that or any other jurisdiction.
- (f) Where a clause in this Agreement is void, illegal or unenforceable, it may be severed without affecting the enforceability of the other provisions in this Agreement.

### Variations

Any variation to this Agreement must be in writing signed by, or on behalf of, both parties.

### Entire agreement

This Agreement replaces all previous agreements between the parties concerning its subject matter and contains the entire agreement between the parties.

### Governing law and jurisdiction

This Agreement is governed by and construed in accordance with the law of Victoria and the parties submit to the exclusive jurisdiction of the courts of Victoria.

### Counterparts

This Agreement may be executed in any number of counterparts and by the parties on separate counterparts. Each counterpart constitutes an original of this, and all together constitute one agreement.

## 12. Definitions and Interpretation

---

### Definitions

**Claim** means a claim, demand, action, or proceeding.

**Confidential Information** means information of the University, including:

- (a) trade secrets;
- (b) Intellectual Property;
- (c) business and financial information, including systems of work, procedures, manuals and training material, reports generated in performance of the Services, products, service costs, prices, profits and sales, new business ideas, business strategies, product and service plans, marketing plans and studies, forecasts, computer programs, databases, computer codes and software ideas, technologies, concepts and designs, research projects and all information connected with research and development, reporting methods and

- (d) any information that is marked, or the Recipient Organisation is told is, confidential.

**Data IP** means all Intellectual Property created, conceived, developed or reduced to practice by or on behalf of the Recipient Organisation using or relating to the Data, including all Intellectual Property in or relating to Derivatives.

**Derivatives** means anything derived by Recipient Organisation from or using the Data, including:

- (a) structural or functional analogues and homologues and purified or fractionated subsets of the Data; and
- (b) any improvements or modifications to the Data

**Intellectual Property** means all present and future rights to intellectual property including any inventions and improvements, trademarks (whether registered or common law trade marks), designs, copyright, any corresponding property rights under the laws of any jurisdiction and any rights in respect of an invention, discovery, trade secret, secret process, know-how, concept, idea, information, process, data, or formula.

**Insolvency Event** means any of the following:

- (a) if a party comes under a form of external administration referred to in Chapter 5 of the *Corporations Act 2001* (Cth) or equivalent provisions in any Law, or has an order made against it for the purpose of placing the party under external administration;
- (b) if a party is unable to pay all its debts as and when they become payable or the party fails to comply with a statutory demand within the meaning of sections 459E and 459F of the *Corporations Act 2001* (Cth);
- (c) proceedings are initiated to obtain an order for winding a party up, or any shareholder, member or director convenes a meeting for the purpose of considering or passing any resolution for winding the party up;
- (d) if a party is a local government organisation, the relevant government takes action to cease the party's operations, or to amalgamate them with the operations of another local government organisation;
- (e) a party becomes bankrupt or enters into a scheme of arrangement with creditors;
- (f) a receiver, liquidator or administrator is appointed for a party;
- (g) an application is made, or proceedings are commenced, with a view to obtaining cancellation of any registration of a party or appointment of an inspector or other officer to investigate any of a party's affairs pursuant to any law; or
- (h) anything analogous to, or of a similar effect to, anything described above occurs in respect of a party.

**Law** means all applicable statutes, regulations, by-laws, ordinances or subordinate legislation in force from time to time anywhere in Australia, whether made by the Commonwealth, a State, a Territory or a local government and, where the context permits, includes the common law and equity.

**Loss** means claims, actions, expenses, losses, liabilities, damages and costs (including legal costs) and indirect losses and damages including those arising from third party claims.

**Personnel** means any employee, officer, agent, contractor, sub-contractor, student or volunteer of the University.

**Privacy Laws** mean Commonwealth and/or State and/or Territory legislation, principles, codes and guidelines in relation to the collection, use, storage and security or disclosure of any Personal Information and/or Health Information (as defined in any applicable Privacy Laws)

**Publication** means any proposed manuscript, abstract, paper, journal article, student thesis, or content of any oral, poster or other presentation.

**Recipient Organisation** means, jointly and severally, the Key Researcher and Recipient Organisation.

**Results** means all outcomes of Recipient Organisation's use of the Data, including all information, data, Derivatives and tangible objects arising from Recipient Organisation's use of the Data.

## 16.1 Interpretation

- (a) Headings and bold type are for convenience only and do not affect the interpretation of this Agreement.
- (b) The singular includes the plural and the plural includes the singular.
- (c) The word 'includes' in any form is not a word of limitation.
- (d) Words of any gender include all genders.
- (e) Other parts of speech and grammatical forms of a word or phrase defined in this Agreement have a corresponding meaning.
- (f) An expression importing a person includes any company, partnership, joint venture, association, corporation or other body corporate and any government agency as well as an individual.
- (g) A reference to a clause, party, schedule or annexure is a reference to a clause, party, schedule or annexure of, or to, this Agreement.
- (h) A reference to any legislation includes all delegated legislation made under it and amendments, consolidations, replacements or re-enactments of any of them.
- (i) A reference to a document includes all amendments or supplements to, or replacements or novations of, that document.
- (j) A reference to a party to a document includes that party's employees, subcontractors, successors and permitted assignees.

- (k) A promise on the part of 2 or more persons binds them jointly and severally.
- (l) No provision of this Agreement will be construed adversely to a party because that party was responsible for the preparation of this Agreement or that provision.
- (m) A reference to a body, other than a party to this Agreement, whether statutory or not:
  - (i) which ceases to exist; or
  - (ii) whose powers or functions are transferred to another body,is a reference to the body which replaces it or which substantially succeeds to its powers or functions.