Data Literacy Training and Dashboard Requirements Gathering

Prepared by Jenny House

December 2023

Table of Contents

1   Introduction

This report outlines the activities and findings of the Data Literacy and Dashboard Requirements Gathering Process conducted in August – December 2023. A Data Literacy and Requirements Gathering Process Manual has also been prepared, which contains a more detailed explanation of the training content, facilitation approach, and explanation of how to repeat this process in the future. This document presents the overall objective and rationale behind the approach, documents the activities conducted, describes the requirements gathered thus far, and makes some recommendations for use of this methodology by Similie in the future, in order to maximise the value provided by this Module.

The Data Literacy and Dashboard Requirements Gathering Module was developed in response to the challenge of determining what elements should be included in a Water Resource Management (WRM) Dashboard. Similie's people-centred approach requires that the development of this dashboard should be driven by the needs and aspirations of the platform users. However, there are a number of prerequisites needed in order for users to identify and articulate their needs and wants from such a dashboard. Therefore, though the ultimate goal was to produce a list of users' requirements for the WRM Dashboard, the approach used in this Module began with some fundamentals of data literacy and working with data visualisations, as well as guided discussions on different aspects of the data pipeline and how the data will be used. It is hoped that this holistic approach to data literacy, equips participants with the necessary skills to understand the purpose of the WRM Dashboard and how it may assist them with their objectives and activities. Furthermore, the Module focussed on building confidence and communication skills around the participants' "data stories", in an effort to support them with understanding what data they need, how it connects to other aspects of their work, and how to think, understand, and talk about data.

As part of this project, the Data Analyst Consultant conducted the Data Literacy portion of the Module with Similie team members. Subsequently, the Module (including dashboard requirements gathering) was delivered by the Water Lead and Customer Success Team, with support from the Consultant, to members of PERMATIL. The Module was collaboratively developed by the facilitation team and was tailored to the needs of the various participants. Given the interactive and flexible nature of the Module, it can be adapted to any audience or level in the future. This report outlines the activities and outputs of this project, as per Service Agreement SA2023014, as well as the findings of the requirements gathering process so far.

2   Activities completed

The Data Literacy and Dashboard Requirements Gathering Process was developed collaboratively by the Consultant and Similie staff, via an iterative process. As briefly outlined in Table 1, the Data Literacy & Visualisation Workshops were conducted with Similie staff in September. This experience was used to inform the material for the Module used with platform users, and highlighted existing issues and requirements for the dashboard, based on previous conversations with users or Similie staff's own experience and knowledge. In addition, the purpose of these workshops was to develop the team's own data literacy skills, as well as their ability to have confident and meaningful discussions about data and how it can be used.

Table 1: Workshops conducted, and their objectives

	Module Activities	Participants	Objectives	Lead Facilitator
14/15 Sept	Data Literacy & Visualisations Workshops	Similie (~11 participants)	1) Strengthen data literacy skills and ability to have data conversations with users 2) Develop understanding of what makes a "good" dashboard (for users & Similie) *3) Gather Module feedback and input for user version*	Consultant
20/21 Nov	Data Literacy & Visualisations Workshops	PERMATIL (4 participants)	1) Understand the importance of data literacy and how it can support other work 2) Explore links between research objectives, parameters, data, and practical application or decision-making 3) Understand basics of data and graph interpretation and use *4) Preliminary dashboard requirements gathering*	Water Lead
24 Nov	Dashboard Requirements Gathering Follow-up Workshop	PERMATIL (3 participants)	1) Synthesise and validate requirements gathered already 2) Solicit feedback on draft ideas and responses to queries about dashboard content/appearance	Water Lead

To deliver effective people-centred workshops on data literacy and requirements gathering in the future, facilitators need to develop their own data literacy skills to a sufficient level, as well their ability to understand and convey the importance of non-technical skills such as communication and critical thinking. Though these workshops successfully covered the basics of data literacy and how to support others in their data literacy training, additional practice and skill development is needed to consolidate these skills and knowledge. Table 2 provides an overview of the workshop content.

Table 2: Overview of workshop content (see Module materials for details)

Workshop	Session Content	Requirements gathering
1 **Working with data**	Introducing data literacy and the skills involved Selecting parameters based on research questions or objectives, by identifying the concepts and variables which need to be understood What makes a good indicator? Proxy indicators & measuring things indirectly	Discussion on what the participants want from the data, what it can be used for, what their objectives are, and what parameters may be needed to address these
2 **Interpreting data**	The journey from raw data to interpreting data Introduction to basic descriptive statistics and why they may be useful Overview of different types of graphs, how to read them, and for which kinds of data they should be used Identifying trends and correlations, the difference between correlation and causation	Discussed examples that are similar to those currently used by Similie as well as other sources, to identify what characteristics are liked or disliked, and what makes a good graph
3 **Understanding and telling data stories**	Introduced the idea of the data story – connecting the objectives/questions to the data, and the data to action or decision-making, and being able to communicate and think those links through clearly Based on practice scenarios, thinking through what could be going on in the data in a graph, and what that might mean in the real world. Dealing with unexpected results	A real example was used from the participants' work, to practise identifying the data and relationships which need to be conveyed in the dashboard (building on Session 1-2) Discussion on what data stories participants currently work with, or would like to be able to tell in the future
4 **Working with dashboards**	Evaluated dashboards from other sources in the WASH sector and practised using them. Discussed what makes a good dashboard and designed a mock-up version (see appendix for results)	Identified liked and disliked characteristics of a dashboard and began selecting desirable graphs or concepts
5 **Dashboard requirements gathering follow-up**	Review of Sessions 1-4 Explanation of WRM Dashboard Scope Presentation of parameters identified, and synthesised into themes Discussion of possible analysis and visualisations Discussion about dashboard features	Presented results from previous findings, gained feedback, and discussed other dashboard features Internal requirements gathering process is recommended to ensure processes with stakeholders are effective

All activities and deliverables were completed, but unfortunately only PERMATIL members were available to participate in the workshops held in November. Therefore, it is suggested that the Module and requirements gathering process be adapted for use with the Government of Timor-Leste, WaterAid or other relevant stakeholders as part of the dashboard development process.

2.1   Requirements gathering process

The requirements gathering process was integrated throughout four Data Literacy workshops and was completed with a follow-up session a few days later (see Module Manual for details of session content). This was achieved by incorporating group discussion and practice exercises throughout the workshops, alongside thorough notetaking by the facilitation team. For example, the culmination of much of this discussion with the PERMATIL team was the development of a "data story" which incorporated some of their main objectives, the link with the water monitoring data, data sources of their own, and the parameters, variables, or relationships of interest for the WRM Dashboard (Figure 1). In between the initial Data Literacy Workshops and the Requirements Gathering Follow-up Workshop, all notes and discussion outputs were collated and synthesised to produce a list of parameters that fall within the scope of the WRM Dashboard, and which meet the needs of the PERMATIL participants. Several draft graphs were also developed to seek feedback. Participants were also asked about relevant topics identified by the Similie team, such as the data to be included in the dashboard (parameters and level of resolution), appearance, user experience, alerts and communication methods, and maintenance data.
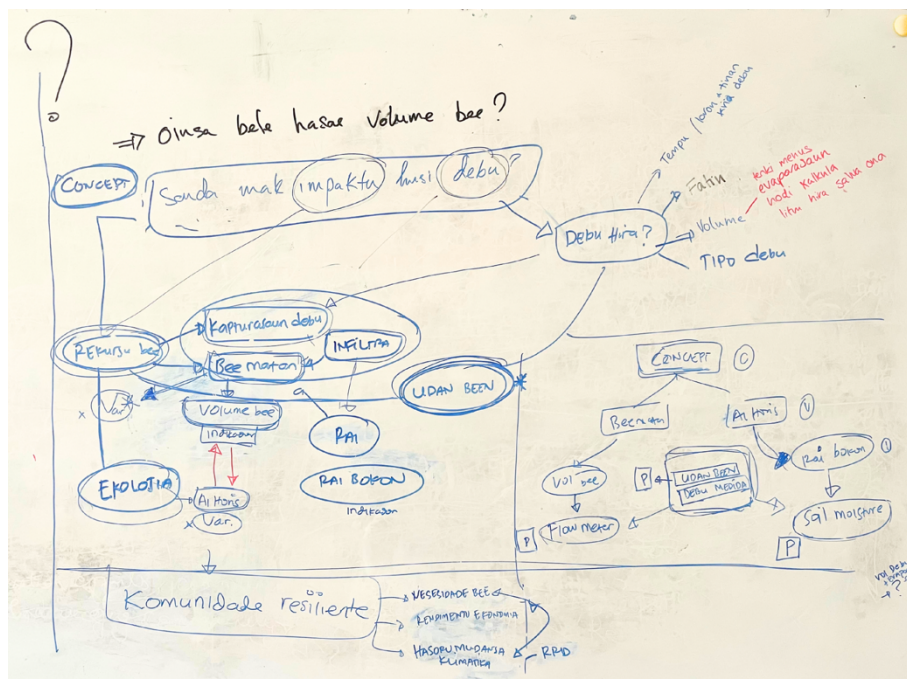

Figure 1: Whiteboard of the data story connecting PERMATIL activities to key variables and parameters from the WRM Dashboard

2.2   Project Outputs

The following materials were produced during this project and can be adapted for use with other audiences for future Data Literacy Training or Requirements Gathering Processes:

- **Data Literacy and Dashboard Requirements Gathering Manual**
  - This Manual includes details of the session content, facilitation guidance, and information on how to adapt the materials to a new audience.
  - In terms of facilitation, the manual includes a Facilitation Agenda, suggestions on facilitation approaches for the different activities, note-taking templates, and steps for preparing for the follow-up session (Session 5).

- **Slide deck & facilitation notes** (Tetun slides & English facilitation notes as requested)
  - Session 1-4 focus on Data Literacy and Visualisations. These should be adapted based on the level and work focus of the participants or the relevant Similie domain in question.
  - Session 5 is the Requirements Gathering presentation and must be re-done based on the findings of the preliminary requirements gathering in the previous sessions. See Manual for guidance on how to do this.

- **Data Literacy Training Overview** (for sharing with participants)

- **Worksheets & activity materials for printing** (English & Tetun)

3   Dashboard requirements gathered

Through the requirements gathering process with PERMATIL, there was much discussion about PERMATIL's objectives and how Similie data may assist with this. The PERMATIL members identified four main priorities or uses for data in their work; 1) to validate their efforts, observe results and evaluate their work or impact; 2) to determine the best mechanism/approach for their work, 3) to influence government and donors, and 4) to communicate results back to the community and other stakeholders. There was specific discussion and interest in how such data may be combined with PERMATIL's own M&E data or other data sources to validate their water conservation activities. Though the WRM Dashboard is related to this topic, this specific objective is outside the scope of the WRM Dashboard. Therefore, the requirements included in this report are limited to the WRM scope, but the notes from the process may inform future projects which are more focused on M&E of the water conservation. The WRM Dashboard covers one part of PERMATIL's data story, and they will be able to combine this with their own data and analysis in the future.

3.1 Parameter selection

PERMATIL identified useful parameters which fit into four themes (Figure 2): water resources (rain, springs, soil moisture), climate (rain, temperature), water use/availability, and contextual information (place, time on various scales). The water resources theme includes water in rain, springs, and soil moisture, but the water use/availability theme incorporates the human and water supply system aspect of the water availability. Consequently, this theme includes the water present in the tank, how much is being used and who is using it. The participants described this part of the data story as understanding if the community has access to the water that they need and how much is being used. Though the PERMATIL participants were more focused on water in the springs and soil overall, the water supply tanks represent an important link between these water resources and the use of the water, which is clearly linked to their research questions and conservation objectives. The WRM Dashboard should play a valuable role in this data story by bringing these different aspects together, to allow a more holistic understanding of the data and water resources overall, the relationships between these themes, and, thus, the significance and usefulness of the monitoring results.
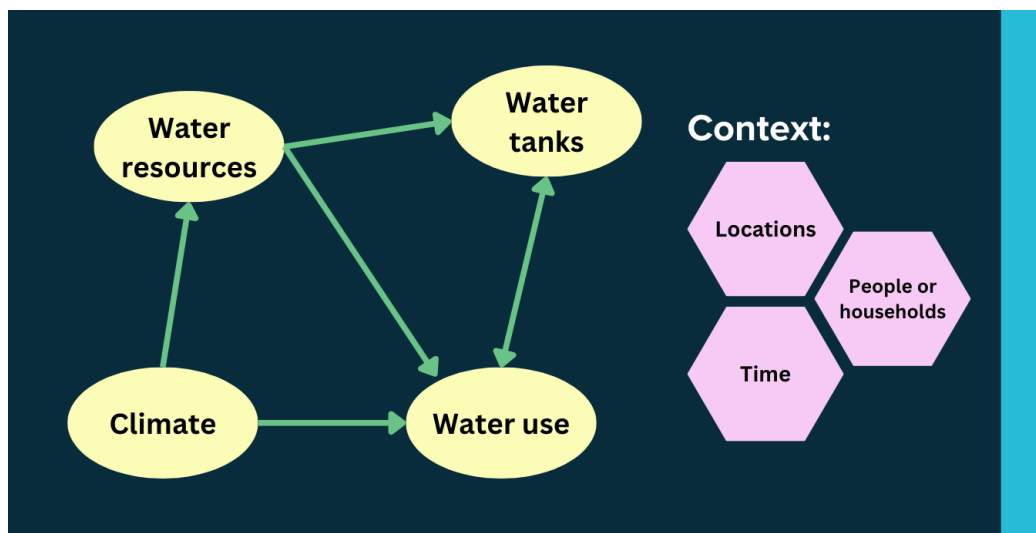

Figure 2: Summary of key themes and relationships identified by PERMATIL for the WRM Dashboard

Several parameters were selected as the most relevant and interesting for the dashboard. These were selected from the available parameters in Similie's platform, as well as from the data stories and group discussions. Table 3 presents a list of the parameters selected for the WRM Dashboard by PERMATIL, and does not include the parameters which relate to other aspects of the water conservation activities. Data relating to the number of people or households in an area or accessing a particular water tank is a crucial component in calculating relevant indicators and interpreting data, so such contextual information is included in this list of parameters.

Table 3: Parameter list gathered from PERMATIL in Session 5

Parameter (theme)	Units	Scale/Resolution	Comments
Rainfall (water resources) (climate)	mm	Daily, Monthly, Annual	Need to compare between years and see rain throughout the year (i.e., seasonality) Would also like to calculate daily and monthly mean
Spring Flow Rate (water resources)	l/s (see comment)	Daily, Monthly, Annual	Need to be able to compare between and within years, a monthly average would also be helpful Units should match field data and be consistent with other parameters
Soil moisture (water resources)	%	Automatic Weather Stations (i.e., catchment areas) By retention ponds	For PERMATIL, the soil moisture near the retention pond locations is important for evaluating and directing their conservation efforts. This would be a useful future step for PERMATIL's work.
Wind speed (climate)	*m/s*		*This variable was discussed but it was decided that it is not a high priority for the dashboard.*
Temperature (climate)	°C	Daily, Monthly, Annual	
Water tank level (water use/availability)	%/cm/l		It is suggested that graphs may show water tank level as cm, litres, or as a percentage, and that the absolute measurement (cm or litres) and the percentage could be shown simultaneously by adjusting the scale and units on the Y-axis so that only one line is needed on the graph to show the data with any units (this comment relates to any graphs on the dashboard or platform that show tank level, as it was noted that the two lines on the default platform graph could be streamlined into a single line by adjusting the axis labels and units). Information on the maximum tank capacity should be included. Participants stated that it is useful to see whether water is going into the tank. Therefore, tank level (or inflow) should be included in the dashboard, as well as the calculated water use (below).
Water use (water use/availability)	L/HH/D L/C/D L/Tank/D	At each water tank	Water use can be indirectly calculated using change in the water tank level (though this includes leakage). This could also be measured using the outflow of the tank. For the dashboard, this calculation would assist with interpreting water supply information. Though water tank level is useful on its own, calculating daily water use using tank level fluctuations would assist users who are not as focused on the detailed functioning and use of the water tank system. Given the lack of accurate population/household data, it is suggested that daily water use could be presented as a total per tank (ensuring that data quality can be controlled by Similie). The water use per household or per capita can also be calculated as this is preferred for end users. Between "per household" or "per person", the recommendation from the discussion is to use per household but state the average number of people per household.
Number of water users (water use/availability)		Households Population	See note on data quality above
Time (context)		Most variables on daily, monthly, and yearly scales.	Overall, the time scales which are most relevant are changes between and within years (i.e., seasonality) Need to compare before and after water conservation
Place (context)		Location	Currently, the location of measurement is available. However, in the future this could include GIS or analysis at the *postu* or *suco* level (e.g., mean daily water use per person in a municipality)

3.2   Data visualisation

Through the group discussion, several ideas for graphs were identified as useful by the participants:

- Change in rainfall and soil moisture over time e.g., Figure 3
  - Given the relationship between rainfall and soil moisture, presenting these variables alongside each other would assist with interpreting the results, i.e., the way in which soil moisture is influenced by the rain may vary depending on the local conditions or conservation activities etc.

- Rainfall and spring flow/volume over time (same format as Figure 3 and for longer time series, e.g., Figure 4), in each place.
  - There is interest in identifying the impacts of climate change on water resources and seeing these alongside each other over a longer time period and with changes in seasonality would be helpful.

- The impact that spring volume has on water in the water tanks, and how this relates to water use in each place.

- The relationship between temperature and rain in each place.

- The relationship between rain, soil moisture, and spring flow.
  - There was interest in the relationship this has with ecological parameters, such as how soil moisture may influence the effectiveness of reforestation. The extent to which the soil remains moist after rain is important for vegetation.

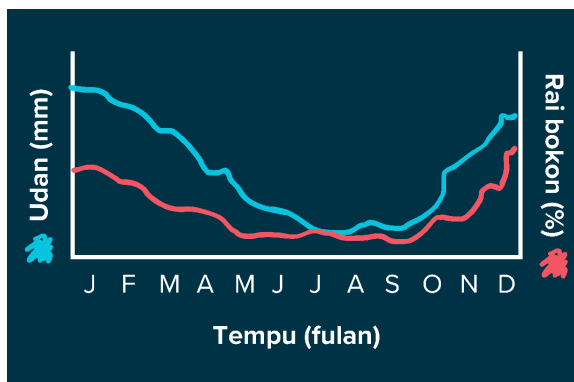- Average daily water use (L/HH/D) for each month of the year

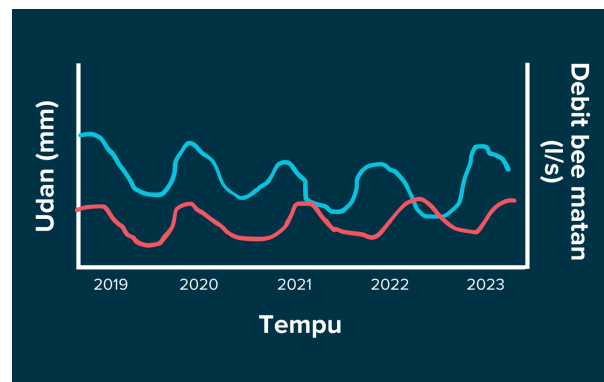*Figure 3: Rainfall and soil moisture throughout the year*   *Figure 4: Rainfall and spring volume over time*

The preferred type of data visualisation is change in particular parameters over time. When asked about how change in variables over time should be shown, participants were shown Figure 5 and Figure 6 as options. Participants stated that Figure 6 is preferred due to the ease of comparing the same time of year, but that if both could be included then that would be good.
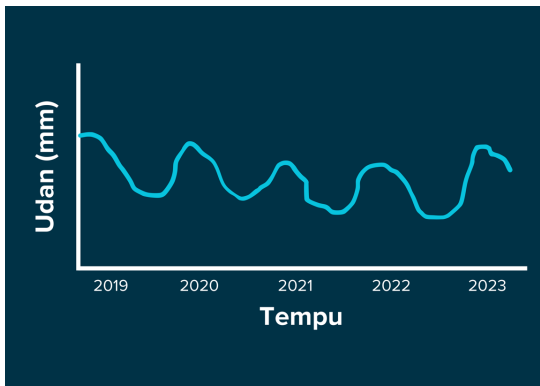
Figure 5: Rainfall as a continuous line, enabling identification of longer trends.
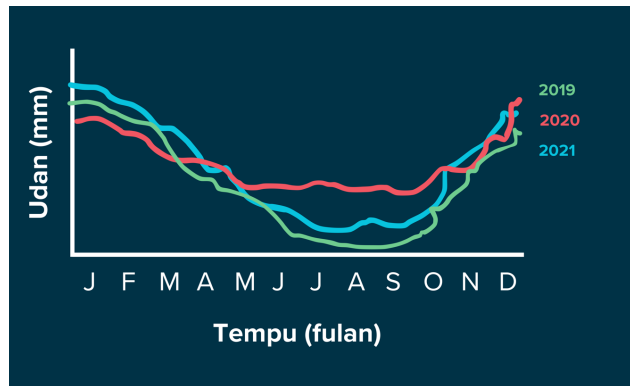

Figure 6: Rainfall with each year as a line, enabling comparison of seasonality

In addition to showing continuous data, the option of presenting monthly averages was also discussed, as per the example in Figure 7. Participants found this helpful and intuitive and suggested that it may be useful for showing available water vs water use in each month. Some ideas for comparing between places were discussed, such as Figure 8, but it was thought that this may not be the most useful for the dashboard at this stage. PERMATIL would like to be able to make such comparisons for the locations where they have conservation activities or to compare the control site with a location that has conservation in place. The option to compare locations would be useful in the future.
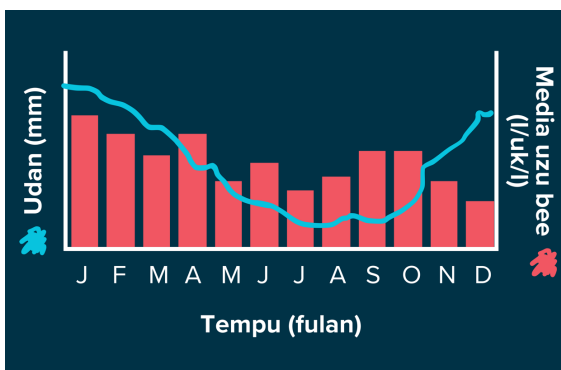

Figure 7: Monthly averages can be combined with line graphs of continuous variables
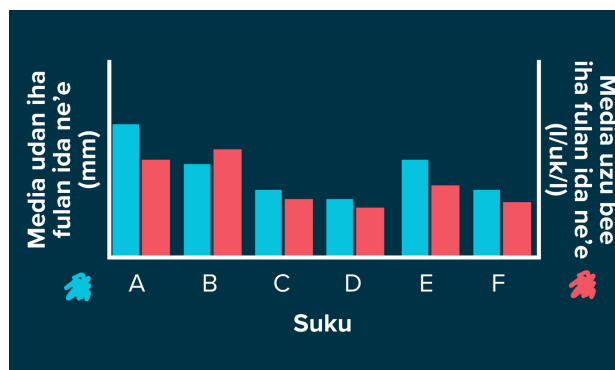

Figure 8: Average rainfall and water use for the month in different places – an example of location comparison

Other forms of graphs, such as box plots were not selected. Though scatter plots were easy to understand and work with, it was decided that being able to see change over time is most useful, so line graphs and bar charts are preferred. Some draft visualisations and dashboard layouts are presented in the Appendix, though some of these are not within the scope of the WRM Dashboard.

The use of black lines and colours make it easy to read graphs. Writing should be black and using standard fonts such as Times New Roman. Text should be orientated to make reading the results easy. Sufficient labels are required (titles and axes) to enable reading of graphs without hovering the mouse. Including comprehensive titles, axes titles, and scales is needed to improve readability and ease and accuracy of data interpretation. Consistent time scales should be used for all time series data, rather than data collection intervals (this is relevant for dashboards and elsewhere in the platform). Furthermore, such features would make the dashboard more accessible to users with less internet access or computer skills.

3.3   User Experience

Some additional features were discussed. Firstly, providing some contextual information on the dashboard as text may be useful for readers. It was suggested that this could explain the data source and location, geographic details, local authorities, and agencies working in the location. Though there was general discussion about this, the general consensus was that the participants would be happy with general information selected by Similie or other users.

Secondly, during the Similie training the idea of adding a "help me" button was discussed. When asked, PERMATIL members thought this may be a useful user guide, especially if the dashboard has interactive features. This guide would be a simple guide to orientate the reader, rather than a guide about how to seek specific data.

The level of interactivity required was also discussed. It was suggested that a basic dashboard is needed which can be used "as is" for the main results. However, having the option to select variables or scales may also be useful. For example, the basic graphs would be there by default, but users could then select the variables of interest for their needs for each graph.

The dashboard must be able to be used offline so that it can be used in locations with poor internet access. It was suggested that a button to save a screenshot or pdf of the dashboard may be useful, so that users can save the latest version before travelling to the field. Therefore, all information needed to read and interpret the graphs must be available without interaction. For example, information provided by hovering the mouse over a data point can be a useful feature, but it should not be required to read the graph.

In terms of maintenance data (battery etc), the participants suggested that the status of the sensor should be included in the dashboard so that users can see if data is currently being collected (similar to the functionality status for water tanks in the water supply platform). However, no additional details is needed in the dashboard because it is still available in the platform.

The use of alerts and messages in the dashboard were also discussed. Alerts are urgent messages which are delivered by email or SMS. Participants suggested that relevant alerts for the WRM system could be about heavy rain (to alert people of risks such as floods or landslides) or very dry conditions (to warn about fire risks of burning, cigarettes etc). These are similar to the alerts provided by the Early Warning System domain. In contrast, status messages may be used in the dashboard for messages where are less time sensitive. Ideas discussed included identifying when conditions are very dry or rain is coming late (so consider delaying agricultural activities or avoiding slash & burn), when there is high precipitation, or when soil moisture or spring volume are particularly high or low. Such messages can also inform people when they should be conserving water. It was noted that the nature of these messages would vary depending on the audience. For examples, if only PERMATIL and other users access the dashboard then these messages should be tailored to the needs of organisations. Otherwise, if communities may access the dashboard directly, then different guidance around behaviour change may be more useful.

4   Conclusion and recommendations

In conclusion, the integrated approach used in this Data Literacy Training and Requirements Gathering Process has enabled a more meaningful conversation about the priorities and needs of the dashboard users. The majority of the findings are well aligned with the available data in the Similie platform, making a relevant and user-friendly dashboard very achievable. The extensive discussion and feedback process should ensure that the participants are able to read and understand the WRM Dashboard and will be more equipped to use such a resource in their work. The ability of users to understand and apply monitoring results may determine whether the data can have an impact. Therefore, user friendliness is paramount in the dashboard design. The focus for the dashboard design should be to produce something which is clear, intuitive, and easy to read, even if that means that some data are not featured.

The Data Story approach, represented by Figure 9, should be the foundation of the dashboard design and future requirements gathering processes. This method uses research objectives and an action-oriented approach to inform the parameter selection and data visualisation process. In addition to the requirements gathered through this process, this project has shown the importance of developing data literacy skills for platform users and Similie staff. Understanding data stories and improving data literacy skills may assist with overcoming silos between teams, where people in particular roles are focused on their part of the data pipeline and may not understand other perspectives. For example, data literacy and storytelling may help field staff, or those who do not usually work with data, to understand their colleagues and their organisational approach more effectively.
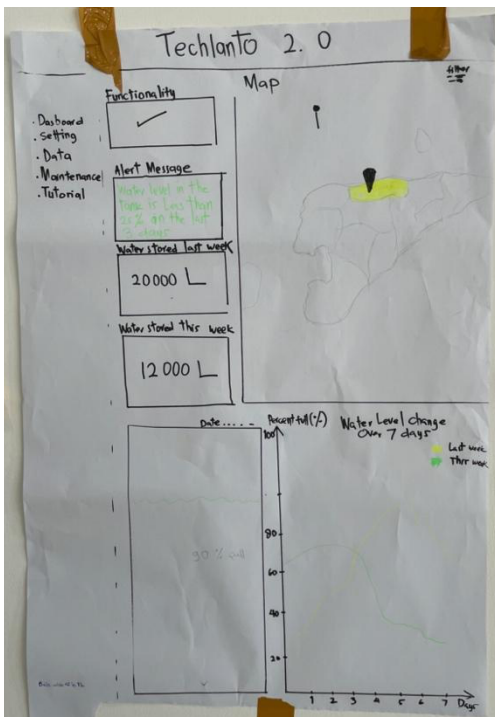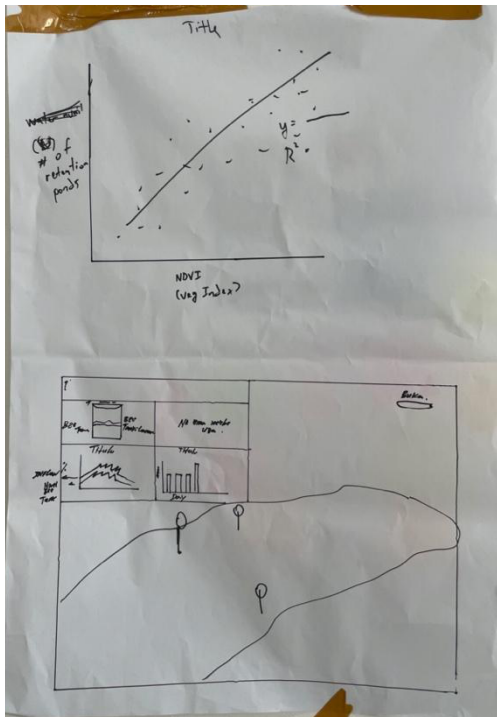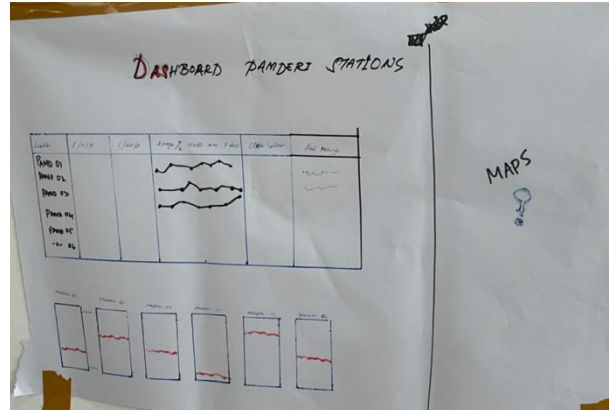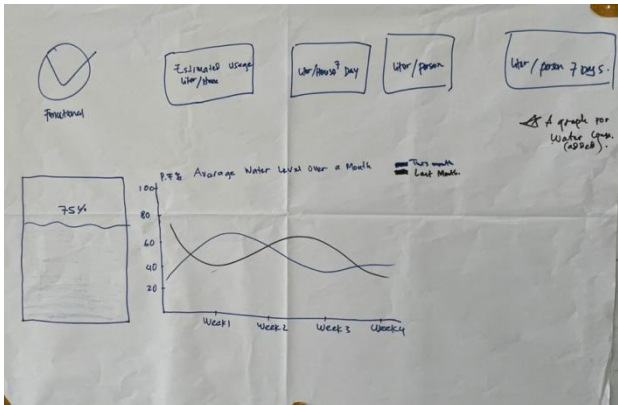
This project has successfully gathered requirements from representatives of PERMATIL. However, it is recommended that an internal requirements gathering process take place, to capture the experience and needs of Similie staff who already have extensive knowledge about the data and user needs. The Module and requirements gathering should also be repeated with other stakeholders who may have different needs and wants from PERMATIL. This approach can also be used when working with new partners or at the beginning of new projects as part of a consultation process. Developing a greater culture of data literacy will provide many benefits to Similie and partner organisations.

Figure 9: Data stories begin with the broad question or objective, before focussing in on the details, and then linking back to real world application

Appendix: Dashboard Mock-ups

Similie Workshop

PERMATIL Workshop