

SOIL CRC

Performance through collaboration

MILESTONE REPORT

Project 2.3.002

**Visualising Australasia's Soils:
Extending the Soil Data Federation**

Impact Research Report



SOIL CRC MILESTONE REPORT

CONTENTS

1. Introduction	3
2. Background	3
3. Methods	6
4. Results	7
5. Discussion	21
6. Conclusion	26
7. Acknowledgements	26
8. References	27
Appendix A.....	28
Appendix B.....	30

AUTHORS

Dr Alison Ollerenshaw & Dr Carolyn Staines (Research data collection and analysis)

Prof Helen Thompson & A/Prof Peter Dahlhaus (Research management and review)

KEYWORDS

Impact research, benchmarking, interviews, national soil data federation, opportunities and challenges with federating soil data

CITATION

Ollerenshaw A., Staines, C., Thompson, H., & Dahlhaus, P. (2024). *Visualising Australasia's Soils: Extending the Soil Data Federation. Impact Research Report. Milestone Report 14.2. Soil CRC Project 2.3.002. Cooperative Research Centre for High Performance Soils, Callaghan, NSW.*

1. INTRODUCTION

Visualising Australasia's Soils (VAS) is an interoperable spatial knowledge system, federating soils-related data and datasets, information, and knowledge for Australasia. Funded by the Cooperative Research Centre for High Performance Soils (Soil CRC), the first phase of the project was completed in 2021 with the launch of the VAS data federation which includes a cloud-based data aggregator to facilitate data access to shared, public and private soils research data.

Funding was awarded for the expansion of VAS (2022-2024) with the objective for it to become an independent and enduring soil research data federation that is self-sustaining and supports the data-management and data-sharing needs of end-users and data custodians. To this end, VAS will be extended, contributing to research and education for Australia's agriculture data through the availability, extension, and application of soils data for the industry.

Research was conducted during the second phase of VAS. Employing social science methods, this research report documents the current role of VAS and its anticipated value, contribution and impact on users. It outlines the barriers and opportunities for VAS, identified by project partners through their interactions with VAS. The research findings, while documenting the insights and experiences of project partners about VAS, also informs the project's direction to ensure the needs of partners continue to be met.

2. BACKGROUND

2.1 VISUALISING AUSTRALASIA'S SOILS: PHASE 1

VAS is a Soil CRC project that commenced in 2019. With initial investment of \$1.18M (including \$510,000 funding from the Soil CRC), the principal aims of the project were to: (a) make Australasian soils data visible and reusable in accordance with agreed governance principles; and (b) maximise the value of research data for the Soil CRC, its members, and the broader agriculture industry.

Fifteen project partners collaborated on the project in Phase 1, including Soil CRC and partners, and organisations representing the agriculture industry in Australia and New Zealand, including: Landcare Research New Zealand, University of Tasmania, University of Southern Queensland, Birchip Cropping Group, Burdekin Productivity Services, Central West Farming Systems, Gillamii Centre, Herbert Cane Productivity Services, Landmark, Liebe Group, Mackillop Farm Management Group, Mallee Sustainable Farming, North Central Catchment Management Authority, Riverine Plains Inc., Southern Farming Systems, Western Australian No-Tillage Farmers Association and Wimmera Catchment Authority.

During initial development of VAS, the project was guided by the following six objectives:

1. Establish a soil research data federation, based on agreed data stewardship and governance frameworks, that allows Australasian soils data from all sources (private and public), to be discoverable to all Soil CRC participants through an intuitive-to-use internet portal.

2. Co-develop, with CRC end-users, data stewardship governance frameworks for sharing data.
3. Collaborate with project participants to develop the capability to interoperably provide their data to the spatial knowledge system according to the rules that they set.
4. Co-design and co-develop training programs with project participants to improve data stewardship and governance, and to improve the collection and quality of data and metadata for inclusion in the knowledge system. Co-design and deliver online educational materials for farmers and researchers to make best use of the data in the knowledge system.
5. Co-develop simple, web-based tools to assist in spatially visualising soil data, searching and filtering data, downloading datasets, and publishing data in VAS, to suit the needs of all Soil CRC participants.
6. Co-develop, with CRC end-users, dynamic models that are applied to the interoperably federated data to answer frequently asked questions such as finding temporal trends in soil performance indicators.

To this end, the project aimed to facilitate the provision and access to soils data and information that was more Findable, Accessible, Interoperable and Reusable (FAIR). VAS federates data from disparate sources in both the public and private sector, providing a foundation for making agriculture and soils data visible. VAS has been integrated with other national and international initiatives that are liberating soils data within the FAIR framework. The availability of soils data through the VAS federation aims to inform decision making for farm management, and develop resilience and profitability of the agriculture industry.

Phase 1 of VAS was completed in mid-2021, having successfully met its original project aims and key project deliverables. In concluding this phase of the project, future directions for VAS were identified alongside recommendations for the continued expansion of the VAS system (Dahlhaus et al. 2021).

2.2 VISUALISING AUSTRALASIA'S SOILS: PHASE 2 (PROJECT EXTENSION)

Following from the success of VAS Phase 1, funding for a three-year extension to VAS was awarded in late 2021, totalling \$2.29M, with \$945K contribution from the Soil CRC. During this second phase of development (2022-2024), VAS will become an independent and enduring soil research data federation, delivering access to data and information to support research and educational activities for the agriculture industry. Additional funding for Phase 3 of VAS (2024-2027) has since been announced.

The extension activities for VAS are led by CeRDI in partnership with 14 project partners (all members of the Soil CRC), with broad representation across the Australian agriculture industry. Project partners include grower groups, farmer data cooperatives, and researchers who will provision data to VAS and actively use VAS for data sharing and data management purposes. Project partners include:

- Birchip Cropping Group
- Burdekin Productivity Services

- Central West Farming Systems
- Facey Group*
- FarmLink Research Ltd.*
- Herbert Cane Productivity Services Ltd
- Holbrook Landcare Network
- Liebe Group
- MacKillop Farm Management Group
- North Central Catchment Management Authority
- Riverine Plains Inc
- Southern Farming Systems
- West Midlands Group
- Western Australian No-Tillage Farmers Association
- Wheatbelt Natural Resource Management Inc.*
- Wimmera Catchment Management Authority

(*Denotes organisations that commenced partnership with VAS during Phase 2).

2.3 IMPACT RESEARCH

A structured approach to conducting research incorporating social science research methods was implemented during the second phase of VAS. The research aimed to examine partners experiences and use of VAS, and the challenges and opportunities for soil data sharing through VAS. The insights from project partners informed the co-design, implementation, data sharing, and participation in VAS, while increasing the breadth and depth of the soil data within the federation. Key questions guiding this research included:

- How is VAS being used and has it met expectations?
- What are the opportunities for data sharing through VAS?
- What are the barriers for data sharing through VAS?
- How can data sharing be facilitated using VAS?
- What features/functions of VAS have facilitated and/or reduced data sharing?

Research interviews were conducted with project partners during the first and final years of VAS Phase 2. An interim report (Ollerenshaw et al. 2022) documented the research findings from nine interviews with project partners during the first year. Interviews with a further five project partners were conducted in the final year. This report documents the findings from both rounds of interviews, conducted during Phase 2 of VAS.

3. METHODS

Between 2022 and 2024, semi-structured interviews were conducted with 18 staff from 14 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 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.

An initial round of project meetings was conducted with project partners. During these meetings, the impact researchers were introduced. Invitations to participate in the research were then circulated via email to key contacts (CEOs, managers, project staff) at each of these partner organisations. 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.

All 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 is 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).

4. RESULTS

4.1 OVERVIEW

Four key themes and related ideas were identified. These are represented in the following figure (Figure 1) and described in detail in the descriptive analysis of the findings and supported by representative quotes.

The analysis commences with findings about the perceptions and usage of VAS. This is followed by an assessment of the current and anticipated benefits and challenges with VAS, and suggestions for improved data reporting and sharing with VAS. The report concludes with a discussion outlining key points arising from this research.

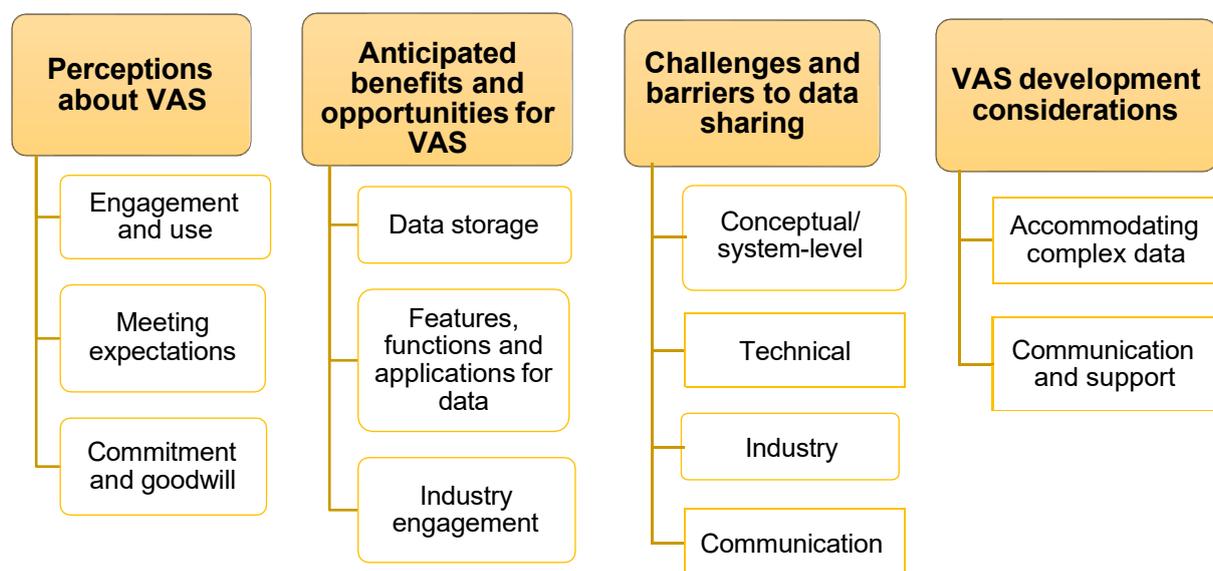


Figure 1. Themes and ideas identified through analysis of interviews with project partners ($n = 14$).

4.2 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 will be examined further, using quotes and excerpts from participant interviews to demonstrate and support these ideas.

4.2.1 Engagement and use

During the first round of interviews, most participants reported low to moderate interaction with VAS, with one participant having 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 being 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 with it (Interview participant, Time 1).

So really looking at the portal from our end, at the moment, we 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 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 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 involved in. 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 have been dealing with as part of a soil moisture probe and a weather station network project we've been rolling out (Interview participant, Time 1).

Thematic analysis from the second round of interviews reveals new insights from participants about VAS. Interviewees have a foundational understanding of VAS and had conducted a preliminary review of it. Moreover, some interviewees had commenced entering 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 that's sort of what the project's about, a bit is us getting familiar with it (Interview participant, Time 2).

4.2.2 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 the primary reasons for claiming this. 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 tempered their feedback and expectations somewhat. 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 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 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 the end of the project to this and if the more we put in, 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 wider 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 strongest expectations for VAS were associated with interviewees that: (a) had been 'long-term' partners on VAS (since Phase 1 of the project); and/or (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 the system appears to have formed. These ideas are evident in the following comments:

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 earlier 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 were anticipated but not yet achieved. In the second round of interviews, the expectations for VAS for one informant had not been fully realised. However, this same informant acknowledged that the contribution of VAS for other stakeholder groups was a strength. 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 ... 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 for us being involved with because anything that helps our farmers is something that ... we will stand behind (Interview participant, Time 2).

4.2.3 Commitment and goodwill

Ongoing commitment and goodwill towards VAS were consistently documented in the first round of interviews. These ideas were associated with the potential value and expectations for VAS that had been largely developed following discussions with the VAS project team. 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 ... VAS is to pull a lot of the other old ... 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 is around looking at having as many layers as we can, and then 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).

4.3 ANTICIPATED BENEFITS AND OPPORTUNITIES FOR VAS

Informants frequently referred to the prospective opportunities and anticipated benefits for VAS. These were focused on three areas, identified during both rounds of interviews, including: (1) data storage; (2) features, functions, and applications for data; and (3) industry engagement. These ideas will be examined in detail, below.

4.3.1 Data Storage

Most informants across all interviews indicated that the data-storage and data-management capabilities were a key benefit of VAS. The capacity to store new and existing data in VAS was important in supporting their own organisations' (current and future) data storage needs, and overarching research capabilities. Interviewees, particularly in the second 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 for identifying 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).

4.3.2 Features, functions and applications for data

In the first round of interviews, participants referred to the practical benefits associated with VAS. This included VAS functionality and technology enabling searches 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, also anticipated using VAS for data extrapolation and application across a range of organisation-specific activities. For some informants, accessing detailed data and information offered the potential for new insights (through interrogation of organisational data) that would inform their research and data-relevant activities, including research reporting and funding applications. It was also identified that the data could be used to convey key soil information to their member-farmers. For these informants, using VAS for their soils data will provide additional mechanisms for communication and information sharing for 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, 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 locally, nationally, much easier or robust (Interview participant, Time 1).

In a perfect world, if you had a lot of data on there you could generate a 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 it'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 ... 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 second round of interviews shared insights about the perceived contribution of VAS. These include the provision for managing their soils data in VAS, and the security associated with this. The application of VAS for discoveries in the data was identified as beneficial to the organisation, and the industry. Creating reports and visualising soils data were additional benefits of VAS. These ideas are represented in the following quote:

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).

Participants outlined their expectations for VAS. The opportunity to conduct extensive reviews of their soils data in VAS was a benefit that would inform their organisations' future research and planning activities. The potential for VAS to enable knowledge transfer within organisations, and between organisational staff was another identified benefit:

Just from a data storage consistency perspective, within the organisation, I think that it'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 was interested in this program mainly because of the visualisation part of it. Like we get all these data sets 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 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 [VAS], 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).

4.3.3 Industry engagement

Interviewees identified other opportunities for industry engagement using VAS. The opportunity for generating industry-relevant information and the prospect of industry engagement, including research collaborations, was a benefit aligned to 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 others, VAS presents an opportunity to engage with their farmer-stakeholders, particularly the member-base at the partner–organisation level. This would necessitate the involvement of farmers to supply and/or enter their data on VAS. Establishing farmer trust in VAS is an important consideration in relation to the project's growth and expansion (N.B. This will be explored further in a subsequent section of this report):

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 [and] 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 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 the industry level, VAS can offer project partners an opportunity for involvement with state and national government agricultural organisations sharing and accessing industry-relevant data. Several interviewees commented on their expectations associated with accessing data from other organisations, including data from government agencies. This would fulfill an important need for some, in providing access to industry-relevant data and information. These ideas are highlighted in the following representative quotes:

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 ... 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 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).

4.4 CHALLENGES AND BARRIERS TO DATA SHARING

Without exception, 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 comprising each category are captured in Figure 2.

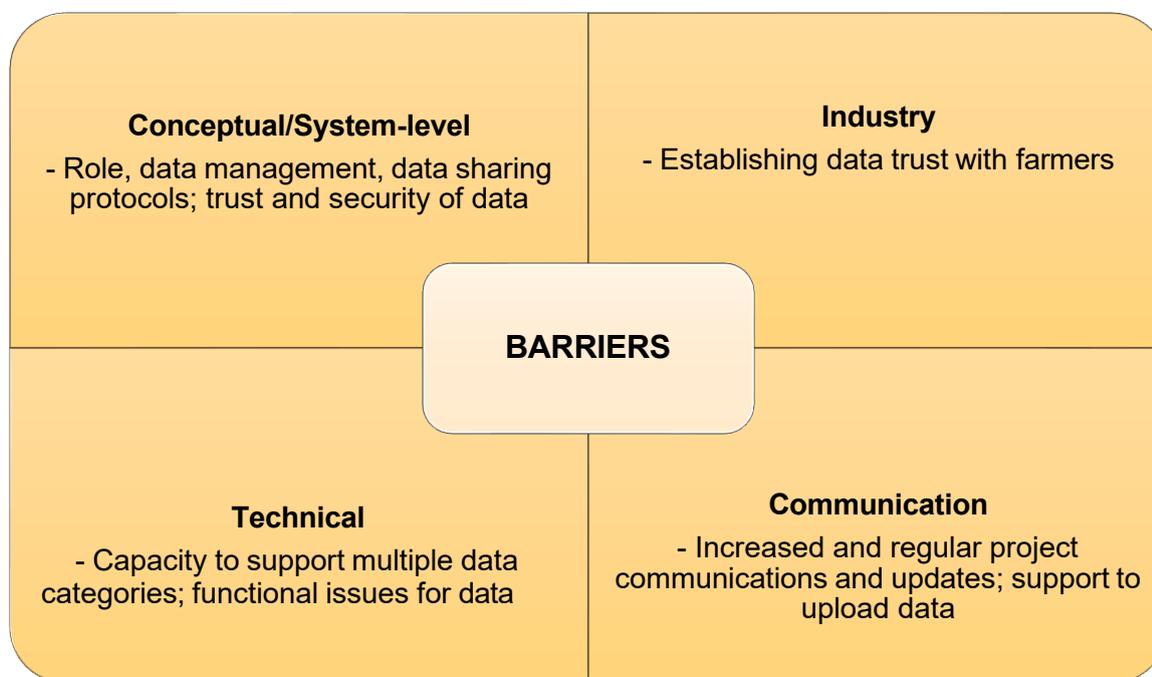


Figure 2. Barriers associated with VAS identified through analysis of interviews ($n = 14$).

Numerous barriers were identified during the first round of interviews. These were wide-ranging and represented actual and perceived barriers with VAS. Many of the reported barriers were representative of 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 Table in Appendix B.

4.4.1 Conceptual/system-level barriers

In this report, conceptual/system-level barriers refer to discrepancies in understanding about the overall project, its operations, objectives, and expectations. During the initial 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 during the second phase of the project, including through the provision of information, and advice about how partner organisations establish their data governance arrangements.

Participants across all interviews remain 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 second round of 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 the first round of 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 about assigning data-sharing permissions, and the complexities associated with assigning the permissions. However, there was an exception to this, with one interviewee commenting that the data-sharing agreement for 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 ... 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 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 comes up a lot (Interview participant, Time 2).

Another issue raised during the first round of interviews was 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 include the ongoing management and maintenance of their data within VAS, data 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 soil's context, and the federation of soils data and its maintenance over time, was raised by some project partners. There remains some ambiguity about the role of VAS within the broader national soil context.

4.4.2 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 organisation's 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 considered a substantial barrier that was limiting the breadth of data that partner organisations could include on VAS. This issue also has research implications: the lack of complex research data in VAS was viewed by some as limiting the scope and breadth of data extraction and discovery. 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 continued to report challenges associated with entering complex soils data into VAS and uncertainty about retaining anonymity and value in the data.

4.4.3 Industry barriers

Industry barriers are challenges associated with issues linked to the agriculture industry, specifically the sensitivities identified by partner organisations with uploading farmer-members' data in VAS.

Recognising and addressing issues of trust associated with data management and sharing farmer data was a persistent issue. The security of soils data shared in VAS and maintaining the trust that exists between partner organisations and their member-farmers were dominant concerns for partner organisations. Consistently, informants commented about the lack of farmer trust in sharing their soils data. It is important to establish 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.

During the second 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. Notwithstanding these concerns, the following comment demonstrates the deeper understanding that interview participants have of VAS. It identifies that, as data custodians, they have the capacity to restrict public access. 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, 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).

4.4.4 Communication barriers

The lack of project communication was widely discussed by participants in the first round of interviews. There were widespread calls for 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 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 project communications and promoting VAS usage.

Another barrier raised during the first round of interviews were 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 second round of interviews. Only a small subset requested assistance with using VAS and entering their soils data.

4.5 VAS DEVELOPMENT CONSIDERATIONS

Interviewees offered suggestions for developing VAS including recommendations to address the challenges and barriers with VAS as previously identified. Two key areas where VAS could be further developed include: (1) accommodating complex soils data; and (2) increased project communications and support. These ideas will be examined in this following section.

4.5.1 Accommodating complex data

Accommodating complex soils data in VAS, and increasing the ease of access, were identified areas for improvement. It was recommended that VAS accommodate complex soils data with areas for expansion including 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 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).

The expansion of available soils data in VAS was also called for:

- 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 comments:

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? Is it because there were 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, like, if we were to import that data, how do you reflect the differences between a known point 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 P was a format that you could use to upload and you could just have some extra columns you got to fill in (Interview participant, Time 2).

4.5.2 Communication and support

The importance of communicating about the project was widely discussed by participants in the first round of interviews. Increasing project communication between partners 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 ... 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 a workshop and/or conference about data management and data agreements that could 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 ... 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 a recommended approach as 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 ... 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 that they could contact regarding queries about using the self-service system:

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).

5. DISCUSSION

Research was conducted to understand the impact of VAS from project partners during Phase 2. Social research methods were implemented, and in-depth interviews with project partners were conducted. Interim findings from nine interviews were previously reported (Ollerenshaw et al. 2022). A further five interviews with project partners were conducted in late 2023 and early 2024. The qualitative data presented in this report offers insights into project partners' experiences with and perceptions about VAS. It includes an exploration of the barriers they have experienced, and the potential opportunities for data sharing that is anticipated with the expansion of the VAS data federation.

5.1 INTERVIEWS: ROUND 1

The thematic analysis from the initial round of interviews ($n=9$) indicated that they had limited engagement and use of VAS. Interactions with VAS were mostly tentative, with many yet to comprehensively use or interact with VAS; some partners had not uploaded data into VAS or had uploaded only minimal data sets into the system. The lack of interaction with VAS was, for some, due to a lack of time to use and interact with it, and others remained unsure of how to use VAS to upload their data.

There was, however, widespread agreement that VAS was meeting expectations, particularly with the development and launch of VAS in 2021. This was widely viewed as the foundation point from which future engagement and data sharing in VAS would occur and be expanded.

In the early stages of VAS Phase 2, informants were cautiously optimistic about the project. This sentiment was closely aligned with the future contribution and value of VAS as a soil data federation for the management of data that will benefit individual organisations and the agriculture industry. The potential for data extrapolation and the sharing of data for organisational and industry purposes were identified as a principal benefit of a fully developed VAS system. There exists a breadth of understanding and widely held anticipation for VAS by project partners. The opportunities to access soils data through the federation would provide new insights and knowledge that could change current practices for soils in agriculture. This would facilitate greater engagement and collaboration for the sector through a shared understanding that the data in VAS can generate.

During the earlier interviews, multiple and various challenges for VAS were identified. These were widely predictable, relative to the project's development. For some partners, there remained a lack of overall understanding of the project-specific data leading to uncertainty in their expectations for the project. This was evident in relation to the concerns raised about the use of farmer data and ensuring data privacy. Project partners were widely concerned about sharing sensitive soils data in VAS. The security and privacy of this data in VAS remains a concern.

There was also uncertainty about data security and data management of VAS over time. In addition, some partners were unsure of the role of VAS as a data federation within the wider industry, in which there are several emerging national data repositories.

There is an opportunity for further communication and information about VAS to improve clarity about its objectives, enabling the differences, similarities, and opportunities of VAS and other national online knowledge systems, to be fully understood.

Project communication was also identified as a barrier with informants reporting a lack of regular communication about the project preceding these interviews.

Six key areas of improvement and development of VAS were identified following the first round of interviews. These provided a framework for recommended activities for VAS during the remainder of Phase 2 (for further details, see Ollerenshaw et al. 2022). These included:

- Promoting greater interaction and use of VAS.
- Increasing data sharing.
- Listening to and obtaining feedback about VAS from project partners.
- Championing the VAS experience within and between partners.
- Developing guidelines to enhance trust associated with member-farmer data.
- Establishing ongoing and regular project communications for project partners.

5.2 INTERVIEWS: ROUND 2

By the second round of interviews, all interviewees reported having interacted with VAS, with some having entered their soils data into the system. The application of data in the portal suggests wider exposure to, and experience with using VAS, while also demonstrating their commitment to the project. There was broad confirmation that the project was meeting expectations (a sentiment most frequently cited by long-term project partners and those that reported greater interaction with VAS). This was supported by statements indicating that VAS is easy to use and assists in visualising soils data.

There was wide-spread agreement that VAS is an important and valued soils data-management system. The established features of VAS, including data security, and tools for generating outputs, including reports and visualisations from the soils data, were widely reported and highly regarded. Widespread benefits – and efficiencies – associated with VAS at the organisational-level were identified. VAS offers improved and long-term access and storage of existing soils data thus delivering organisation-wide data access and soils data knowledge while also streamlining the transfer of soils data to other stakeholders and project partners. The potential contribution of VAS to generating new knowledge from the soils data, leading to new industry-wide insights, was also identified.

Although numerous barriers for VAS were identified during the second round of interviews, the reported challenges were noticeably fewer than for the first round. This is despite there being a smaller number of participants involved in the second round. There remains continued uncertainty about VAS in relation to the soil data sharing and data ownership, especially when soils data has been collected with other stakeholders and involving funding bodies. Sharing private data, where a trusted relationship exists between the partner organisation and key stakeholders (project funders), was also raised. Sensitivities associated with managing and sharing farmer data remains a

persistent concern. There is however an understanding that VAS allows partners to restrict access to safeguard sensitive and private soils data.

Challenges were identified with assigning permissions for sharing data. This included uncertainty and difficulties in assigning ownership and access to soils data in VAS. The information about data-sharing permissions was described by some as overly complex. Further, and although not widespread, there remains some uncertainty about the role and context of VAS within the national soil data environment.

The technical challenges identified the importance of accommodating complex data in the portal. Soil data is complex and often includes potentially identifiable information (e.g., GPS coordinates). Data custodians are uncertain about how to assign complex data in VAS and how to de-identify it.

Various recommendations for VAS were offered by participants in the second round of interviews. Prioritising development of the system to accommodate diverse and complex soil data was identified. Streamlining the user experience associated with entering complex data was also identified.

5.3 SIMILARITIES AND DIFFERENCES BETWEEN INTERVIEW ROUNDS

Similarities exist 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 soils data in VAS remains a challenge. The soils data owned by project partners is intricate and cannot easily be assimilated into VAS. 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. 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 a continued area for 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. 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).

Furthermore, many of the recommendations outlined in the interim impact research report (Ollerenshaw et al. 2022) have been actioned which may have led to fewer challenges being reported in the second round of interviews. For example, the recommendations for increased project communications led to the regular preparation and distribution of the VAS newsletter, Soil CRC conference presentations, information sheets, and videos to support data entry into VAS.

5.4 FUTURE CONSIDERATIONS FOR VAS

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 addition to the continuation of activities recommended in the interim report, and undertaken in the next phase of VAS:

1. Uncertainty about sharing soils 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 flows charts 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: Develop resources and information about VAS for farmers. Developing fact sheets about VAS and outlining the ways in which it safeguards soils data and allows soil data custodians to restrict access to private soils data are important to 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 soils 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 soils 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 soils 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 will closely align with the activities for VAS towards an enduring soil data federation, in Phase 3. The value proposition of VAS is integral to its longevity, necessitating ongoing sharing of soils data by custodians. This will lead to broader industry benefits overtime. Incentivising data sharing in VAS through 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 system users and project champions from each user/stakeholder group.

Understanding the impact of VAS in this third phase 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.

6. CONCLUSION

Social research methods were implemented to document the user experience and impact of VAS during this second phase of VAS. A range of insights about the use and perceptions of VAS were collected through interviews with project partners. Both rounds of interviews highlighted the complexity of available soil data, and numerous uncertainties in sharing this data in VAS. The insights documented in the first round of interviews led to a selection of recommended activities. Insights from the second round of interviews revealed some differing perceptions and experiences with VAS. There was an increased use of VAS, and soil data entry into VAS had commenced. There was broad agreement about the value of VAS, particularly its contribution to soils data management. Fewer challenges with VAS were also reported. As Phase 3 of VAS commences in late 2024, and as the federation of soils data becomes fully realised, 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 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.

7. 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 wish to thank those individuals who participated in an interview for this research.

8. REFERENCES

Dahlhaus P.G., 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. Cooperative Research Centre for High Performance Soils, Callaghan, NSW. 14p.

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). 119p.

Ollerenshaw A., Staines, C., Thompson, H., & Dahlhaus, P. (2022). Visualising Australasia's Soils: Phase 2. Project 2.3.002. Interim Impact Research Report. Cooperative Research Centre for High Performance Soils (Soil CRC). 32p.

Saldana, J. (2016). The coding manual for qualitative researchers (3rd ed.). Singapore: Sage.

APPENDIX A

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

Table 1. Actual and potential barriers to data sharing (and participation) in the VAS project

Category	Barriers	Representative quote*
Conceptual/ System-level	Data management agreements and permissions	<p><i>Yeah, I think there's always been some trust issues with some farmers. Probably unavoidable, but yeah, and I suppose I've been reassured by the (VAS project) team a few times ... that it is secure. But yeah, uh, how locked down that data is to access, and what people can do with the data, and you can you got different levels of access to ... (Interview participant, Time 1).</i></p> <p><i>Because if we're going to put things in there, we need to know that a) we can get them out and b) that its functionality remains post project (Interview participant, Time 1).</i></p> <p><i>I found it incredibly difficult to know what permissions to assign or not assign to data. Some of the language and terminology made no sense to me. I was like 'I can't remember what they were', but there were these check boxes, so I put everything into the Excel template and uploaded it, and then it asked me all these different permissions things and I went 'well I'm not that into all the legal stuff' and I just had no idea. I don't know what to tick there. So, like I think in terms of having a clear understanding of what the actual things that are taking off mean would be helpful because more often than not, I think the people uploading data are going to be people like me, they're not going to be someone who reads legal papers and understands all the terminology. So, I think like it needs to be written in layman's terms and you know clearly stated what it is that you're agreeing to (Interview participant, Time 2).</i></p> <p><i>I guess the tricky part that we've come up against is just that the ownership and the access requirements of you know, just knowing who, like what levels of access we should be putting on to the data. Yeah, it's probably not that complicated, but when I looked at it last year, I thought, oh, this is going to be put off because it's a bit hard (Interview participant, Time 2).</i></p> <p><i>Maybe we need to put the thinking into who we want to give permissions to, to access it. Whether it is at a public level or just a group level ... and I guess maybe some more information around what each of those different access levels ... means (Interview participant, Time 2).</i></p>
	Uncertainty around data ownership and data sharing	<p><i>We're still publishing papers from that dataset, so we've held off doing much and really going into it much because it's tightly held (Interview participant, Time 1).</i></p> <p><i>I was just going to say it's probably the biggest barrier is our data sharing agreement with farmers at the moment and that it's their data, it's not ours. And anything that we have collected as part of a project, it's on a farmer's property and ... it's collected using government funds, who funds the projects generally. So, data ownership is probably a</i></p>

		<p><i>big barrier and our current sharing arrangements with farmers is probably the biggest barrier at the moment (Interview participant, Time 1).</i></p> <p><i>Well, we know that producers will have concerns about ... they don't want people to know if there's areas that are really acidic or really salty that might affect their land prices. So that's kind of a bit of our ... from their point of view ... if there's any issues there, they're not going to want to potentially share that data (Interview participant, Time 1).</i></p> <p><i>... it's a double-sided sword, this view I've got. At the moment all our soil data is currently stored in Excel workbooks ... Excel spreadsheet and we very rarely dive back into it. And with some of those soil tests we've been given permission to use those results for further research or whatever. The growers signed it off. But then there's the catch of sometimes the soil testing is subsidised by a funding body or another group, and so then you go "Well, who actually, at that point, owns the data and what flexibility have we got to put it into portals like this? (Interview participant, Time 2).</i></p> <p><i>... when you collect data for a project, where does that data go? Or how can we then display it? So yeah, it's just knowing each different funding source, what's allowed and not allowed with the data (Interview participant, Time 2).</i></p> <p><i>I guess the tricky part that we've come up against with is just that the ownership and the access requirements of ... just knowing who, like, what levels of access we should be putting on to the data (Interview participant, Time 2).</i></p>
	Data trust and data privacy	<p><i>Farmers have generally got, I suppose trust issues, but they've got to be, have, you know, they're so used to big companies stealing their data because now all the John Deere's and everything and yield maps, using their software and all that kind of thing and in the cloud. So, their data gets stolen all the time. So, I think if we can get that trust into that system (Interview participant, Time 1).</i></p> <p><i>There is always a level of ... not necessarily uncertainty, but the willingness to share data freely, particularly by farmers. ... essentially, what they would describe as kind of intellectual property or commercial kind of in business data. Of course, the contradiction is that most of the data that any large multinational corporation wants to get, they could get, largely from satellite information anyway. So, I think the whole trust debate of data and then will the data actually be used against them. So, if you look at it from a climate change perspective ... the people get very, very nervous and yes, so that it's almost that kind of public, private, good kind of discussion when we come into areas such as climate change or sequestering carbon and land use planning and all those things where data can be misinterpreted or paint the picture for a particular audience (Interview participant, Time 1).</i></p> <p><i>Growers are very, very conscious of their privacy, and so we have been treading very lightly here and I've requested to Peter and Andrew that any information is locked down really heavily because there's a perception from farmers ... that if we were responsible for releasing any individual data, which I know this project isn't, it's not the intent at all, but if somehow, someone got access to that, then you know it could be seen as us endorsing the government regulations and making the farmers life harder. So, it is probably another reason why I haven't, you</i></p>

		<p><i>know, dived right into it and really got that excited just because there's always this thing weighing on my mind that our members (Interview participant, Time 1).</i></p> <p><i>...there's always a level of trust that farmers will put into things. So, if there's that trust in there (in VAS), they'll probably do it. But yeah, there are a lot of concerns around data privacy with farmers at the moment. You know, they've always been open about sharing things, but I think there is a bit more of a realisation that maybe they should put a bit more value on what they do put out there (Interview participant, Time 2).</i></p>
	Data interoperability with other national soil systems	<p><i>So, it's about ensuring its interoperability with other systems, whether that be the national soil strategy, or we were talking about connectivity to Decipher Ag for example so it's about how can it fit with the other systems around it and into the future (Interview participant, Time 1).</i></p> <p><i>I'm probably a little curious as to how this might be similar or dissimilar to the CSIRO product in terms of a soils database and why or how this is actually different (Interview participant, Time 2).</i></p>
	Data governance arrangements	<p><i>... in terms of that data sharing and getting support with managing you know how we set up agreements and contracts with our farmers and with how we share data out of ... VAS, that was another major issue (Interview participant, Time 1).</i></p> <p><i>Because people would be sensitive about sharing yield information ... and then soil stuff is probably just put in the same bucket even though it is harder to turn soil information into economic value, yes, you couldn't rule it out. ... It really means that we need probably to be a lot more hot on agreements up front about, before people engage in projects and delivering things, that they're happy for their data to be publicly available (Interview participant, Time 1).</i></p>
	Project longevity; ongoing management, and maintenance	<p><i>I think you want to get to that stage ... the federal government's going to fund it ad infinitum. It's going to have its own self-perpetuating ... because mostly things happen, for while they're really good, big database, and then no one uses it ... I've seen a lot of those over the years that I like. Peter's got some ideas around that, but having it big enough to people say, 'this is this is the place to go to' (Interview participant, Time 1).</i></p> <p><i>And there's ongoing maintenance to the site because it's not owned by us and therefore we can't take it and get any old technical person in to fix any glitches that might occur, which then is a risk in terms of the safety of our data and being able to access it and manage it the way we need to because, post the project, we're going to still ... want to put data in there constantly and data requirements may change et cetera. And so, it's got to have that long term, ongoing maintenance for it to be less of a risk for us to put things in (Interview participant, Time 1).</i></p>
Technical	Perceptions, difficulties, and uncertainty in using VAS	<p><i>In terms of it's a functionality that's probably not what I expected. It's yeah, 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).</i></p>

		<p><i>So visual, like actually visualising the soil data, like even in the tabulated form that it comes up in...that's quite hard to read. And also like the graph component where you can graph soil data and soil properties. We were hoping, our expectation was, that it could graph soil properties between years, and I haven't been able to work out how to do that (Interview participant, Time 1).</i></p> <p><i>Data access – I'm not sure what functions it will have in terms of being able to get data to farmers. Like if farmers can login or you if you know it's got tools where we can, I guess, it will have tools work and export data in an easy way that'll get to a farmer (Interview participant, Time 1).</i></p> <p><i>So, we've got the data prepared but we're still, I suppose, we haven't uploaded it yet and we're a bit unsure of how to do it (Interview participant, Time 1).</i></p> <p><i>I found it really hard to know what I should call the samples because sometimes we collect a lot of our samples based on growing and paddock. But then you don't necessarily want to put the grower details in there because you want to keep it a bit nondescript. But then the paddock can still sort of give it away and then it's pinpointed to a location so people would probably going to know who's it is, if you're looking at local context (Interview participant, Time 2).</i></p>
	<p>Limited data categories/parameters</p>	<p><i>I think yield...probably, trends of yield overtime are important for farmers. I mean the other point of view is you don't need to go to that extent, it's a good soils database for Australia. It will be there forever more and be added to so the value in that's huge. It depends what level you're looking at, but from a farmer point of view you want to have a look at the data and come up with a management solution to issues and that require doing like an economic analysis on that yield. It's pretty hard to do to do that (Interview participant, Time 1).</i></p> <p><i>It needs contextual information around the soil data so that we can say if it's in a permanent pasture or whether it's under tree lines or if its intensive management or has had ten tonnes of lime put on it or whatever. It needs that next level of management data to provide context for soil data interpretation (Interview participant, Time 1).</i></p> <p><i>And I was thinking ... what about other types of data that could overlay this. You know, topographical maps and land class systems or something like that. Would that help with contextualising as well? Vegetation map to sit behind it or in front of it? I would have thought topographic at least. Like, yeah, digital and elevation, digital elevation model or something (Interview participant, Time 1).</i></p> <p><i>But I think it's really important that regardless of whether we've got that information or not that the data can still be put up because I hate a whole data set to not be able to be put on because we're missing one relatively minor piece of information (Interview participant, Time 1).</i></p> <p><i>... sometimes you have point data which have really got a lot of attributes associated, but then sometimes we get spatial data which is geographically rich but attribute poor. ... I'd like a few more examples of how those other data</i></p>

		<p><i>sets get swallowed up and get, not swallowed up, absorbed into bit the project going forward and how even the template had had like for GPS points, had one point where sometimes our data will be collected over a transect and just how do we account for those variations in that there's quite a diversity in for our data looks like and just trying to massage it so it fits I think a useful exercise (Interview participant, Time 1).</i></p> <p><i>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).</i></p>
	<p>Diversity and complexity of data (unsupported in current structure)</p>	<p><i>We've still got a lot of work to do with my data set to try and make it extractable for the data. As I said, there's 13 years of research trials embedded, overlaying fairly complex design. So, try extracting and let's just bring out the effect of these treatments on PH, for example, or soil carbon, that's still being proving a bit tricky to get hold of (Interview participant, Time 1).</i></p> <p><i>... a lot of our data is single point lots of treatments and kind of temporal data, so it's many years' worth of data of, say, treatment AB&C and its effect on soil properties and when we kind of put that in, we I think both us and the VAS team realised that, well, how do we deal with this and how do we make sense of it within the layers and things like that (Interview participant, Time 1).</i></p> <p><i>The data import template was that you sent I got was as good as any I've seen. It's just the nature of soil data. Once you get away from the state soil database you get all sorts of soil data like I've got, I'm starting to get biological numbers. I'm starting to get lots of diverse data. Especially my project, we're not very descriptive about what people measure, so I can sometimes end up with quite a diverse range of numbers, and there might only be, for this thing, there might only four examples of it. And so then how do you massage that into the rest of the database? The levels of information jump around quite a lot (Interview participant, Time 1).</i></p> <p><i>... 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 ... in lots of different formats, and lots of different organizations 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></p> <p><i>...I've showed it (VAS) to a few farmers as well and I guess farmers work on a paddock scale so they're looking at how accurate is this for my paddock, which might be 100 hectares or something like that ... They question whether ... they can pull any value out of out of it because the data, whether the data is going to be specific enough for their paddock or for the paddock, so at this stage they're seeing it as perhaps you know, the data is perhaps more general related to an area rather than much more specific sort of honed down into you know that paddock level ... (Interview participant, Time 2).</i></p>

Industry	Farmer data	<p><i>... like we're sort of conscious of, I guess, or mindful, of how we how we use the farmers' data. We're seen as like a trustworthy organisation where we're storing their data securely and not giving it out willy nilly to anyone so. Yeah, it's sort of, you know, that's written down in the data sharing agreement, they sign it saying, you know, sign an agreement saying we will store their data and we won't give it out or we only use it on a regional scale, blah blah blah. But yeah, sort of, I guess we don't want to be seen that we are going to give out data or sell their data ... you want to keep the trust with the farmers because you want them to keep actually testing soils. Yeah. I guess if we approach them all the time and I ask if we can use the data for here or there and then they'll sort of get Jack of it and won't do any future testing through us, I suppose, so that's a big barrier (Interview participant, Time 1).</i></p> <p><i>So if the VAS project is asking growers to hand, oh, or to a certain degree submit soils information off their farm, there has to be an economic return coming back to the grower for them to be able to justify it because at \$150 a soil sample then, and if they've been on that farm for 10 years, and maybe if they've sampled 5 paddocks every year, that's a significant investment that has contributed to their decision making on farm. But if you ask a funding body to hand over data that they have invested in, they're going to want to know what the financial return is going to be from them contributing that that investment or sharing that that investment with someone else. So that's really something that I think the project needs to nut this out really well before they try and push too hard down the line of contributing farmer data (Interview participant, Time 1).</i></p> <p><i>...there's always a level of trust that farmers will put into things. So, if there's that trust is there, they'll probably do it. But yeah, there are a lot of concerns around data privacy with farmers at the moment. You know, they've always been pretty open about sharing things, but I think there is a bit more of a realisation that maybe they should put a bit more value on what they do put out there (Interview participant, Time 2).</i></p>
	Time constraints at the organisational level	<p><i>I think the only limiting factor will be our time to actually organise ourselves to input data into it. So I think if we if we can make a conscious effort to get it in there then it'll be fine, but it's just at the minute, just capacity, within our own organization (Interview participant, Time 2).</i></p>
Communication	Assistance with using VAS and uploading data	<p><i>... there will always be a need for someone to be able to call to be able to help organise the data if it is in his self-service platform, especially as we move to an automated system (Interview participant, Time 1).</i></p> <p><i>...so, I couldn't fill out the data sheet quite that easily and I think they got lots of feedback. So, then they made changes to it and they've sent it back to us, but there is... it looks like it might contain like 10 or 12 tabs of information that you need to include. So, it's just not something that you can do quickly. You kind of need to be concentrating on it, I think, especially being the first time to use it so we've not really had the time to do that and work it out and I think in terms of a time saving for us is that we don't muck around trying to do it ourselves and work it out, that we just try and do it up there at FedUni and if we get stuck well, there's someone there to help us (Interview participant, Time 1).</i></p> <p><i>Maybe we need to put the thinking in as to who we want to give permissions to, to access it, whether it is at a public level or just a group level or, you know, umm and I guess maybe some more information around what each of those different access levels, umm what it means and yeah (Interview participant, Time 2).</i></p>

	Project communications, updates, and promotion	<p><i>I think one of the challenges around COVID that was just kind of continuity of, you know, keeping in touch and kind of meetings and things like that which we didn't feel so much at our end... and that which kind of interrupted the regularity of communications between the project team and us (Interview participant, Time 1).</i></p> <p><i>I mean, when I did a Google search for it (VAS) before, I knew what to search for. I found it relatively easy, and now I've got it saved as one of my favourites to be able to access, but I don't know ... If you Googled soil information or anything, whether it would necessarily come up (Interview participant, Time 1).</i></p> <p><i>... have examples of what's being developed, what's new (with VAS). What are people using it for and then probably some face to face like at the start of the project. Everyone was very strong on face to face and then when COVID came down, it's not. And yeah, it made such a difference that the recent conference in Adelaide, when you can actually see people and have a proper conversation... (Interview participant, Time 1).</i></p>
	Managing expectations about the data	<p><i>... I've sat in on a few of Pete's public webinars where he's gone, through the Soil CRC, giving a project update and has highlighted certain groups that he's worked with, where he said we've uploaded 100 sets of data for this particular group. And I'm sitting there going 'How did that happen? How did they get 100 sets of data uploaded and I'm sitting here with twelve sets of data in a queue waiting?' so I don't think that's been communicated very well ... (Interview participant, Time 1).</i></p>