



British
Geological
Survey



Updating the SADC Grey Literature Archive Final Report



British
Geological
Survey



Updating the SADC Grey Literature Archive

Final Report

Colin Blackburn, Simon Burden, Ken Lawrie, Martin Nayembil, Andrew Simister, Raquel Sousa, Arnaud Sterckx, Kirsty Upton

BGS Commercial Report CR/20/066

October 2020

Bibliographic Reference:

Upton, K., Nayembil, M., Burden, S., Lawrie, K., Blackburn, C., Sterckx, A., Sousa, R. & Simister, A. 2020. Updating the SADC Grey Literature Archive: Final Report. British Geological Survey Commercial Report, CR/20/066. 90pp.



British
Geological
Survey



Foreword

This report is the final product from a commercial project by the British Geological Survey (BGS) and International Groundwater Resource Assessment Centre (IGRAC), on behalf of the client, the Southern African Development Community Groundwater Management Institute (SADC-GMI). It summarises the main activities and outputs from the project – primarily the development of a new groundwater literature archive for the SADC region – the SADC-GLA. The project was led by BGS. Development of the new database was led by Martin Nayembil, Ken Lawrie and Andrew Simister; development of the application programming interface and user interface was led by Patrick Bell, Colin Blackburn and Simon Burden; the project was managed by Kirsty Upton. New data was compiled by Arnaud Sterckx and Raquel Sousa from IGRAC, in collaboration with Young Professionals from across the region.

Acknowledgements

In addition to the BGS and IGRAC staff acknowledged in the Foreword, the project team would like to acknowledge the input of SADC-GMI staff, particularly project manager James Manda. The following 25 Young Professionals from across the SADC region played a crucial role in the success of this project by collecting new data from their respective Member States: Avelino Chilete Cacoma, Katia Antonio Melo, Ndineni Ngayaya, Keneilwe Kenny Thibelang, Mahamoud Ahamada, Antoissi Ahamada, Sean Ginindza, Thandeka Ngobe, Kabelo Setalan, Motlatsi Pheko, Patrick Mlomba, Vitumbiko Mwausegha, Omar Sirage, Nelson Malikito, Reginalda Joseph, Emanuel Kahuva, Ngumu Bonaje, Eliofoo Hamis Hango, Qunitne Chibesa, Moonga Himulonga, Shephard Shereni, Percy Mugwangwavari, Devisha Ittoosingh, Lindelani Lalumbe and Thandilizwe Bengeza. The team would also like to acknowledge Alan MacDonald (BGS) for his advisory role in the project, Carl Watson (BGS) for his input to the user consultation process, and Antonio Ferreira (BGS), Ahmed Zeggan and Claudinei Nunes da Silva for their translation services.



British
Geological
Survey



Contents

Foreword.....	3
Acknowledgements.....	3
1 Introduction.....	5
2 Background.....	5
3 Project Overview.....	5
4 Project Structure, Milestones & Deliverables.....	5
5 Summary of user consultation (WP1).....	7
6 Summary of data and system review (WP1)	7
7 Summary of system update (WP2).....	8
8 Summary of data update (WP2).....	8
9 Summary of training (WP3).....	9
10 Long-term strategy (WP3)	9
11 Reflections and Recommendations	10
12 Project team & Contact Persons.....	11
Annex 1: Project Kick-off Meeting Summary & Minutes.....	12
Annex 2: Situational Analysis Report	19
Annex 3: Project Update	41
Annex 4: System Documentation	49
Annex 5: Report on data collection.....	66
Annex 6: Training Summary.....	69
Annex 7: Governance Strategy	77

FIGURES

Figure 1 SADC-GLA Architecture Overview



British
Geological
Survey



1 Introduction

The Southern African Development Community Groundwater Management Institute (SADC-GMI) is a regional centre of excellence for groundwater management, whose vision is to ensure the equitable and sustainable use and protection of groundwater in the SADC region. In 2019, SADC-GMI initiated a project to update the SADC Grey Literature Archive, to make it a fully functional online archive of groundwater literature for use by all groundwater stakeholders in the SADC region. The project started in February 2020 and will finish in October 2020. It was implemented by the British Geological Survey (BGS), in collaboration with the International Groundwater Resources Assessment Centre (IGRAC).

This report provides a summary of the project process and outputs.

2 Background

In 2010, BGS developed the SADC Grey Literature Archive, an online catalogue with almost 3000 unpublished reports, which were searchable through a web portal by title, author and country. In 2014, BGS began development of the [Africa Groundwater Literature Archive \(AGLA\)](#), which built on the SADC-GLA, providing improved functionality and almost 1000 additional documents for the SADC region. Since 2017, IGRAC and SADC-GMI have developed the SADC Groundwater Information Portal (SADC-GIP), a web portal for sharing groundwater related data and information in the SADC region. The SADC-GIP is being improved by IGRAC as part of a parallel project with SADC-GMI.

3 Project Overview

The key aim of this project was to provide access to groundwater literature in the SADC region to all relevant stakeholders, to support the equitable and sustainable use and management of groundwater in this region. To achieve this, the project team have developed a new system – the SADC Groundwater Literature Archive (SADC-GLA) – which can be easily updated and searched, is technologically advanced, and contains up to date information from stakeholders across the SADC region. The SADC-GLA can be accessed through a new web portal:

www.sadc-gla.org

The SADC-GLA consists of a new database, containing around 4,500 references to groundwater literature for the SADC region. Many of these are available to users, either as downloadable PDFs or through links to external websites. The database is accessed by the web portal (address above) through an API, providing interoperability and flexibility for future development and hosting. Users in the region have been introduced to the new system through an online training workshop, and a long-term governance strategy for the SADC-GLA has been developed and initiated with SADC-GMI.

4 Project Structure, Milestones & Deliverables

The project was structured into 3 work packages:

1. User consultation and review of the current data and system (WP1)

2. Development and implementation of a new system to meet the needs identified in WP1 within the time and budget constraints of the project (WP2)
3. Training of users in the SADC region and development of a long-term governance strategy to ensure the system can be updated and maintained beyond the end of project (WP3)

WP1 was carried out during February and early March 2020. A kick-off meeting was held on 5th February, attended by representatives of SADC-GMI and the BGS project team. At this meeting, the scope of the project was discussed and agreed, including: (1) identification of the main user groups of the SADC-GLA (represented by SADC-GMI for the purposes of user consultation); (2) confirmation that BGS would be required to host the system in the short- to medium-term, with a strategy developed for long-term governance of the system; (3) agreement on a terms of reference for the Young Professionals in the region, who would be sub-contracted to collect new data for the SADC-GLA, under the management of IGRAC. A meeting summary and minutes can be found in Annex 1.

The user consultation process, summarised in Section 5 below, was implemented through a workshop on 14th February, which was attended by SADC-GMI, IGRAC, and the BGS project team. In parallel, a review of the existing AGLA system and SADC data was undertaken, as summarised in Section 6. AGLA was reviewed instead of the previous SADC Grey Literature Archive as this was a more up-to-date system in terms of technical design and content. The results from WP1 were detailed, along with recommendations for the design of the new SADC-GLA, in a Situational Analysis Report, delivered on 11th March and provided in Annex 2.

WP2 kicked-off in late March – slightly delayed due to the developing global pandemic – with a 2-day internal workshop, attended by the BGS project team and IGRAC. The purpose of this workshop was to: (1) agree final decisions on the design of the new SADC-GLA system, including the database, API and web portal; (2) define the specific tasks required, timeline and responsibilities within the team; (3) agree on key aspects for the governance model. The outcomes of this project workshop were summarised in a project update report, delivered on 31st March and provided in Annex 3.

Development of the new system continued through April and May with the demo system going live for review by SADC-GMI on 18th May. Comments from this review were addressed and the first version of the new SADC-GLA system, as described in Section 7, was launched on 30th June, alongside delivery of the technical documentation (provided in Annex 4).

In parallel to development of the new system, IGRAC led the data collection activity, which has provided nearly 500 new documents to the archive, as summarised in Section 8 and detailed in Annex 5.

WP3 involved the delivery of a two-day online training workshop to stakeholders from across the region, including SADC-GMI staff. The training workshop involved use of the system by a wider range of people than had previously been involved in testing, providing valuable feedback that was addressed over the following weeks. This is summarised in Section 9, with a full report in Annex 6.

A governance strategy was finalised in September, outlining how the system would be hosted, maintained and updated in future. This has led to a maintenance contract being established between SADC-GMI and BGS for a two-year period starting in November 2020, after which

the system will be handed over to SADC-GMI. This is summarised in Section 10 and the full governance strategy provided in Annex 7.

The project was initially due to finish in June 2020, but due to the global pandemic, a three-month contract extension was granted, largely due to difficulties experienced by the Young Professionals in accessing and gathering new data. The restrictions to global travel also caused the training activities to be revised, with a shorter online training run as part of this project and a longer in-country training planned as part of the final handover of the system at the end of the two-year maintenance contract.

5 Summary of user consultation (WP1)

Initial discussions on the user consultation process took place during the project kick-off meeting on 5th February (see Annex 1), at which it was agreed that SADC-GMI, IGRAC and BGS project partners could fulfil the roles of website administrators as well as proxies for real world end users of the SADC-GLA. A User Research Workshop took place on 14th February via skype, attended by BGS, IGRAC and SADC-GMI. Miro – an online whiteboard for visual collaboration – was shared prior to the workshop and used during the workshop to discuss different target audiences, the user journey (using AGLA as a template), administration and site monitoring, and IPR issues.

Analysis of the Miro board after the workshop defined key user requirements, which were used to inform the design of the new system, as outlined in the Situational Analysis Report (Annex 2). A summary of these requirements is provided below:

- The web portal should be easy to find via common search engines
- The search functionality should be simple and intuitive
- It should be easy to filter for entries in the database where a document is available
- As many documents as possible should be available/accessible
- All documents should be discoverable by a range of search criteria
- There should be a well-defined and straightforward process for updating or adding new documents to the archive
- The system should be interoperable with other document databases

6 Summary of data and system review (WP1)

A comprehensive review of the AGLA system, including existing data for the SADC region, technical implementation, functionality, and presentation was also undertaken as part of WP1. The full details of this review are provided in the Situational Analysis Report (Annex 2). Some of the key findings and recommendations of this review were:

- There were a significant number of entries in AGLA without a downloadable PDF or link to external site; other metadata was also missing
 - All existing data requires a general QA, the database should be standardized, and more dictionary constraints added to provide better control of metadata
- The keyword dictionary was highly granular and the most common search terms used were country names and very general water- and geology-related search terms (e.g. water, groundwater, hydrogeology, geology)
 - The keyword dictionary should be rationalized

- Most search terms are entered into the title textbox, rather than using individual search criteria
 - A google-style simple search box should be implemented
- A small percentage of documents were georeferenced (although almost all with country information)
 - This should not be a priority and the map search should be removed
- The architecture was old and inflexible using a combination of HTML, CSS, JavaScript and ColdFusion
 - This should be updated to a three-tiered software system with presentation, application and data tiers (as described in more detail below)
- There was no easy way for users to provide new data
 - A draft process should be developed, which can evolve over time as the governance model is formalized

7 Summary of system update (WP2)

The new SADC-GLA portal is implemented as a three-tiered software system comprising presentation, application and data tiers (Figure 1). In the current system the presentation tier is a web portal and the data tier is an Oracle 11g relational database system, although this can easily be migrated to an open source system. The application tier, which mediates between the web portal and the database, is a Representational State Transfer (REST) application programming interface (API) server.

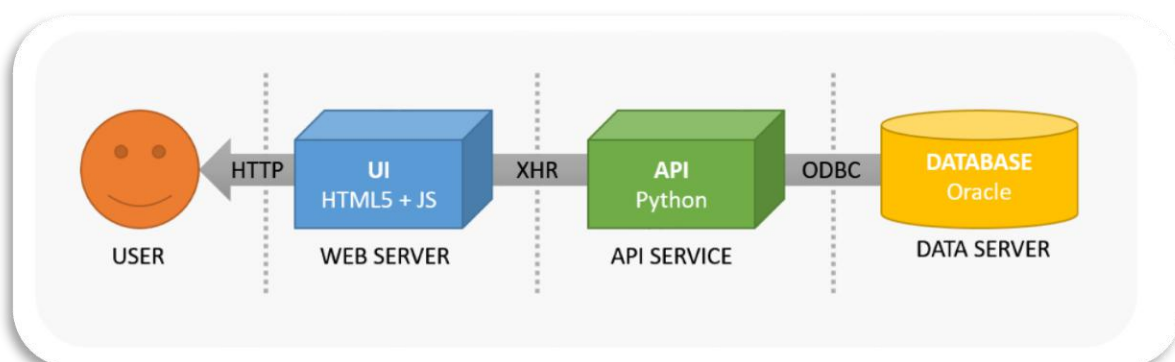


Figure 1 SADC-GLA Architecture Overview

The benefits of this approach is that the three tiers are only loosely coupled to each other. If the database structure changes the application can manage those changes without affecting the web portal. The design of the web portal could be modified, or a new portal or app developed, without requiring changes to the application. The API can also be extended to provide additional requests without affecting the existing functionality. Technical documentation of the system is provided in Annex 4 (note this is in English, but French and Portuguese versions are also available).

8 Summary of data update (WP2)

Existing data within the SADC-GLA was quality assured (QA) by the BGS team – this is an ongoing process. Protocols for adding new data, initially using an excel template then adding

functionality with an online form, were developed as defined in the governance strategy (see below).

The compilation of new data for the SADC-GLA was led by IGRAC and had 2 components:

1. Migration of data from the SADC-GIP and Ramotswa Information Management System (RIMS)
2. Compilation of new data by young professionals in the SADC Member States

Data from the SADC-GIP and RIMS were compiled first and used as a test of the data entry process using the excel template (the online form had not been developed at this point). A total of 69 new records were added to the SADC-GLA from these sources.

Following this initial data entry, the excel template was updated and provided to 25 young professionals from 16 Member States in the SADC region (DRC, Madagascar and Seychelles did not engage in the process). In total, 476 new records were compiled by the young professionals, with a significant amount of QA undertaken by IGRAC project staff (see Annex 5 for further details).

The BGS team used the new data provided by IGRAC to develop a methodology for entering data to the database – this will continue to be refined during the two-year maintenance period prior to final handover of the system.

9 Summary of training (WP3)

An online training for the SADC-GLA took place on Wednesday 19th and Thursday 20th August 2020. The aim was to introduce the SADC-GLA web portal, database, and data update protocols to different stakeholders – managers, users, and contributors – across the region. The training was attended by 26 participants from 12 countries within the SADC region (Malawi, Namibia, Zambia, Botswana, Comoros, Eswatini, Lesotho, Mauritius, Mozambique, South Africa, Tanzania, and Zimbabwe), as well as Raquel Sousa from IGRAC and James Manda and Brighton Munyai from SADC-GMI. The training was led by the BGS project team. A full summary of the training is provided in Annex 6 and all training materials are available through a google drive:

<https://drive.google.com/drive/folders/12iPUMcq7A7mzJv9ChO36N9kBYTqUpWcn?usp=sharing>

10 Long-term strategy (WP3)

As mentioned above, it was agreed early in the project that, due to lack of capacity, BGS should host and maintain the SADC-GLA in the short- to medium-term before handing over to SADC-GMI. The terms of this maintenance period and handover are defined in a governance strategy co-developed by BGS, SADC-GMI and IGRAC (see Annex 7).

The governance strategy outlines:

- Definitions for hosting and maintenance
- Protocols for data entry during the maintenance period
- Cost of maintenance



British
Geological
Survey



- Handover to SADC-GMI
- BGS' and SADC-GMI's obligations

As a result of the governance strategy, a two-year maintenance contract has been established between BGS and SADC-GMI, part-funded by unspent travel funds from the current project and part-funded by BGS as part of our ODA remit, beginning in November 2020.

11 Reflections and Recommendations

The SADC-GLA project has been undertaken under challenging circumstances for all those involved. The timescale of the project was already very tight and this was further exacerbated by the pandemic, with most staff having to manage the transition to home-working and a reduction in capacity within the first two months of the project. Even in the absence of an unforeseen global pandemic, a longer time-frame would have allowed more opportunity for user consultation as part of a more iterative development process.

Despite the challenges, a new SADC-GLA system has been developed to the specified requirements, with more than 500 new entries, and with a strategy in place for its long-term governance. The two year maintenance period will provide time for BGS to refine the data entry protocols prior to final handover. It will also provide time for SADC-GMI to build the required capacity to host and maintain the system into the future.

The feedback received from users during the online training was very positive, with recognition of the usefulness of such a resource for sustainable groundwater development and management in the region. The future success of the system will very much depend on continued and increased buy-in from a range of stakeholders, including the young professionals. As owners of the system, it is expected that SADC-GMI will continue to engage and encourage stakeholders to use and contribute to the system so that it remains an up-to-date and relevant resource into the future.

The pandemic was clearly very prohibitive for the young professionals' data collection activities and ongoing engagement with them could potentially elicit more data as restrictions are relaxed. They could also be involved in the final handover training workshop as a way of engaging them again further down the line.

It is expected that SADC-GMI will make provision for sufficient QA of new data, which can be a time consuming process, to ensure that a high standard of metadata is maintained so new documents are easily discoverable through the web portal.

One of the key issues for the SADC-GLA is IPR. As developers of the system, BGS and IGRAC have made every effort to ensure that copyright has not been breached through the sharing of documents without required permissions. Feedback from the young professionals indicated that some data providers were scared by the IPR permission form or unsure whether they were entitled to sign it. Further engagement by SADC-GMI with some of the main data providers in the region could potentially help to overcome this issue going forward.

The new SADC-GLA has been designed as a flexible, open and interoperable system, which has the potential to be linked to other knowledge platforms into the future. When the system is handed over to SADC-GMI at the end of the maintenance period, it is anticipated that some level of linkage be maintained with the Africa Groundwater Literature Archive so that updates relevant to the SADC region, whether obtained through the SADC-GLA or AGLA, are







British
Geological
Survey



discoverable through both systems. There may also be opportunity for the system to be linked to other regional or continental initiatives in future, so that the significant investment made during this project has as broad an impact as possible.

12 Project team & Contact Persons

The project team was composed of senior leaders and supporting scientific and technical staff from BGS and IGRAC. The key experts have many years' experience in developing agile systems and online databases and in working on groundwater-related projects with partners across the SADC region.

 	<p>SADC-Groundwater Management Institute (SADC-GMI) Mr. James Manda (Project Coordinator SADC-GMI) jamesm@sadc-gmi.org +265 991742582</p>
 <p>British Geological Survey Expert Impartial Innovative</p>	<p>British Geological Survey (BGS) Dr. Kirsty Upton (Team leader) kirto@bgs.ac.uk + 44 131 6500 210</p>
	<p>International Groundwater Resources Assessment Centre (IGRAC) Dr. Arnaud Sterckx arnaud.sterckx@un-igrac.org +31 6 82104674</p>



British
Geological
Survey



Annex 1: Project Kick-off Meeting Summary & Minutes

Updating the SADC Grey Literature Archive Project Kick-off Meeting: Summary & Minutes

5th February 2020



British
Geological
Survey



Project Summary

The Southern African Development Community Groundwater Management Institute (SADC-GMI) is a regional centre of excellence for groundwater management, whose vision is to ensure the equitable and sustainable use and protection of groundwater in the SADC region. SADC-GMI has initiated a project to update the SADC Grey Literature Archive (SADC-GLA), to make it a fully functional online archive of groundwater literature for the SADC region. The project started in February 2020 and will finish in June 2020. It will be implemented by the British Geological Survey (BGS), in collaboration with the International Groundwater Resources Assessment Centre (IGRAC).

In 2010, the BGS developed the SADC Grey Literature Archive (SADC-GLA¹), an online catalogue with almost 3000 unpublished reports, which were searchable through a web portal by title, author and country. Since 2014, BGS has built upon the SADC-GLA to develop the Africa Groundwater Literature Archive (AGLA²), which has improved functionalities and contains almost 1000 additional documents for the SADC region. Since 2017, IGRAC has developed with SADC-GMI the SADC Groundwater Information Portal (SADC-GIP), a web portal for sharing groundwater related data and information in the SADC region. The SADC-GIP is being improved by IGRAC as part of another project with SADC-GMI.

The project team will improve the functionalities and the content of the SADC-GLA. A new system will be developed, which can be easily updated and searched, is technologically advanced, and contains up to date information from stakeholders across the SADC region. The key aim of this project is to provide access to groundwater literature in the SADC region to all relevant stakeholders, to support the equitable and sustainable use and management of groundwater in this region.

Kick-off Meeting Summary

A kick-off meeting for the project was held on 5th February 2020 and attended by representatives of SADC-GMI and the BGS project team. Detailed Minutes of the meeting can be found below.

During this meeting, 3 main user groups of the SADC-GLA were identified: academia, policy makers and consultants. SADC-GMI act as product owners and proxies for these users, therefore it was agreed to hold a follow-up user consultation workshop with the SADC-GMI team, in order to define user needs in more detail, so these can be reflected in the design of the new GLA. The user consultation workshop was scheduled for the 14th February, the results of which are discussed in the output from Work Package 1: Situational Analysis Report.

It was also agreed at this meeting that BGS will host the updated SADC-GLA in the short- to medium-term, with the ambition that it will be transferred to SADC-GMI in the long-term. This will be addressed during the development of the governance model during Work Package 2 of the project.

¹ <https://www.bgs.ac.uk/sadc/index.cfm>

² <https://www.bgs.ac.uk/africagroundwateratlas/archive.cfm>



British
Geological
Survey



There was some initial discussion of data collection using the Young Professionals Model, and it was agreed that a Terms of Reference should be drafted for the Young Professionals prior to the user consultation workshop on 14th February.

Project Kick-off Meeting

Minutes

Venue IGS Board Room and Through Skype

Date: 05/02/2020, Time: 15:00Hrs CAT

Attendees

SADC-GMI: James Manda (JM), James Sauramba (JS), Brighton Munyai (BM), Mampho Ntsekhe (MN), Thokozani Dlamini (TD)

BGS: Kirsty Upton (KU), Colin Blackburn (CB), Ken Lawrie (KL), Carl Watson (CW)

Summary of Key Actions

1. All to attend next meeting on Friday 14th February to continue user needs consultation
2. Young Professionals process to be initiated through drafting of ToR
3. Contracts to be finalised

Opening and Introduction

- JM welcomed everyone and introductions made

Project Scope

- JM outlined the scope of this project – to update the existing GLA in connection and collaboration with SADC Member States – visibility of the SADC GLA is a key requirement

Detail on the current Africa Groundwater Literature Archive

- KU outlined the existing functionality of the existing Africa Groundwater Literature Archive, which contains around 4000 entries for the SADC region
- KU highlighted that IPR is the main barrier to allowing downloadable PDFs in AGLA – this is to be addressed during the project
- ACTION: KL to check how many of the SADC entries currently have a PDF or external link available

Work packages and key deliverables

- KU outlined the work packages and key deliverables (see Annex 1)

- All agreed that the content of the GLA will be determined by what literature is available, rather than user needs
- There are 3 main user groups, who may have different requirements in terms of how they access the GLA:
 - Academia
 - Policy makers (largely within government)
 - Consultants (private sector/industry)
- The GLA will be hosted by BGS in the short- to medium-term, with the ambition that it will be transferred to SADC-GMI in the long-term – this will be addressed during the development of the governance model and may involve, for example, an MoU between SADC-GMI and BGS
- SADC-GMI are the managers of the GLA as a service for Member States – JM and BM are the primary contacts
- ACTION: ALL to join meeting on Friday 14th February to continue consultation to define user needs, led by CW
- ACTION: CW to send some initial questions to SADC-GMI before next meeting
- The young professionals (YP) model should be initiated as soon as possible; IGRAC should be involved in the meeting on 14th February and prior to this a ToR for the YPs should be drafted (question: will they be using the same YPs as the GIP project?)
- ACTION: KU to follow up with IGRAC and SADC-GMI to begin drafting the YP ToR and work plan

Clarifications and Amendments

Contractual

- ACTION: KU to provide BGS' bank details for the contract, which can then be counter-signed by SADC-GMI
- All agreed that SADC-GMI will manage the YP budget (USD 6400) – this should be removed from the Reimbursables prior to counter-signing
- Reimbursables will be paid upon submission of invoices
- All agreed on deliverables/milestones, as reflected in Annex 1

Terms of reference

- All happy with the ToR, as reflected in the Work Packages and Deliverables (Annex 1)

Progress Meetings

- To take place monthly, following the initial meeting on 14th February

Any Information requests

- No additional information requests from BGS
- SADC-GMI request draft ToR for the YPs – see action above



British
Geological
Survey



Immediate next steps, tasks and deliverables

- Summary of initial meeting (ACTION: KU)
- Meeting on Friday 14th February to continue to the consultation process with SADC-GMI (ACTION: ALL)
- Initiate YP model (ACTION: KU, IGRAC & SADC-GMI)

Date for next meeting

- Friday 14th February

ANNEX 1 WORK PACKAGES AND KEY DELIVERABLES

Work Packages

1. Understand user needs for an updated system and assess the functionality of the existing GLA/AGLA
2. Improve the functionality of the GLA to meet the needs identified in WP1 within the time and budget constraints of the project
3. Train users in the SADC region and develop a long-term strategy to ensure the system is updated and maintained beyond the end of project

Work Package 1

- Consultation with SADC-GMI to define user needs and start developing governance model
 - Inception meeting with SADC-GMI (online)
 - OUTPUT > Summary Report (14 February)
- Review of existing data & system
 - OUTPUT > Situational Analysis Report (28 February)
- Initiate young professionals model

Work Package 2

- Update existing data (AGLA & SADC-GIP)
- New data
- Update database and migrate content
- Define protocols for adding/editing data
- Develop API and web portal
- Host archive on agreed infrastructure
- OUTPUTS
 - Progress Report (31 March)
 - Functional system with new web portal (15 May)
 - Final Report (30 June)

Work Package 3

- Deliver user manual and training to SADC staff
- OUTPUTS
 - User manual (30 June)
 - Training (30 June)



British
Geological
Survey



Annex 2: Situational Analysis Report

Updating the SADC Grey Literature Archive Situational Analysis Report

March 2020

Authors:

**Patrick Bell, Colin Blackburn, Simon Burden, Ken Lawrie, Martin
Nayembil, Andrew Simister, Arnaud Sterckx, Kirsty Upton, Carl
Watson**



British
Geological
Survey



Table of Contents

Introduction	21
Background.....	21
Objectives and Scope	21
Project Structure	21
User Consultation	22
Approach and methodology.....	22
Key User Requirements	23
Review of the current system	24
Data Review	24
System review: Technical implementation	Error! Bookmark not defined.
System review: Functionality and presentation	26
System Review: AGLA website usage statistics for 2019	32
System Review: Analysis of user search terms	33
Recommendations	34
Data Updates	34
Software System Updates	34
Search-related Updates.....	36
Administration-related Updates.....	36
Young Professionals Model.....	37
Work schedule	39
Project team.....	Error! Bookmark not defined.
Contact persons.....	Error! Bookmark not defined.



British
Geological
Survey



Introduction

The Southern African Development Community Groundwater Management Institute (SADC-GMI) is a regional centre of excellence for groundwater management, whose vision is to ensure the equitable and sustainable use and protection of groundwater in the SADC region. SADC-GMI has initiated a project to update the SADC Grey Literature Archive (SADC-GLA), to make it a fully functional online archive of groundwater literature for the SADC region. The project started in February 2020 and will finish in June 2020. It will be implemented by the British Geological Survey (BGS), in collaboration with the International Groundwater Resources Assessment Centre (IGRAC).

This report outlines the key findings of a rapid consultation process, carried out with SADC-GMI to define the key user requirements of the new SADC-GLA. It summarises a review of the existing system, against these requirements, and makes recommendations for improvements.

Background

In 2010, the BGS developed the SADC Grey Literature Archive (SADC-GLA³), an online catalogue with almost 3000 unpublished reports, which were searchable through a web portal by title, author and country. Since 2014, BGS has built upon the SADC-GLA to develop the Africa Groundwater Literature Archive (AGLA⁴), which has improved functionalities and contains almost 1000 additional documents for the SADC region. Since 2017, IGRAC has developed with SADC-GMI the SADC Groundwater Information Portal (SADC-GIP), a web portal for sharing groundwater related data and information in the SADC region. The SADC-GIP is being improved by IGRAC as part of another project with SADC-GMI.

Objectives and Scope

The key aim of this project is to provide access to groundwater literature in the SADC region to all relevant stakeholders, to support the equitable and sustainable use and management of groundwater in this region. To achieve this, the project team will develop a new system, which can be easily updated and searched, is technologically advanced, and contains up to date information from stakeholders across the SADC region.

Project Structure

The project is structured in 3 work packages:

4. Work package 1 – Review the current data and system
5. Work package 2 – Improve the functionality of the SADC-GLA to meet the needs identified in WP1 within the time and budget constraints of the project

³ <https://www.bgs.ac.uk/sadc/index.cfm>

⁴ <https://www.bgs.ac.uk/africagroundwateratlas/archive.cfm>

6. Work package 3 – Train users in the SADC region and develop a long-term strategy to ensure the system is updated and maintained beyond the end of project

This report outlines the results of Work Package 1, which has the following components:

1. User consultation: understanding user needs and frustrations with the current system
2. Review of the current system:
 - a. Data
 - b. Technical implementation
 - c. Functionality and presentation
3. Recommendations: how the system should be updated to better meet user needs and to enable it be more easily maintained in the future (the focus of work package 2).

In addition, work package 1 had the goal of initiating the Young Professionals Model as a mechanism for supporting the long term sustainability of the SADC-GLA (the focus of work package 3).

User Consultation

Approach and methodology

The following activities were undertaken to understand user needs and frustrations with the current system:

- Initial discussions on the user consultation process took place during the project Inception Meeting during a Skype video conference call on February the 5th 2020.
- It was agreed that, for the purposes of user consultation, SADC-GMI, IGRAC and BGS project partners could fulfil the roles of website administrator users as well as proxies for real world end users who search for documents in the literature archive.
- A User Research Workshop was scheduled for February the 14th, this was attended by BGS, IGRAC and SADC-GMI
- In order to facilitate the workshop across distributed physical sites the project partners agreed to use Miro, an Online Whiteboard for Visual Collaboration, and Skype.
- The Miro board was created in advance of the meeting and circulated along with pre-workshop instructions.
- During the workshop at least one person per site connected to the Miro board and proactively contributed to the session, see Figure 1.
- During the workshop, the following points were discussed:
 - Target audience(s)
 - User journey, including:
 - Search Engine Optimisation
 - Search form design
 - Metadata / document details
 - Search by map
 - Administration and site monitoring
 - IPR issues

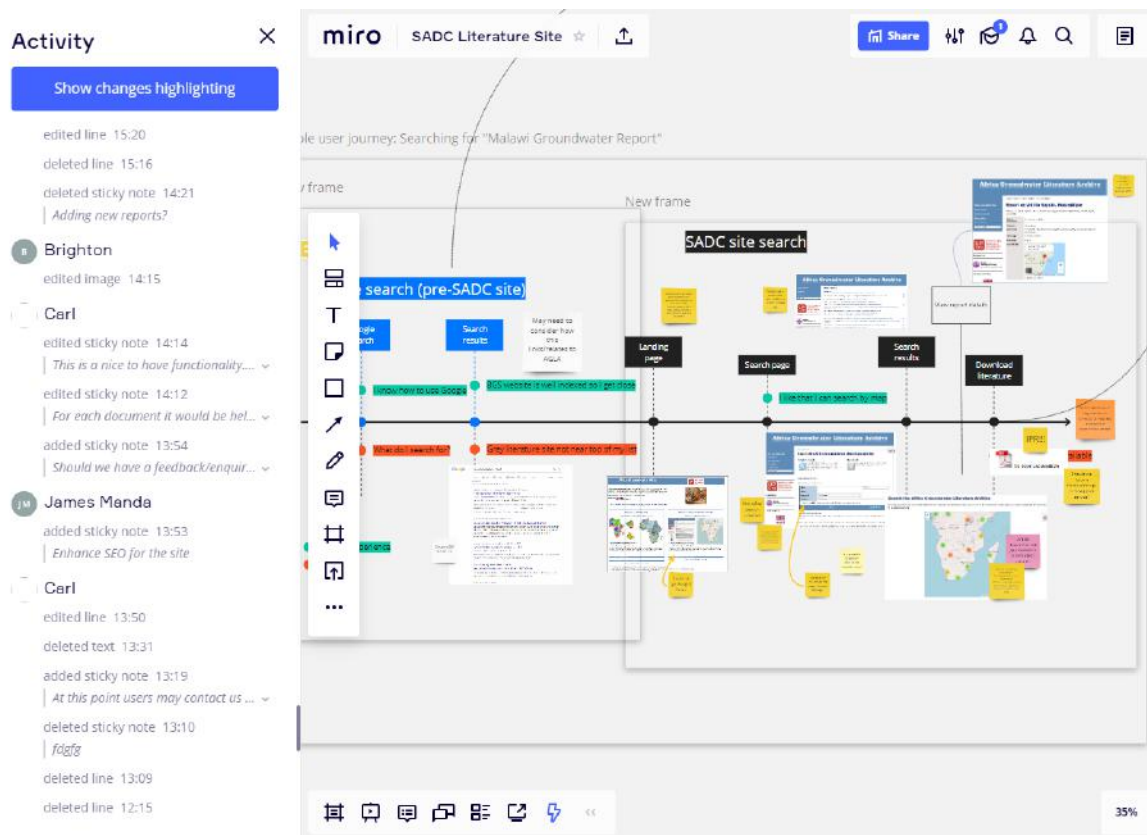


Figure 1: Miro board used in User Research workshop.

In the days after the workshop the Miro board and content was analysed to define the Key User Requirements outlined below. Finally, key findings and recommendations were produced to inform the design of the new version of the website and associated database, as outlined in the Recommendations Section below.

Key User Requirements

Website users looking for documents

Although there are a host of external users including policy advisers, commercial consultants, academic researchers and others it became apparent during the workshop that there was little to no significant difference in their requirements.

Goals

- To find groundwater related documents for the Southern Africa region

Frustrations

- Struggles to find the SADC website
- Confused about whether to use the map or search form
- Struggles to know which parts of the search form to use
- Fails to notice the checkbox that filters out results that have no accessible documents
- Regularly finds document metadata that states 'document not available'

Opportunities

- There may be many more potential users than the ones currently finding the SADC or AGLA websites
- Many of the documents which are flagged as not available are actually available elsewhere and we could help users to access these
- If the new site is more enjoyable to use then existing users may use it more often

Website administrators

It may be possible to subdivide these into more focused categories but during the workshop there was a lot of discussion on the need to improve how the content of the site is maintained in the future and how it would be beneficial to monitor the usage of the website to inform future improvements.

Goals

- To improve the way documents are described to in turn improve their discoverability and usefulness
- To add new documents to the archive
- To increase and monitor the impact of the archive

Frustrations

- No way for users outside of the BGS database team to add new documents
- No way for users outside of the BGS database team to improve metadata
- Unclear why so many documents are restricted access
- Unclear how to resolve the IPR concerns about making more documents openly available
- I don't know how many users visit the site or what they are searching for

Opportunities

- Would like to make SADC archive interoperable with other associated document databases
- More partners could add new content making the archive larger and more up to date

Recommendations regarding how to improve the current system to address these findings are discussed later in the report.

Review of the current system

In order to understand what work was required to improve the system implementation, the following review activities were undertaken:

1. Data review
2. System review: Technical implementation
3. System review: Functionality and presentation

Data Review

All entries within the SADC-GLA and AGLA have been reviewed to determine the level of information that currently exists for each entry within the archive. This review identifies the total number of entries for the SADC region – both those from the original SADC Grey Literature Archive and AGLA – and identifies the key missing metadata.

The data review categorised the existing entries in the database to provide a summary of the metadata as follows:

	SADC	AGLA
Total number of entries	2997	988
Entries with Keywords (%)	89%	71%
Entries with Authors (%)	99%	95%
Entries with Language (%)	4%	91%
Entries with Country (%)	92%	100%
Entries with Georeferencing (%)	2%	35%
Entries with Live document link (%)	14%	44%
Entries with Link to external site (%)	18%	41%

Keywords are semi-hierarchical and have parent groupings. The Keyword dictionary is extensive and very granular. Geographical extents are classified in two ways: by country and location (with georeferencing). The Country classification has a dictionary and almost all entries include this information. Georeferencing is either by a point or bounding box and is much less complete in the current Archive. Just over 40% of entries have a live document link, and a similar proportion contain links to an external site. BGS has PDFs available for a further 178 documents, which are not made available for IPR reasons. From the initial review, several duplicate entries have been identified in the archive.

System Review: Technical Implementation

The AGLA portal was reviewed as this has the most recent and most modern codebase and provided a higher level of functionality. The AGLA portal is currently implemented in HTML, CSS and JavaScript. Data access is provided by ColdFusion applications reading from an Oracle database. The Map Search uses the open source [Leaflet Map](#) JavaScript plugin. Individual records use an embedded Google Map plugin (which is no longer supported for free). This technology stack can be improved upon to provide a more modern architecture that is more flexible and easier to maintain. It can also be implemented utilising more modern web development technologies in line with current industry standards and the skills base of present-day web developers. Recommendations for modernising the technical implementation are provided later in this document.



British
Geological
Survey



System review: Functionality and presentation

The following pages make up the current system's user interface:

Home Page

<https://www.bgs.ac.uk/africagroundwateratlas/archive.cfm>

Links to:

- “**Keyword Search**” page >>>
- “**Map Search**” page >>>

Country dropdown jumps to “**Search Results**” results for selected country >>>

This page is available in both the English and French language.

Keyword Search

<https://www.bgs.ac.uk/africagroundwateratlas/atlas.cfc?method=search>

Africa Groundwater Literature Archive

Home » Search

Africa Groundwater Atlas

Archive home

About the Archive

External links

Search Archive

Select Country

Search the Africa Groundwater Literature Archive

Keyword search

To search for references on a particular topic or combination of topics, use the keyword search option. There is also a free text option to search titles and authors.

Map search

You can also search references that have been georeferenced using an interactive map.

Keyword search form

Use this form to search the archive.

Title	<input type="text"/>	(All of these)
Author	<input type="text"/>	
Keyword	<input type="text"/>	
Country	<input type="text"/>	

☐ Search to **ONLY** include documents with PDF download or external resource

Documents with no online link

If there is no document download or online link for a document in the Archive, it means we do not know of an online version of the document.

There may be a digital version, but if the Archive does not hold copyright for the document, unfortunately we can't make it available.

Many of the documents in the Archive are only available in hard copy, and are often difficult to obtain. Unfortunately, the Archive does not store hard copies of any of the documents.

If there is no online link in the Archive, you could try to search for a document on the wider web using a search engine such as Google. If information on the publishing organisation is available, you could try to contact the publisher directly to obtain a copy. If you can't find any more information elsewhere, please feel free to contact us via the contact page and we will try to help you find it if we can, but we may not be able to help further!

If you know of an online version of a document which doesn't already have a web link in the Archive, or you have digital versions of documents which the copyright holder agrees can be made available through the Archive, please get in touch via the contact page and let us know, so we can update and improve the Archive.

Prepared by:

British Geological Survey
NATURAL ENVIRONMENT RESEARCH COUNCIL

Jointly funded by:

UKaid
from the British people

NERC
NATURAL ENVIRONMENT RESEARCH COUNCIL

ESRC
ECONOMIC & SOCIAL RESEARCH COUNCIL

Copyright | Contact | Privacy policy | Terms & Conditions

© BGS, DFID, NERC, ESRC 2020. This site is hosted by the British Geological Survey.
Questions, suggestions or comments regarding the contents of this site should be directed to AfricaGWAtlas@bgs.ac.uk.

Search by one or more of:

- Title (*free-text*)
- Author (*free-text*)
- Keyword (*list*)
- Country (*list*)
- ANY/ALL/EXACT match (*select*)
- Only return matches with file links (*checkbox*)

Submits to **"Search Results"** >>>

This page is available in both the English and French language.

Map Search

https://www.bgs.ac.uk/africagroundwateratlas/leaflet_map.cfm

Africa Groundwater Literature Archive

Home | Search

Search the Africa Groundwater Literature Archive

The map shows the locations of 2189 references that have been georeferenced so far, out of a total of 8710 references in the archive. Documents marked with have a link to an external resource.

Africa Groundwater Atlas

Archive home

About the Archive

External links

Search Archive

Select Country

Unlocking the Potential of Groundwater for the Poor

Prepared by:

British Geological Survey
NATURAL ENVIRONMENT RESEARCH COUNCIL

Jointly funded by:

UKaid
from the British people

NERC
NATURAL ENVIRONMENT RESEARCH COUNCIL

ESRC
ECONOMIC & SOCIAL RESEARCH COUNCIL

Copyright | Contact | Privacy policy | Terms & Conditions

© BGS, DFID, NERC, ESRC 2020. This site is hosted by the British Geological Survey.
Questions, suggestions or comments regarding the contents of this site should be directed to AfricaGWAtlas@bgs.ac.uk.

Click an item on the map

Links to **“Full record”** >>>

This page is available in both the English and French language.

Search Results

https://www.bgs.ac.uk/africaGroundwaterAtlas/atlas.cfc?method=listResults&title_search=kenya&authorBoolean=AND&titleboolean=AND&pageSize=10&start=80&reportpdf=Y

Africa Groundwater Atlas

Archive home

About the Archive

External links

Search Archive

Select Country

Unlocking the Potential of Groundwater for the Poor

Prepared by:

British Geological Survey

NATURAL ENVIRONMENT RESEARCH COUNCIL

Jointly funded by:

UKaid

From the British people

NERC

THE NATURAL ENVIRONMENT RESEARCH COUNCIL

ESRC

ECONOMIC & SOCIAL RESEARCH COUNCIL

Africa Groundwater Literature Archive

Home » Search » Search results

Search results

Kenya

Hughes, IG . 1981. *Report on advisory/inspection visits to Pakistan and Kenya: 1981*

Hughes, IG . 1982. *Report on advisory/inspection visits to Malawi, Swaziland and Kenya: Feb/Mar 1982*

Willmson, D, Taleb, M, Dammati, B, Icole, M, Thouvey, N . 1993. *Equatorial extension of the younger Dryas event: rock magnetic evidence from Lake Magadi (Kenya)*

Institute of Hydrology . 1985. *Groundwater conditions in Baringo District, Kenya*

JICA . 2013. *National Water Master Plan (Kenya)*

Cassidy, R, Soule, H, Robins, N, Join, JL, Bantson, O, Marobhe, I, Obando, J, Comte, JC, Makokha, M, Shauri, H, Melchior, S, Mjemah, JC, Kassam, IA, Mohamed, I . 2014. *Towards Groundwater Security in Coastal East Africa - an interdisciplinary study in Comoros, Kenya and Tanzania*

Jones, BF, Eugster, HP, Rettig, SL . 1977. *Hydrochemistry of the Lake Magadi basin, Kenya*

Williams, TM, Rees, J, Kairu, KK, Yabe, AC . 1996. *Assessment of contamination by metals and selected organic compounds in coastal sediments and waters of Mombasa, Kenya*

Rees, JG, Williams, TM, Ngugi, MM, Kairu, KK, Yabe, AC . 1996. *Contaminant transport and storage in the estuarine creek systems of Mombasa, Kenya*

Mumma, A, Lane, M, Kairu, E, Tuinhof, A, Hirji, R . 2011. *Kenya Groundwater Governance, Case Study*

Showing 81 - 90 of 194 [10] [50] [100] [All]

Previous 4 5 6 7 8 9 10 11 12 13 Next

Copyright | Contact | Privacy policy | Terms & Conditions

© BGS, DFID, NERC, ESRC 2020. This site is hosted by the British Geological Survey. Questions, suggestions or comments regarding the contents of this site should be directed to AfricaGWAtlas@bgs.ac.uk.

Shows list of matched reports grouped by country:

- Author(s)
- Year
- Title
- Icon indicating file availability (**NOT A LINK!**)
- Link to **"Full Record"** >>>

This page is available in both the English and French language.

Full Record

<https://www.bgs.ac.uk/africaGroundwaterAtlas/atlas.cfc?method=ViewDetails&id=KE0009>

Africa Groundwater Literature Archive

Home » Search » Search results » Document details

Africa Groundwater Atlas

Archive home

About the Archive

External links

Search Archive

Select Country

Kenya Rift Valley Geothermal Project, Report on a visit 28 April to 22 May

Darling, WG. 1987. Kenya Rift Valley Geothermal Project, Report on a visit 28 April to 22 May. British Geological Survey (WD/OS/87/12).


Link to document: No download available.

Country: Kenya


Keywords: Geothermal,


Item type: Publication - Report


Prepared by:

 **British Geological Survey**
NATURAL ENVIRONMENT RESEARCH COUNCIL

Jointly funded by:

 **UKaid**
from the British people

 **NERC**
NATURAL ENVIRONMENT RESEARCH COUNCIL

 **ESRC**
ECONOMIC & SOCIAL RESEARCH COUNCIL

Copyright | Contact | Privacy policy | Terms & Conditions

© BGS, DFID, NERC, ESRC 2020. This site is hosted by the British Geological Survey.
Questions, suggestions or comments regarding the contents of this site should be directed to AfricaGWAtlas@bgs.ac.uk.

Shows full details of a report:

- Author(s)
- Year
- Title
- Abstract
- Countries
- Report Type
- Language(s)
- Link to "File" if available >>> URL or DOI
- Coverage (embedded Google Map – **NOT WORKING** CONSISTENTLY FOR ALL USERS)

This page is only available in English.

Recommendations for how the functionality and presentation of the current system can be improved to better meet the user needs extracted during the user consultation exercise are discussed in the Recommendations section of the report.

System Review: AGLA website usage statistics for 2019

ALGA web site usage statistics for the year 2019 were analysed to understand the level of usage of the site and to help corroborate any findings resulting from the user consultation process.

Daily Visitors

Visitors are actual people looking for information and include desktop browser and mobile users.

Average Visitors per Day 59

Page Views per Day 84

Total Visitors for 2019

There were a total of 13111 visitors in 2019 reaching the various pages in the AGLA website as broken down below:

Home page (EN or default) 9167

Home page (FR) 429

Search page (EN) 12071

Search page (FR) 1040

Map page (EN) 839

Map page (FR) 67

Details page (EN only) 959

Document Searches

These are search queries actually submitted by users.

279161 document searches BY **13111** visitors = **~21** searches per visitor

Website hits

Hits are access to any file forming part of the site e.g. individual images, css file etc



British
Geological
Survey



Total Hits	2611476		
Visitor Hits	427128	17% of total	
Spider Hits	2184348	83% of total	(from Google/Bing etc)
Average Hits per Day	7154		
Average Hits per Visitor	20		

Website usage statistics notes

- The comparison of search page visitors against home page visitors indicates that a lot of people have bookmarks to the search page
- The French versions of the pages are requested on average by 7% of visitors
- Nearly 50% of hits are from online marketing company "AhrefsBot" - these could be blocked if wished (<https://botcrawl.com/ahrefsbot/>)
- Nearly 20% of hits are from commercial UK distributed search "MJ12bot" - these could be blocked if wished (<https://mj12bot.com/>)

System Review: Analysis of user search terms

We also analysed user search terms entered into the web portal in an attempt to better understand the topics of interest to users.

The following is a previously undertaken detailed snapshot analysis of the most popular keywords used to search the SADC web portal in an attempt to better understand the topics of interest to users. This snapshot analysis of search terms covers the time period from April to September 2011. This analysis was a time consuming process and time wasn't available to repeat it for 2019 within the timeframe of this situational analysis report. A rapid analysis was undertaken to confirm that the outcomes of the analysis from 2011 were still relevant to the searches being undertaken at the present time.

The search terms listed in Table 1 are a combination of the terms entered into all of the available search fields on the search form (title, author, keyword, country). However, it should be noted that 90% of all search terms are entered into the "Title" textbox suggesting that a simpler search form with a Google-style single search box might be considered in the redevelopment of the portal.

Country Search Terms	Water-related Search Terms	Geology-related Search Terms	Organisation Search Terms	Report Type Search Terms
(231) mozambique	(337) water	(296) geology	(84) resources	(42) report

(115) Malawi	(230) ground	(187) geological	(77) survey	(26) literature
(102) madagascar	(203) hydro	(167) map	(55) sadc	(22) review
(65) zambia	(187) groundwater	(68) mineral	(52) department	(20) download
(62) africa	(131) hydrogeology	(46) exploration	(21) british	(14) study
(53) south	(42) hydrogeological	(42) supply		(8) master
(41) namibia		(24) borehole		
(40) botswana		(23) drilling		
(35) zimbabwe				
(32) tanzania				
(30) angola				
(22) swaziland				
(20) lesotho				
(11) congo				

Recommendations

The following recommendations are made to enhance the SADC-GLA portal in line with user requirements and to improve its maintainability and sustainability going forward.

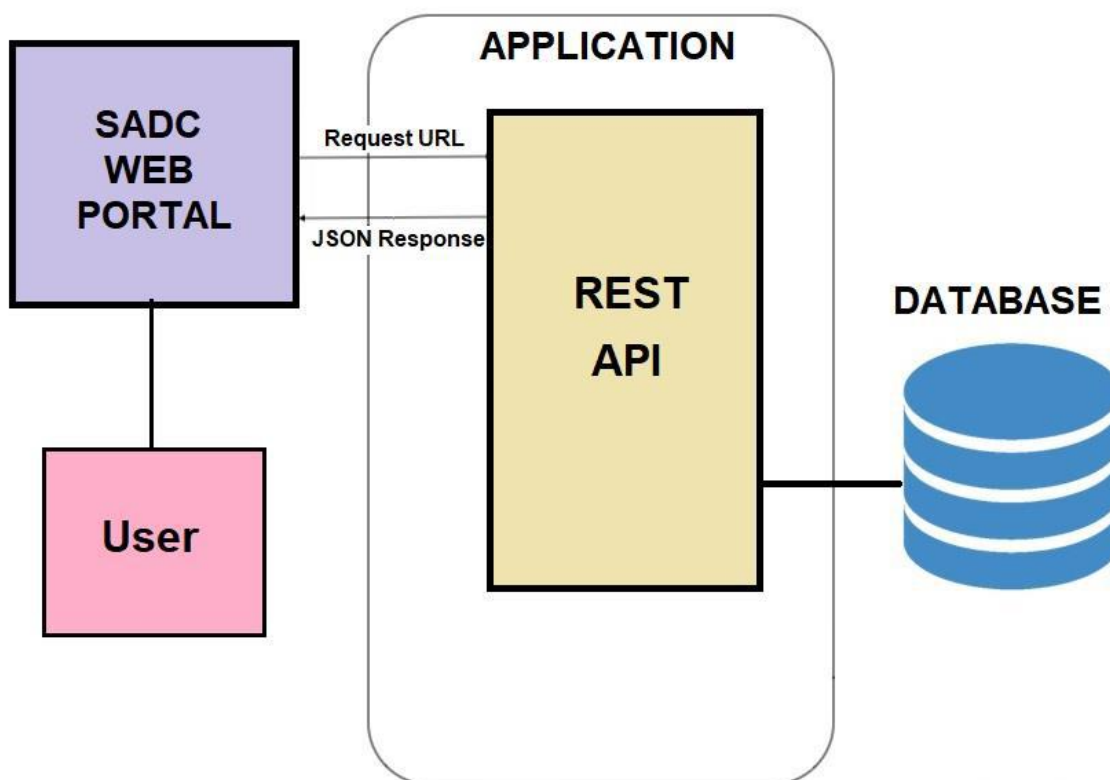
Data Updates

A general QA of all existing SADC entries will be undertaken to fill the identified data gaps, where possible, and remove duplicates. This will focus on updating the live document links and link to external sites to improve the availability of literature through the Archive. Georeferencing will not be undertaken for existing entries in the database, but Country information will be completed. The Keyword dictionary will be rationalised.

The database will be standardised so that any data manipulations are accurately audited with a full and complete audit trail. Additional dictionary constraints will be added to the database to provide better control of metadata.

Software System Updates

In order to modernise the technical implementation of the SADC-GLA portal, a three-tiered software system is recommended comprising a presentation, application and data tier as outlined in the diagram below.



The presentation tier in this instance will be a web portal while the data tier will be the revised Oracle database system. All development will utilise the GitLab source code management system and be deployed by test-driven CI/CD pipelines.

Web portal

The web portal will provide the user interface and be responsible for presenting the interface that allows users to define their search requirements. The web portal will interact with the application (described below) in order to return and present the information requested by the user. The web portal will be implemented using html, javascript (using libraries such as jQuery) and css and will be responsive to ensure it works on a range of screen sizes.

Application

The application sits between the web portal and the database. It provides an Application Programming Interface (API) that takes requests from the web portal, queries the database and then returns data to the web portal. The application will implement a standard Representational State Transfer (REST) API, taking requests in the form of Uniform Resource Locators (URLs) and returning data as standard JavaScript Object Notation (JSON) text.

Database

Generally, the database will be rationalised into a robust structure with an architecture that allows us to continue to host it as now but also lends itself to being run on a different platform if required.

Benefits

The benefits of this approach is that the three tiers are only loosely coupled to each other. If the database structure changes the application can manage those changes without affecting the web portal. The design of the web portal could be modified, or a new portal or app developed, without requiring changes to the application. The API can also be extended to provide additional requests without affecting the existing functionality.

Search-related Updates

The following recommendations are ordered in the user journey sequence captured in the User Research workshop and are not prioritised:

- i. Improve search engine optimisation
- ii. Consider spell checking search criteria
- iii. Simplify the search form. This could mean a single text box for search criteria that was then used in queries against the key metadata fields.
- iv. Make the 'exclude results where document not available' more obvious
- v. Potentially remove or simplify the map interface. Most documents relate to a country rather than a specific location and most users appear to search by country name rather than a more targeted geographic location.
- vi. Refresh the search results page and document details pages taking into account the most well populated fields as informed by the Data Review work stream.
- vii. If a document is not available for download then offer users an easy way to perform a search for the document elsewhere – for example many of the documents come with DOIs or academic references which are highly likely to result in the document being found via popular search engines.
- viii. The user review omitted to consider the need for the system to be in both the English and French language. We will return to the user proxies to query that requirement.

Administration-related Updates

- i. Make use of Google Analytics and create a dashboard that can be accessed by SADC-GMI, BGS and IGRAC
- ii. Create a draft process for entering new documents that involves SADC-GMI and their trusted partners but expect this process to evolve over time as the governance model is formalized.

- iii. Define the governance model for adding, editing and quality control for the content on the new site.
- iv. Investigate how to overcome IPR issues and ultimately make more documents publically accessible.

Young Professionals Model

Since September 2019, SADC-GMI has initiated a project to expand the SADC Groundwater Information Portal (SADC-GIP), which is being implemented by IGRAC. The SADC-GIP, launched in 2017, is a web portal, consisting in a map viewer and a document database for storing and sharing groundwater related data and information in the SADC region. It is managed by SADC-GMI with the technical support of IGRAC. It was established to make information easily available for the SADC Member States, research institutions and other international organisations. It contains, among others, maps compiled during the SADC Hydrogeological Mapping Project, a borehole database, the SADC Groundwater Risk maps, etc.

As part of the project, young professionals from the SADC members were engaged to be capacitated on groundwater data storage and sharing aspects. They work in ministries or agencies responsible for groundwater management and are involved in topics like groundwater data collection, data management, policy development or policy implementation. Most of them are below 35 and have a higher education degree. Young professionals are involved in the following project activities:

- Contribute to the situational analysis of groundwater databases in SADC.
- Complete an on-line training course on QGIS.
- Participate in a knowledge management workshop;
- Participate in a training workshop on Geographic Information Systems (GIS), Spatial Data Infrastructure (SDI) and the SADC-GIP.

Focal persons from the member states were asked to nominate 2 young professionals per country. To date, 25 young professionals are part of the project, coming from 13 SADC member states (out of 16).

Member state	Number of young professionals
Angola	2
Botswana	2
Comoros	2
DRC	\
eSwatini	2

Lesotho	2
Madagascar	\
Malawi	2
Mauritius	1
Mozambique	2
Namibia	2
Seychelles	\
South Africa	2
Tanzania	2
Zambia	2
Zimbabwe	2
Total	25

The young professional will be offered to extend their contract to also contribute to the SADC-GLA project. They will be tasked to collect documents and bibliographic records related to groundwater in their country. These can be reports, policy documents, manuals, scientific papers, studies, books, etc. coming from national and international organizations. IGRAC will validate the records collected and upload them in the SADC-GLA. In order to monitor the engagement of the young professionals, they will report on their bibliographic research. Their reports will include: organizations consulted; list of data collected from each source; main challenges to collect documents; potential future sources of information and suggestions for the SADC-GLA.

Young professionals will be given an allowance of USD 200.00 upon completion of all tasks, which should be due by the 30/04/2020. The young professional will be supervised by the project team and focal persons.

Work schedule

	Task Description	Jan	Feb	Mar	Apr	May	Jun
Work Package 1	WP1.1 Data Review						
	1. Review existing data						
	2. Initiate young professionals' model						
	WP1.2 System Review						
	3. Identify key requirements and functionality						
	4. Develop a governance model						
	5. Review the current system (database and portal)						
Work Package 2	WP1.3 Report						
	6. Report on WP1 including plan for WP2						
	WP2.1 Data Update						
	7. Update existing GLA data						
	8. Prepare new data from AGLA and SADC-GIP						
	9. Compile new data from Member States						
	WP2.2 System Update						
Work Package 3	10. Update and populate database system						
	11. Develop protocols for adding/editing/updating data						
	12. Develop a web-enabled API						
	13. Develop a new web-portal						
	14. Host the updated SADC GLA system						
	WP2.3 Report						
	15. Report on WP2, including long-term strategy						
Other	16. User Manual						
	17. Undertake extensive training of key system users						
Ongoing Project Management							



British
Geological
Survey





British
Geological
Survey



Annex 3: Project Update

Updating the SADC Grey Literature Archive Project Update

31st March 2020

Actions and updates from previous meeting (17th March)

ACTION 1: Domain name: BGS team to send domain options and cost implications to SADC-GMI; SADC-GMI to finalise name of archive and URL

UPDATE 1: From IONOS, who BGS have used previously for registering domain names:

- .co.uk - £10 per year
- .org.uk - £10 per year
- .com - £13 per year
- .org - £16 per year
- .info - £16 per year
- Existing domain names can be checked on the [IONOS website](#)

ACTION 2: Progress report: BGS/IGRAC to send short progress report by 31st March

UPDATE 2: Completed

ACTION 3: Governance document: BGS/IGRAC to send draft document for comment by 15th April, for discussion during next meeting on 17th April

UPDATE 3: In progress

ACTION 4: Young Professionals: SADC-GMI/IGRAC to finalise contracts and instructions for YPs, including approaching countries currently missing from the GIP project

UPDATE 4: YPs now have contracts and instructions to begin work

Internal Project Workshop

The project team held a two-day project workshop on 25-26th March. The workshop was meant to be held at BGS' head office in Keyworth, but instead was conducted online via ZOOM. The aim of the workshop was to:

- Based on the outcomes of the Situational Analysis Report, agree final decisions on the design of the new SADC-GLA system, including the database, API and web portal
- Define the specific tasks required, timeline and responsibilities within the team
- Agree on key aspects for the governance model

Day 1 was attended by BGS staff only, while day 2 also involved IGRAC staff.

Database

The structure of the new database has been finalised (Annex 1) and is now in the process of being implemented. Work has begun on updating existing data within the

AGLA database so it can be migrated to the new database structure. This will involve reviewing and updating the Keyword dictionary into a rationalised “parent” and “child” structure, then mapping all existing keywords within the AGLA database to the new SADC-GLA keyword dictionary. The database updates are due to be completed by 17th April (Figure 1).

After this, the database team will focus on further QA of existing data within the SADC-GLA – primarily filling in key missing data – and the addition of new data – initially from the SADC-GIP, then from the YPs and IGRAC.

The database team are also in the process of finalising a new metadata template to be used and tested by IGRAC and the YPs to collate new data for the SADC-GLA.

Web Portal & API

The structure of the new web portal and API has been finalised (Annex 2) and preparatory work is underway so the main phase of development can begin as soon as the database development is complete. The majority of the web portal and API development will take place as a development sprint during the last two weeks of April (Figure 1). During the sprint, the web and application staff will work intensively on the project, with input from the database team when required.

Prior to this, the web team will produce final layouts and mock-ups for review by SADC-GMI. Text for the web pages will also be drafted for review by SADC-GMI.

Review Process

Following the development sprint, the new system will undergo an internal review process within BGS during the first two weeks of May (Figure 1). The new functional system will then be released for external review by 15th May, allowing feedback from SADC-GMI to be addressed prior to project completion at the end of June.

New data from the SADC-GIP and Young Professionals

Instructions were issued to the YPs on 25th March with a provisional metadata template, which will be updated with the new template at the beginning of April. BGS are also working on an IPR authorisation template, which will be completed by the YPs to ensure the required permissions have been given for documents to be shared in the GLA.

A deadline of 22nd May has been set for the YPs to return completed templates and reports to IGRAC. IGRAC will follow up with the YPs during this time to ensure the deadline is met. IGRAC will then QA the data provided by the YPs so it can be loaded directly to the GLA by BGS staff.

Prior to the YP deadline, IGRAC will use the new metadata template to provide data for all relevant documents currently stored in the SADC-GIP. This metadata will be uploaded to the GLA (Figure 1).



British
Geological
Survey



Governance

Governance was discussed during the internal project workshop and work has begun on developing a governance document, which will be circulated to SADC-GMI on 15th April.

Other

A French translator is confirmed and the project team are in the process of sourcing a Portuguese translator.

The potential impact of Covid-19 on project delivery was discussed during the project workshop. The project team do not foresee any delays to the system development. It was suggested there could be delays to data collection by the YPs, but this should be minimal, as the majority of data collection will be done online/remotely. The training workshop, which was scheduled to take place in South Africa in June will not be possible. This should be discussed further between SADC-GMI and BGS to determine whether this could take place at a later date, or online.



British
Geological
Survey



FIGURE 1: PROJECT TASKS & TIMELINE

	30 Mar- 3 Apr	6-10 Apr	13-17 Apr	20-25 Apr	27 Apr- 1 May	4-8 May	11-15 May	18-22 May	25-29 May	1-5 Jun	8-12 Jun	15-19 Jun	22-26 Jun
DATABASE													
Update database structure													
Fix and migrate existing data													
QA existing data													
Add existing data from SADC-GIP													
Contribute to SPRINT													
WEB & API													
Preparation work													
SPRINT													
Internal Review													
External Review													
FINAL SYSTEM													
Finalise based on external review													



British
Geological
Survey

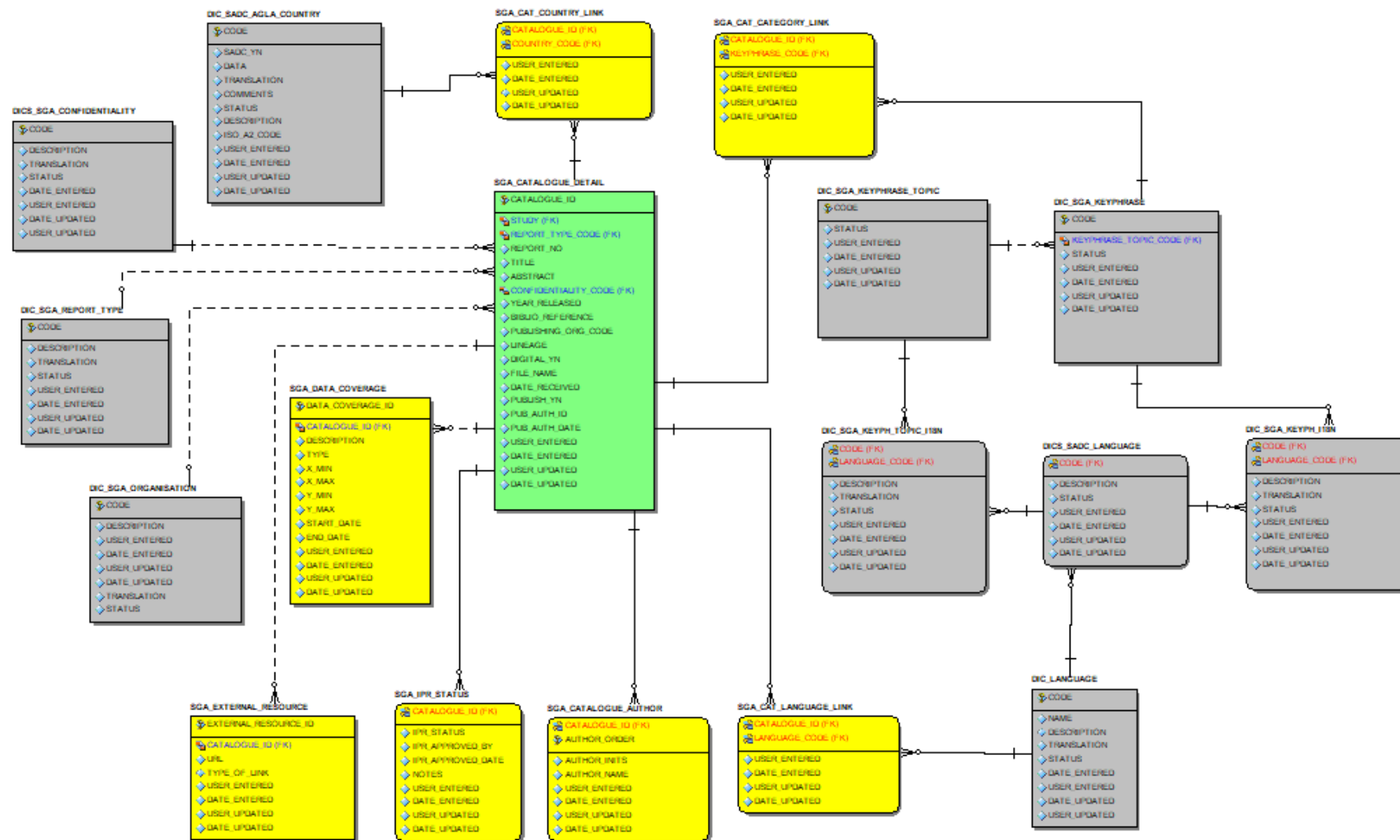


Final Report & User Manual													
GOVERNANCE													
Draft governance doc for SADC-GMI													
YOUNG PROFESSIONALS													
Receive metadata templates													
QA metadata templates													
Add data													
DELIVERABLES	1		2				3						4

DELIVERABLES

1. Progress report – 31/3
2. Draft governance document (15/4) and progress meeting (17/4)
3. Functional System – 15/5
4. Final System, report and user manual – 30/6

Annex 1 SADC-GLA Relational Database Model Design



Annex 2 SADC-GLA Web Portal Design

- **Home**
 - Intro text
 - Phrase search {TEXT} >> **Search Results**
 - Links to country pages (x16) >> **Country List**
 - Link to >> **Advanced Search**
 - *Contact details*
- **Advanced Search**
 - Title {TEXT}
 - Author {TEXT}
 - Country / Countries {LIST OF CODES}
 - Hierarchical (2 tier) Keyword / Keywords {LIST OF CODES}
 - Only Linked {BOOLEAN}
 - {GO BUTTON} >> **Search Results**
 - *Allow multiple countries and keywords*
- **Search Results**
 - Table of citations with column sort options and paging links
 - Citations show live links where available
 - *Highlight search match criteria (i.e. word in title, keyword etc)*
 - *Filtering capability*
 - Rows link to >> **Citation Details**
- ACTION: ALL to decide on list of fields to be displayed in search results
- **Citation Details**
 - Table of all available metadata for a citation
 - Citations show live links where available
 - If spatial data present, uses JavaScript progressive enhancement to add map
- **Country List**
 - Single page list of all citations for a single country
 - Intended for easy indexing by search engines
 - Rows link to >> **Citation Details**
- **About**
- **T&Cs, Privacy & Copyright**

SADC Groundwater Literature Archive System Documentation

June 2020

Authors:

**Patrick Bell, Colin Blackburn, Simon Burden, Ken Lawrie, Martin
Nayembil, Andrew Simister, Kirsty Upton**

Table of Contents

Introduction.....	51
System Overview.....	51
User Interface	51
Design.....	51
Folder Descriptions	52
JavaScript Libraries Used.....	52
Application JavaScript.....	53
Data Management.....	53
Map Service	53
API.....	53
Design.....	53
Using the API	54
Software implementation.....	54
Server deployment	55
Glossary.....	55
Database	55
Database overview.....	55
Database Structure	56
Dictionaries – Vocabularies	56
SGA_CATALOGUE_DETAIL.....	58
Database Environment.....	60
Summary of Data in the Database.....	61
Contact persons.....	Error! Bookmark not defined.

Introduction

The Southern African Development Community Groundwater Management Institute (SADC-GMI) is a regional centre of excellence for groundwater management, whose vision is to ensure the equitable and sustainable use and protection of groundwater in the SADC region. In 2020, SADC-GMI initiated a project to update the SADC Grey Literature Archive (SADC-GLA), to make it a fully functional online archive of groundwater literature for the SADC region. The project started in February 2020 and was implemented by the British Geological Survey (BGS), in collaboration with the International Groundwater Resources Assessment Centre (IGRAC).

The new SADC-GLA is implemented as a flexible and interoperable system, comprising a database, application, and online user interface. The system will be hosted by the British Geological Survey for 2 years, and then handed over to SADC-GMI. This report provides the technical specification for the system.

System Overview

The SADC-GLA portal is implemented as a three-tiered software system comprising presentation, application and data tiers as outlined Figure 1. In the current system the presentation tier is a web portal and the data tier is an Oracle 11g relational database system. The application tier, which mediates between the web portal and the database, is a Representational State Transfer (REST) application programming interface (API) server.

The benefits of this approach is that the three tiers are only loosely coupled to each other. If the database structure changes the application can manage those changes without affecting the web portal. The design of the web portal could be modified, or a new portal or app developed, without requiring changes to the application. The API can also be extended to provide additional requests without affecting the existing functionality.

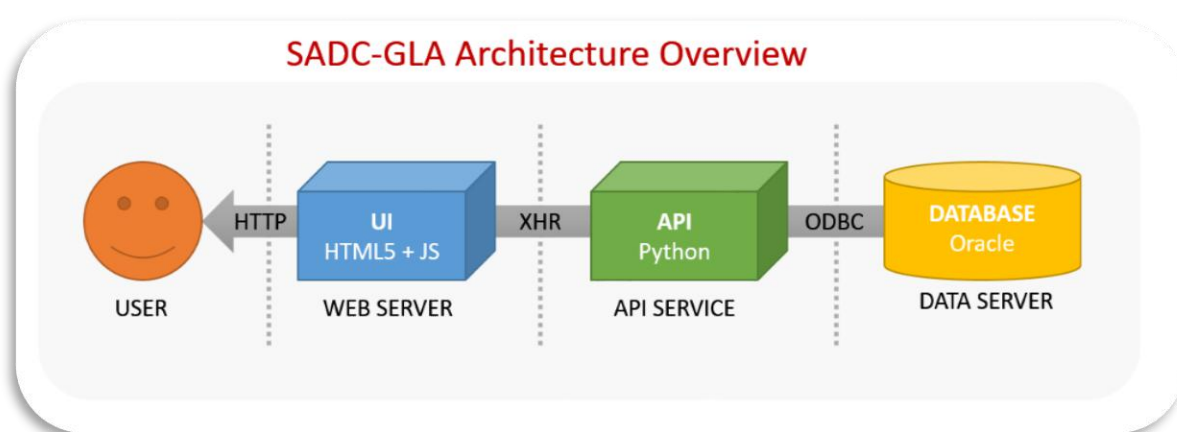


Figure 1: SADC-GLA Architecture Overview

User Interface

Design

The SADC-GLA User Interface is implemented in HTML and JavaScript and accesses data using XHR requests to the API.

- The UI and database are not directly linked – all data requests use XHR to call the API
- All data access is read-only and all requests will use HTTPS
- The UI code is designed to be self-contained and does not link to any external resources – no 3rd party services are used
- The UI does not use Cookies – so no “consent” popup is required
- Wherever possible W3C standards will be followed
- The JavaScript libraries used (jQuery and Leaflet) will be stored locally
- All resources (CSS, images ...) will be stored locally
- The host webserver is used to provide server-side-includes for the header, footer and navigation. This webserver is currently IIS, but Apache and NGINX could provide the same functionality.
- No other code (Java, PHP etc.) will need to be installed on the webserver to deploy the UI.

This should allow the UI to be deployed on any webserver as long as it has HTTP network access to the API – which could be a local machine or a remote server anywhere in the world.

Folder Descriptions

Folder	File Type(s)	Contents
/resources/images/...	PNG, SVG, JPG & GIF	Logos, Country flags, UI icons
/resources/CSS/...	CSS	Core layout styles
/resources/cache/...	JSON	Cached country boundaries as GeoJSON
/resources/JS/...	JS	JavaScript Libraries & core JavaScript code
/SADC/...	HTML5	Webpage templates
/SADC/includes/...	HTM	HTML5 header, navigation and footer includes
/SADC/CSS/...	CSS	Website font and colours
/SADC/docs/...	XLSX, PDF	Downloads e.g. Excel data entry file
/SADC/js/...	JS	Webpage initialisation code

JavaScript Libraries Used

- **jQuery 3.4.1** – event handling and XHR requests
<https://jquery.com/>

- **Leaflet 1.5.1** – shows map on detail page
<https://leafletjs.com/>
- **Leaflet Lat/Lng Graticule Plugin** – adds grid overlay to map
<https://github.com/cloudybay/leaflet.latlng-graticule>

Application JavaScript

File	Description
agw.core.js	Wrapper for modular components
agw.api.js	Handles all API calls
agw.ln.js	Handles multi-lingual functionality
agw.detail.js	Handles display of a detailed record
agw.detail.map.js	Handles map shown for a detailed record
agw.searchForm.js	Handles interaction with search form
agw.phraseSearch.js	Handles phrase search functionality
agw.searchResults.js	Handles display of search results
agw.memberCountries.js	Handles display of country lists

Data Management

A download of the Excel sheet (as sent to the young professionals) will be available to download for those wishing to enter multiple records. A data entry form (with the same fields as the Excel file) is provided – data entered is serialized into an easy to “cut-and-paste” format that can be copied directly into the Excel sheet. New entries entered in this way will be emailed to SADC-GMI for QA. A “report an error” popup is available on each records’ detail page. This will allow users to report errors in the existing data. Error reports will be emailed to SADC-GMI for QA.

Map Service

The Leaflet map used on the document details page uses an open WMS map service provided by BRGM on behalf of OneGeology (<https://mapsref.brgm.fr/wxs/refcom-brgm/refign>). The topography data shown is sourced from Open Street Map.

API

Design

The API server implements a standard REST API:

- taking HTTP requests in the form of Uniform Resource Locators (URLs) from a client,
- using the parameters in those URLs to form database queries,
- querying the database,
- returning data as standard JavaScript Object Notation (JSON) text to the client.

Using the API

The API responds to a number of queries from clients. A query is formed by adding structured text to the base URL of the API.

The following queries are available:

- get a list of SADC countries
- get a list of keywords
- get the summary details of a document using its unique ID
- get the full details of a document using its unique ID
- get all the summary details for a single SADC country
 - search the database for the summaries that match any or all of a list of search terms:
 - country, keyword, author, title, report type
 - general phrase search
 - narrow by a range of years of publication + choose only those documents with available links
 - get the results in manageable batches

For each of these queries, the API returns a structured JSON response containing the data and any relevant metadata or any error information, if the query has failed. Full documentation of all the queries and their responses, including examples, can be found on the API documentation page of the server. Further documentation is available as an OpenAPI JSON document, though this is more useful to automated systems to aid in creating queries and parsing responses.

Software implementation

The API is implemented using the language Python 3 using a number of standard packages. The main package is FastAPI which is used to define routes (queries) and return defined models (responses) as JSON. FastAPI makes use of uvicorn to run an API service. FastAPI also automatically validates the incoming queries, returning errors if the queries are incorrect in some way. The package etlhelper is used to handle database connections and SQL queries, returning rows from the database in response to SQL queries.

The software is structured into a number of layers:

- route functions and supporting data functions
- model definitions
- database handlers
- SQL queries

For a typical query, the API will forward the query to the appropriate route. The route will automatically parse and validate the URL using FastAPI. If the query is correct, the route function structures any query parameters and asks its data function for a response. The data function forwards the structured query to a database handler. The handler uses the query parameters to build the appropriate SQL query and send that to the database using etlhelper.

Once the database has returned a number of rows in response to a query, the handler makes some structural changes and sends the list of rows back to the data function. The

data function uses the rows and constructs model objects from the data in those rows. These data items are added to a response model object and sent back to the route. The route then returns that response with FastAPI automatically transforming the model into JSON text before returning it to the client.

Server deployment

The server can be deployed in a number of ways. Once all of the Python packages are installed it can simply be run from the command line using uvicorn. However, it is more practical to deploy the server using a Docker container. Such a container is defined within the implementation and can be run in a number of ways.

There is some configuration mainly relating to the target database. While the current deployment uses Oracle it would be straightforward to switch to a different database system, PostgreSQL for example, by modifying the configuration and some of the SQL. The remainder of the code would not need altering.

Glossary

API - Application Programming Interface

HTTP - HyperText Transfer Protocol

JSON - JavaScript Object Notation

REST - Representational State Transfer

SQL - Structured Query Language

URL - Uniform Resource Locator

Database

Database overview

The database is implemented within an Oracle 11gR2 relational database management system and catalogues groundwater literature relevant to the Southern African Development Community (SADC). It contains general metadata about the literature including references for thousands of reports, articles, books, conference papers, policy documents, and maps. The database is one component of a 3-tiered architecture system which includes an API layer and GUI interface that offers a catalogued and searchable library including downloads either directly within the system or an external web link of groundwater literature relevant to SADC. The database offers extra metadata about the literature that enables the SADC-GLA system to be easily searchable by author, title, country and keywords.

The content of the database mainly originates from the existing Grey Literature Archive (GLA) archive that is maintained by the BGS. The document metadata are maintained in their original language in the catalogue but also implemented is an internationalised set of Keyphrase dictionaries (controlled vocabularies) in English, French and Portuguese that are used to tag the catalogue items to aid searching and retrieval from within the portal. New data as acquired will undergo QA and QC procedures and be included in the database as appropriate.

Database Structure

The database is implemented within an Oracle 11gR2 relational database system and compatible with other relational database systems. The database is a normalised design and contains a main index table (SGA_CATALOGUE_DETAIL) which holds the singular instance metadata about a catalogue item, and then links to other tables that capture metadata where there can be multiple instance(s) of various attributes pertaining to the catalogue item. These other attributes of a catalogue that could have more than one instance include: Authors, countries, keywords-keyphrases, Language of the catalogue document, links to external resources, IPR status and data coverage.

The database also employs denormalisation techniques to aggregate data outputs into views/query objects to serve as a database middle layer for easy access by the API and also for optimising query performance. This abstraction also allows for cases where the main database structure could be updated without affecting the live system until when it was appropriate to deploy the changes to the live system.

The database consists of 19 tables (entities) of which 11 tables are controlled vocabularies with subsets. There are also 11 query objects that form part of the database query/middle layer. The database also employs these object types as part of the design; sequences, triggers and views. The database objects are named according to a consistent naming standard across the whole data model.

The database also incorporates a full audit trail as part of its implementation with history tables and triggers. All changes to data are recorded including who and when that action was performed and for all new data entries, the user and date for that action are also recorded.

Dictionaries – Vocabularies

The database uses a number of controlled vocabularies (terms and their definitions) to maintain a level of consistent terminology across the database for the tagging of data. These terms are used to organise the information for subsequent easy retrieval and/or data entry. They are predefined, authorised terms, carefully selected list of words or phrases, which are used to tag the data so that they may be easily retrieved by a variety of search techniques.

These reduce the ambiguity inherent in free text tagging of data where the same term is consistently applied and reused across the database. They are also generally used to populate drop down lists in applications or checklist for the user to select for data entry as well as data retrieval.

The SADC-GLA database has implemented 11 dictionaries to cover the following topic areas: Keywords- key phrases, countries, language, confidentiality, IPR status, document types and the types of document links.

The key phrase dictionaries have been internationalised for 3 agreed languages; English, French and Portuguese to allow for tagging the catalogue items appropriately and their presentation on the web portal for searching and retrieval in each of these languages.

Table 1: Entities/Tables and their definitions

TABLE	DEFINITION
SGA_CATALOGUE_DETAIL	Table that holds the main metadata of a catalogue document that has relevance or association to one or more SADC countries
SGA_CAT_COUNTRY_LINK	Table that holds the list of countries that pertain to a document. That is the association between a document and country.
SGA_CAT_CATEGORY_LINK	Table that holds the list of relevant key phrases that are associated with the document
SGA_CATALOGUE_AUTHOR	Table that holds the list of authors (in publication order) against each document
SGA_CAT_LANGUAGE_LINK	Table that holds the original language(s) of a document
SGA_EXTERNAL_RESOURCE	Table that holds a link to the location of a document or to a page where such a link may be accessible
SGA_IPR_STATUS	Table that holds a reference to the IPR of a document
SGA_DATA_COVERAGE	Table that holds the geographical extent(s) that defines the area of influence of a document
<u>Dictionaries – Controlled Vocabularies</u>	
DIC_SADC_AGLA_COUNTRY	Dictionary table of country names including their ISO codes
DIC_SGA_KEYPHRASE	Dictionary table of groundwater key phrases. Used to tag catalogue items (documents) for searching and retrieval
DIC_SGA_KEYPHRASE_TOPIC	Dictionary table of key phrase topics. These are categorisations of the keyphrases in higher level terms for filtering/sub-setting in searches.
DICS_SADC_LANGUAGE	Dictionary table subset of languages.

SADC-GLA HIGH LEVEL DATA MODEL

Project	SADC-GLA HIGH LEVEL DIAGRAM DMT
File Name	SADC-GLA_HIGH_LEVEL_DIAGRAM.DMT
Submodel	Main Model
Author	
Company	
Version	1.0 Modified: 11/05/2023
Copyright	© 2023

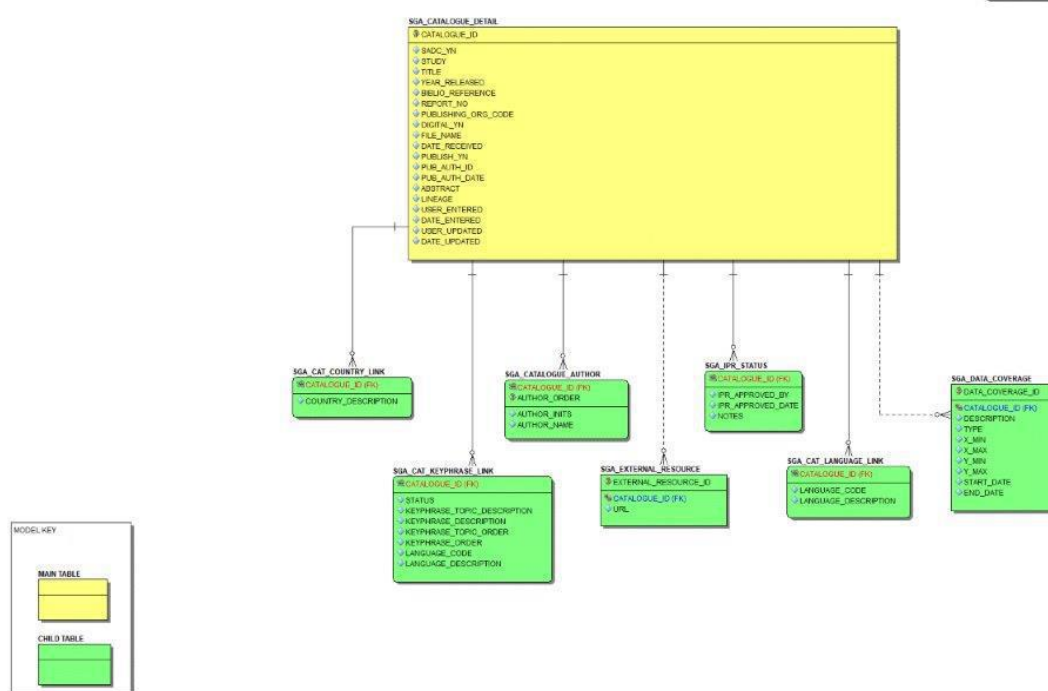


Figure 2: High level Data Model (only showing the high level entities).

Project	SADC-GLA
File Name	SADC_GLA_DEFAULT
SubModel	Main Model
Author	IGS
Company	IGS
Version	1.0
Modified	11/06/2020



60

The database is based on the relational model and implemented within an Enterprise Oracle 11gR2 relational database management system hosted at the BGS. The implementation of the database uses the ISO/ANSI SQL standard and compatible with most relational database management systems. Also program units have been defined in Oracle PL/SQL to enforce auditing as part of a full audit trail implementation.

The database includes the use of the following database object types: Tables, views, sequences and triggers.

The character set of the Oracle database is AL32UTF8 and runs on a Linux x86 64-bit platform. Also compatible on Windows platforms.

Details of the object creation SQL and PL/SQL code are all included in the Appendix.

Summary of Data in the Database

The data contained in the new SADC-GLA catalogue has been largely derived from the existing GLA archive, which has been added to by the BGS over many years to include a large number of relevant documents. Any new data acquired through the SADC-GLA work, will also undergo QA and QC procedures to be included in the database to grow this resource for SADC.

The data consists of a list of documents relevant to the SADC countries and is stored in a series of normalised database tables along with other tables. These tables hold metadata pertaining to these documents and are required to define them within a catalogue system that is both searchable and updateable.

A summary of the main data shown in the tables below for illustration. There are just over 4000 documents in the database (Table 2), mainly for SADC countries (Table 3), in 7 languages (Table 4) although the majority in English.

Table 2: Summary of Document Types

Document Types	Number of Documents
Reports	2124
Articles	1031
Book Sections	238
Maps	182
Conference Documents	176
To be categorised	110
Books	45
Theses	39
Other Documents	24

Dissertations	19
Digital Data Output	10
Policy Documents	6
Total	4004

Table 3: Documents by country

Country	Number of Documents
Botswana	731
South Africa	707
Zimbabwe	637
Malawi	431
Tanzania, United Republic of	352
Zambia	283
Madagascar	257
Namibia	235
Lesotho	157
Mozambique	122
Congo, The Democratic Republic Of The	99
eSwatini	98
Angola	88
Kenya	36
Mauritius	32
Uganda	31
Comoros	30
Seychelles	29
Ethiopia	28

Ghana	22
Sudan	15
Eritrea	13
Rwanda	12
Nigeria	12
Egypt	11
Burundi	11
Senegal	10
Djibouti	9
Niger	7
Mali	7
South Sudan	6
Saudi Arabia	4
Sierra Leone	4
Algeria	4
Cameroon	4
Burkina Faso	4
Morocco	3
Benin	3
Congo	3
Chad	3
Gabon	3
Tunisia	3
Cote d'Ivoire	3
Sri Lanka	3
Libya	2

Somalia	2
Gambia	1
Papua New Guinea	1
Reunion	1
Central African Republic	1
Netherlands Antilles	1
Guinea	1
Mauritania	1
Togo	1
Liberia	1

Table 4: Documents by language

NAME	No of Documents
English	3824
French	136
German	21
Portuguese	20
Afrikaans	2
Chichewa; Chewa; Nyanja	2
Spanish; Castilian	1

The data has been through several stages of QA and QC to remove duplication, correct errors, expand metadata and add value to the data. The new SADC-GLA system has been updated to reflect this extensive work.

Among some of the processes and procedures carried out were:

1. textual checks

- Language specific where possible
 - Structure of titles
 - Spelling/Rationalisation of Authors Names and initials
- 2. duplication detection and rationalisation (on-going)
 - Multi language duplication
 - Structural duplication (extra spaces, use of acronyms, punctuation etc.)
- 3. metadata assignment and checking
 - Countries
 - Key Word/Phrases
 - Original Language correction
- 4. Document links
 - validity
 - type

Updating the SADC Groundwater Literature Archive

Report on data collection, 28 September 2020, IGRAC

Introduction

The Southern African Development Community Groundwater Management Institute (SADC-GMI) is a regional centre of excellence for groundwater management, whose vision is to ensure the equitable and sustainable use and protection of groundwater in the SADC region. SADC-GMI has initiated a project to update the SADC Groundwater Literature Archive (SADC-GLA), to make it a fully functional online archive of groundwater literature for the SADC region. The project started in February 2020 and was planned to finish in June 2020, but the end of the project was postponed to October 2020 because of the Covid19 pandemics. The project was implemented by the British Geological Survey (BGS), in collaboration with the International Groundwater Resources Assessment Centre (IGRAC). The role of IGRAC was to migrate bibliographic records from the SADC-GIP and to engage young professionals in the SADC member states to collect additional records in their countries. This report provides an overview of these project activities. It outlines the achievements and the challenges met and provides recommendations for the development of the SADC-GLA in the next years.

Migration of data from the SADC-GIP and Ramotswa Information Management System

Bibliographic records were migrated from the SADC Groundwater Information Portal (SADC-GIP) and the Ramotswa Information Management System (RIMS). Both systems were hosted in the Global Groundwater Information System (GGIS), managed by IGRAC. In 2020, decision was made to migrate the map layers contained in these viewers to a new version of the SADC-GIP (<https://sadc-gip.org/>) and the bibliographic records to the SADC-GLA. In May-June 2020, 44 records from the SADC-GIP and 25 records from the RIMS were migrated to the SADC-GLA. IGRAC filled in the upload templates prepared by BGS with all 69 records. The job was done by two staff members, as a mean for testing the completeness, robustness and user-friendliness of the template, as well as the consistency of the inputs. Some minor flaws were identified in the template. Some inconsistencies were also identified between the inputs of the two staffs, regarding for instance the use of keywords, editors and publishers, as well as the use of organisation names as authors. The template was updated accordingly, where possible, however several inconsistencies resulted from subjective choices from the two staffs.

Young professionals' program

As part of another SADC-GMI project (Expansion of the SADC Groundwater Information Portal, SADC-GIP), young professionals from the national water departments in the member states were engaged to take part in several assignments and capacity-building activities. 25 young professionals were engaged, from 13 member states (Table 0-1). The same young professionals were engaged in the current project. DRC, Madagascar and Seychelles were invited again to nominate someone (they hadn't before), but they didn't react. The young professionals were tasked to collect documents in their countries to be

uploaded in the SADC-GLA. They were asked to fill in the upload templates and to write a short report explaining their approach and their efforts to collect documents. They were also asked to collect more than 50 documents in their countries. An IPR form was shared to collect the permission of the data providers for sharing the documents in the SADC-GLA. The original deadline for this activity was May 2020 but, because of the lockdown, it was pushed back to August 2020.

Table 0-1 Number of young professionals engaged per member state.

#	Member state	Number of young professionals
1	Angola	2
2	Botswana	2
3	Comoros	2
4	DRC	-
5	Eswatini	2
6	Lesotho	2
7	Madagascar	-
8	Malawi	2
9	Mauritius	1
10	Mozambique	2
11	Namibia	2
12	Seychelles	-
13	South Africa	2
14	Tanzania	2
15	Zambia	2
16	Zimbabwe	2
<i>Total</i>		25

In total, 476 new records were collected (Table 0-2). 59% of them were already available online, while 9% of them were made available for download in the SADC-GLA (IPR form completed). 19% of the documents were shared without IPR authorization and 13% records were shared without hyperlink or digital copy. Several young professionals reported issues with the IPR forms: they said that the data providers were either scared by the form or unsure whether they were entitled to sign it. Although the young professionals from Tanzania dropped out and several young professionals didn't meet the threshold of 50 records, the number of records collected in this activity was satisfying.

It is important to consider the challenges faced by the young professionals because of the pandemics. Most of them were locked at home, with poor internet connection. They were not able to visit the offices where archives are stored and to interview the colleagues from other organisations. These organizations could be engaged in the next years to contribute to the SADC-GLA.

Moreover, the SADC-GLA became available online only near the end of the project. The young professionals were introduced to the new system during a virtual workshop in August 2020. The workshop was very helpful to explain the purpose of the data collection and answer the questions from the young professionals, such as the use of the IPR form. It would have been easier for the young professionals if the training had happened before they start the assignment, and the overall quality and consistency of contributions might

have been better. The young professionals also noted that the online upload form is user-friendlier than the upload template in Excel. It is likely that the online form will become the main channel to collect new records in the future.

For the project team, it was challenging to provide clear instructions to the young professionals before the database was fully developed (e.g. two versions of upload template were shared, IPR form was shared after the upload template), which was the consequence of a tight workplan.

Next to these data collection challenges, IGRAC had a long and difficult time reviewing the contributions of the young professionals (several weeks in total). The main issues were the quality of the data entered in the upload templates, in particular the identification of publishers and the filename of the digital copies submitted (often too long and difficult to track). Incomplete or invalid records were corrected. These quality checks also useful to improve the consistency of all contributions. In the next years, it is expected that the administrators of the SADC-GLA will have to dedicate a significant amount of time for reviewing new entries.

Table 0-2 Number of records per member state.

Country	New records	Records with hyperlink	Records with digital copy		Records without hyperlink or digital copy
			with IPR consent	without IPR consent	
Angola	29	-	19	4	6
Botswana	29	-	-	-	29
Comoros	20	-	18	2	-
eSwatini	13	9	-	3	1
Lesotho	26	10	6	9	1
Malawi	43	31	-	12	-
Mauritius	57	32	-	-	25
Mozambique	43	35	-	9	-
Namibia	44	44	-	38	-
South Africa	59	59	-	-	-
Tanzania	-	-	-	-	-
Zambia	26	14	-	12	-
Zimbabwe	48	47	-	-	1
Total	476	281	43	89	63
Total (%)	100%	59%	9%	19%	13%

The final upload templates, digital copies, reports and IPR forms are available in Google Drive:

https://drive.google.com/drive/folders/15P_crxvdXV1GLF2ZEGRRykvqCryfqzhp?usp=sharing

Annex 6: Training Summary

SADC Groundwater Literature Archive Training Summary

Overview

An online training for the SADC-GLA took place on Wednesday 19th and Thursday 20th August 2020. The aim was to introduce the SADC-GLA web portal, database, and data update protocols to different stakeholders – managers, users, and contributors – across the region. The training was attended by 26 participants from 12 countries within the SADC region (Malawi, Namibia, Zambia, Botswana, Comoros, Eswatini, Lesotho, Mauritius, Mozambique, South Africa, Tanzania, and Zimbabwe), as well as Raquel Sousa (RS) from IGRAC and James Manda and Brighton Munyai from SADC-GMI. The training was led by Kirsty Upton (KU), Ken Lawrie (KL), Simon Burden (SB) and Martin Nayembil (MN) from BGS. The full training schedule is provided in Annex 1.

Summary of comments/issues raised

General Data Entry

- Most people, but not everyone, preferred the online form – it is important to keep both data entry methods
- Clarify data entry process for excel template and online form
 - Do SADC-GMI process the data from emails, or just check and forward emails to BGS?
 - Should emails from the online form be processed into a different format that provides direct entry to tables, rather than using the excel template as an intermediate step?
 - Clarify how users submit PDFs when using the online form – attach to emails? – and provide the IPR template as a download (also to be attached to the email)
 - ACTION: expand on workflow showing steps in data upload/entry process to be shared with SADC-GMI
- Clarify difference between contribution and IPR consent on the online form and excel template

Publishers

- The database and online form only allows one, while the excel template allows multiple

Keywords

- These are difficult to enter in the excel template
- ACTION: KL to include a dropdown menu for ALL keywords
- There may be some missing (participants to provide feedback)
- ACTION: KL to add guidance on how many should be entered

Online Form

- Requires more guidance/notes next to each field
 - ACTION: SB to add guidance based on pop-up notes in the excel template
- ACTION: SB to provide functionality to remove an individual author/keyword/country from the selection list

Excel Template

- Participants have difficulty editing the template due to locking
 - ACTION: KL/KU follow up with RS to discuss what kind of edits are required and explore whether these can be enabled
- Date formats – issues if the date formats don't match (e.g. DD/MM/YYYY or MM/DD/YYYY)
- The pop up guidance boxes get in the way
 - ACTION: KL to explore moving these (could also add note on front page that ESC closes the box)
- Many participants tried to replace the example rows
 - ACTION: KL to add guidance to start below these, or move to a separate sheet?

Data Quality

- Issues highlighted by IGRAC:
 - Confusion between author names and surnames
 - Completing the wrong fields for the document type
 - Limited keywords
 - Duplicates (only around 10%)
- A full list of issues highlighted by the participant exercise are provided in Annex 2
- Common issues include:
 - Spelling/punctuation etc
 - Missing data
 - Data in fields not relevant for the document type
 - Including author initials for an organisation
 - Too few keywords
 - Missing contributor name
 - IPR consent where no PDF is required or contributor is not the owner
 - Subjectivity with keywords

Summary of participant reflections and feedback

- The SADC-GLA is a very valuable resource for the region for sharing and accessing information – it has been significantly improved from the previous SADC Grey Literature Archive and Africa Groundwater Literature Archive
- The training provided insight into the new system
- Those who participated would like to act as trainers for colleagues, using the resources from the training session – can they be translated?
- The participant exercise was useful

Summary of trainer reflections

- The training was extremely useful from a user testing perspective – we gained a lot of very useful feedback on how the online form and template can be improved

- The training was well attended and there was a good level of participation – most completed the exercise and actively joined the discussion

Annex 1: Training Schedule

Time (SAST)	Time (BST)	Session Content	Session Lead
Wednesday 19th August			
10:30-10:50	09:30-09:50	Introduction Aims of the SADC-GLA, user requirements and general features of the new system	Kirsty
10:50-11:30	09:50-10:30	Introduction to the web portal Tour of the web portal, including search functionality, search results and user input functionality	Simon
BREAK			
11:45-12:30	10:45-11:30	Introduction to the database Introduction to the data model, metadata and QA issues	Martin & Ken
LUNCH			
14:00-15:00	13:00-14:00	Updating the SADC-GLA Introduction to data entry protocols. Introduction to participant task: each participant is given a set of documents to be entered through the online form and excel template	Simon & Kirsty
BREAK			
15:15-16:45	14:15-15:45	Updating the SADC-GLA (cont.) Participant task: participants do exercise individually; ZOOM meeting remains open for participants to ask questions; participants send completed excel templates when finished	Kirsty & Ken
Thursday 20th August			
11:00-11:45	10:00-10:45	Updating the SADC-GLA (cont.) Feedback on participant exercise	Kirsty & Ken
11:45-12:30	10:45-11:30	Updating the SADC-GLA (cont.) From the online form & excel template to the database – what is required?	Ken
12:30-13:00	11:30-12:00	Closing	Kirsty

Annex 2: Data Quality Issues from Participant Exercise

DOCUMENT 1: Botswana Integrated Water Resources Management & Water Plan

- Document Type
 - Report (IIIIIIIIII)
 - Policy Document
- Publisher
 - Government of Botswana (III)
 - UNDP, GEF
 - Department of Water Affairs T Ministry of Minerals, Energy & Water Resources (IIII)
 - Ministry of Mineral, Energy and Water Resources(I)
 - Centre for Applied Research and Aurecon(I)
- Author
 - None!
 - Department of Water Affairs... (III)
 - Centre for Applied Research and Aurecon Botswana (IIII)
 - Aurecon Botswana (I)
- Keywords
 - General Water Supply
 - IWRM (III)
 - General Groundwater Protection (II)
 - Drinking Water
 - Climate Change
 - General Sanitation
 - General Groundwater Quality
 - General Ecology
 - Livelihoods
 - Groundwater Resource Management (IIII)
 - Policy(I)
 - General Aquifer Characterisation
 - General Water Supply (I)
 - Legislation

DOCUMENT 2: Reconnaissance hydrogéologique pour l'alimentation en eau d'une plaine littorale en milieu semi-aride, Madagascar

- Keywords
 - Hydrology
 - General Engineering Hydrogeology
 - Salinity
 - Climate

DOCUMENT 3: Challenges in groundwater resource management in coastal aquifers of East Africa: Investigations and lessons learnt in the Comoros Islands, Kenya and Tanzania

- Keywords

- Groundwater Resource Management (IIII)
- Groundwater Resource Assessment
- Drinking Water
- Urban Groundwater
- General Groundwater Protection
- General Groundwater Quality (II)
- General Aquifer Characterisation (III)
- General Environmental Change (II)
- Vulnerability
- Population
- General Water Supply
- Geophysics
- Salinity
- Governance (II)
- General geology or Aquifer Type
- General Engineering Hydrogeology
- Authors
 - Missing Authors
- Journal Details
 - Missing journal title
 - Mixed up article title and journal title

DOCUMENT 4: CUVELAI-ETOSHA GROUNDWATER INVESTIGATION

- Publisher
 - Ministry of Agriculture, Water and Forestry (IIII)
 - Department of Water Affairs, Ministry of Agriculture... (III)
 - Bittner Water Consult (II)
 - Missing
 - Republic of Namibia
- Author:
 - Bittner Water Consult (IIIIII)
 - Department of Water Affairs, Forestry, Hydrology Division (III)
 - Department of Water Affairs, Ministry of Agriculture...
- Keywords
 - General Groundwater Quality (II)
 - General Development
 - General Water Supply (II)
 - Urban Groundwater
 - Groundwater Data
 - Groundwater Levels
 - Groundwater Resource Management
 - General Groundwater Protection
 - General Borehole Development
 - Geophysics
 - General Geology or Aquifer Type
 - Aquifer Properties
 - Groundwater Monitoring (II)

- Groundwater Resource Assessment (II)
- General Groundwater Exploration (II)
- General Aquifer Characterisation
- Geology
- Rainfall
- Hydrology

DOCUMENT 5: Striving for Professionalism in Cost Effective Boreholes Angola

DOCUMENT 6: GROUNDWATER DATA MANAGEMENT BY THE MALAWIAN MINISTRY OF IRRIGATION AND WATER DEVELOPMENT – A CASE STUDY

- Publisher
 - Missing
- Keyword:
 - Groundwater Data (II)
 - General Water Supply
 - Groundwater Resource Management (IIIIII)
 - General Groundwater Exploration
 - General Environment
 - General Governance (II)
 - General Groundwater Protection (II)
 - General Geology or Aquifer Type

GENERAL

- Issues where example rows have been replaced, rather than starting new rows below the examples
- Contribution dates
- Missing Document Entry Numbers
- Missing subtopic keywords
- Spelling mistakes
 - Changed capitalisation (data should be entered exactly with no changes)
 - Reason – can create duplication issues
 - makes searching harder by creating differences
 - Substitution (changing “&” for “and”)
- Additional punctuation/characters when copied from a document
- Data in columns not relevant to the document type
- For external resource – URL – document and IPR not required
- Don't give approval for IPR without consent from owner
- Digital Copy – only yes if providing a PDF – if entering a URL to another website where the doc is available, this (and Doc name) is not required
- For all fields with a drop down list you must only select from that list (no changes or additions)
- Remember to fill in Contributor Name/Date
- Missing Country
- Missing part of URL
 - Must give instruction on how to copy the link from the browser window bar!
- Organisation names in Column J only – not Initials Column (I)

- Great to get comments about documents!
- Missing URLs!
 - Must give instruction on how to copy the link from the browser window bar!
- Fill in all relevant columns!!!!
- Aim for more than 1 keyword – most publications will have the potential for more than one
- Where there are multiple (Authors, Keywords, Languages, Countries etc.) each must be on a new line.
- Don't leave spurious characters in the spreadsheet (if you put something in a field either temporarily or by mistake then remove it)
- DO NOT USE AUTOFILL especially in dropdown box controlled fields as it can allow non listed values to be entered
- DO NOT USE copy and paste when entering data into the drop down list fields as this will stop the checking.

SADC Grey Literature Archive Governance Strategy

September 2020

Authors:

Patrick Bell, Martin Nayembil & Kirsty Upton (BGS)

James Manda (SADC-GMI)

Table of Contents

Purpose of this Document.....	79
Project Background	79
Previous Governance Discussions.....	79
Key Components of Governance	80
Hosting & Maintenance	80
Upload of Data	80
Handover to SADC-GMI	81
Maintenance Contract	81
Cost of Maintenance	82
BGS Obligations.....	83
SADC-GMI Obligations	83
Data Entry Protocols	83
Direct Data Entry to the SADC-GLA	83
Entry of Data from the SADC-GIP	84
Links to Other Databases	84
Others (tbc by YPs)	84
Monitoring Use of the SADC-GLA.....	85
Contact persons.....	86
Annex 1: Data Entry Stage 1 (User to SADC-GMI)	87
Annex 2: Data Entry Stage 2 (SADC-GMI to BGS)	87
Annex 3: Data Entry Stage 3 (Entry to live database)	88

Purpose of this Document

This Governance Strategy outlines how the SADC Groundwater Literature Archive (SADC-GLA) will be hosted, maintained and updated beyond the end of the development project (from 31 October 2020 onwards). A maintenance contract will be in place between the British Geological Survey (BGS) and SADC Groundwater Management Institute (SADC-GMI) for a two-year period from 1 November 2020 – 31 October 2022. This Strategy defines the roles and responsibilities for BGS and SADC-GMI during the maintenance contract period, including the mechanisms by which new data are entered into the SADC-GLA. It also outlines the handover process to SADC-GMI at the end of the maintenance period. The Strategy has been prepared by the BGS and SADC-GMI with input from project partners, the International Groundwater Resources Assessment Centre (IGRAC).

Project Background

SADC-GMI is a regional centre of excellence for groundwater management, whose vision is to ensure the equitable and sustainable use and protection of groundwater in the SADC region. In 2019, SADC-GMI initiated a project to update the SADC Grey Literature Archive, to make it a fully functional online archive of groundwater literature for the SADC region. The project started in February 2020 and will finish in October 2020. It is being implemented by the BGS in collaboration with IGRAC.

In 2010, the BGS developed the SADC Grey Literature Archive⁵, an online catalogue with almost 3000 unpublished reports, which were searchable through a web portal by title, author and country. Since 2014, BGS has built upon this to develop the Africa Groundwater Literature Archive⁶ (AGLA), which has improved functionalities and contains almost 1000 additional documents for the SADC region. Since 2017, IGRAC has developed, with SADC-GMI, the SADC Groundwater Information Portal (SADC-GIP), a web portal for sharing groundwater related data and information in the SADC region. The SADC-GIP is being improved by IGRAC as part of another project with SADC-GMI.

The key aim of the SADC-GLA project is to provide access to groundwater literature in the SADC region to all relevant stakeholders, to support the equitable and sustainable use and management of groundwater in this region. To achieve this, the project team have developed a new system, which can be easily updated and searched, is technologically advanced, and contains up to date information from stakeholders across the SADC region.

Previous Governance Discussions

The **Terms of Reference** for the project, as defined by SADC-GMI, stated that the key objective is to “establish a fully-functional and well stocked archive that becomes a resource for researchers, sector players and all other stakeholders in the SADC region”. Following key tasks to update the GLA, the consultants were required to: “provide for user roles to be

⁵ <https://www.bgs.ac.uk/sadc/index.cfm>

⁶ <https://www.bgs.ac.uk/africagroundwateratlas/archive.cfm>

applied for different administrative levels of the archive” and “investigate the possible migration of the archive to a suitable SADC-GMI server”.

In response to this, the **Consultant’s Proposal** outlined a plan to develop a modern, flexible and portable database system, with protocols for adding, editing, and updating data, with consideration of different user roles, IPR, and quality control. The consultants agreed to host the updated system on the agreed infrastructure, with the system initially hosted by BGS, with the option to transfer to SADC-GMI if/when their capacity allows for this.

This was consolidated during the **Project Kick-off Meeting** (February 2020), where it was agreed that BGS would host the SADC-GLA in the short- to medium-term, with the ambition that it will be transferred to SADC-GMI in the long-term. It was also agreed that SADC-GMI are the managers of the GLA as a service for Member States, with James Manda and Brighton Munyai taking primary responsibility for the archive.

Governance was explored further during the User Consultation Process (Work Package 1), the results of which were summarised in the **Situational Analysis Report** (March 2020) and discussed at the Project Meeting on 17th March 2020. During this process, system users (represented by SADC-GMI staff) highlighted the need to:

1. Improve how the content of the SADC-GLA is maintained, including the addition of new documents and updates to existing entries within the database
2. Monitor how the website is used to inform future improvements

It was also agreed that:

1. It is beyond the scope of this project to develop a web front-end that provides users with a login and ability to upload directly to the BGS-hosted database.
2. BGS and IGRAC should explore options for linking the SADC-GLA with the SADC-GIP, which is potentially an ongoing source of documents/publications that could be included in the GLA.
3. The API allows other systems to link to the system and equally can link to others that are discoverable through an API.

Key Components of Governance

Hosting & Maintenance

As outlined above, it has been agreed that BGS will continue to host the SADC-GLA system for a two-year period, according to a maintenance contract (described below). At the end of this two-year period, the SADC-GLA will either be transferred to SADC-GMI, if their capacity allows this, or the contract will be renewed.

Ongoing hosting and maintenance of the web interface, API and database will involve responding to updates needed to address any security vulnerabilities that come to light and periodically moving to a new infrastructure in line with any step changes in BGS’ underlying protocols. It will not affect the functionality of the system.

Upload of Data

As outlined above, it is beyond the scope of the development project to provide remote access for users to upload data directly to the SADC-GLA database while it is hosted by

BGS because of the significant time required to develop a client-server or web-based data entry method, with the required security, quality assurance and quality control.

BGS will therefore continue to upload new data and documents, which have been fully checked and quality assured by SADC-GMI, and are provided through the data entry protocols outlined below. Data upload will be included in the maintenance contract, as described below.

Handover to SADC-GMI

An initial handover will take place at the end of the development project contract in October 2020, with full technical documentation provided to SADC-GMI. It was initially proposed that a training workshop would take place in South Africa, however given the current global Covid-19 pandemic this was not possible. Instead, a 1.5-day online training took place in August 2020, which was attended by 30 stakeholders from 12 countries across the SADC region, including SADC-GMI. The online training provided an introduction to the web portal, database and data entry methods, including detailed instructions on using the online form and excel template and in-depth discussion of metadata and data quality.

The second phase of the handover will take place when the full system is transferred to SADC-GMI at the end of the maintenance contract. As part of this handover, BGS will review SADC-GMI's technical capacities, package the database and application accordingly, and provide guidelines and training on installation, ongoing maintenance, and data upload to the system. Budget is included in the maintenance contract for a longer, in-country, technical training as part of this handover.

Maintenance Contract

BGS will host, maintain and upload data to the SADC-GLA according to a maintenance contract. The maintenance contract will be based on the following:

- BGS will host the web portal, API, and database on our own servers for two years, at which point the contract will either be reviewed and renewed or terminated;
- If the contract is terminated after two years, SADC-GMI will take full responsibility for the system;
- The contract can be terminated earlier than this if SADC-GMI are in a position to take responsibility for hosting and maintaining the GLA;
- Termination will require at least 3 months' notice to allow time for the database and application to be packaged with updated guidelines for installation and preparation for in-country training;
- During the maintenance contract period, ongoing hosting and maintenance of the web interface, API and database will involve responding to updates needed to address any security vulnerabilities that come to light and periodically moving to a new infrastructure in line with any step changes in BGS' underlying protocols; it will not involve any change to the system functionality;
- During the maintenance contract period, data upload will also be carried out by BGS; data will be provided to BGS according to the data entry protocols outlined below; this data should be fully checked and quality assured to help maintain the integrity of the database; depending on the frequency of new data being provided, BGS will either upload new data on an ad hoc basis, or at fixed 2 monthly intervals if this becomes unmanageable;
- When SADC-GMI take full responsibility for hosting and maintaining the SADC-GLA at the end of the maintenance contract, BGS will provide support with the

handover of the system; the technical requirements of the handover will be informed by SADC-GMI's internal capacities; the database and application, with updated guidelines for installation, will be packaged accordingly.

Cost of Maintenance

Unspent reimbursables from the development contract (Table 1) will be transferred into the maintenance contract to cover part of the cost of hosting, maintenance, data entry and final handover, with the remainder funded by BGS. The full costs associated with hosting, maintenance, data entry and final handover are shown in Table 2, with the costs for SADC-GMI and BGS clearly defined.

Item	Unit Cost	Quantity	Total (USD)
Per diem allowances	266	22	5,852
International flights	1463	4	5,852
In/out airport transportation	322.5	4	1,330
IT Infrastructure			4,655
TOTAL			17,689

Table 1: Unspent reimbursables from the current contract, largely due to travel restrictions imposed by Covid-19

Item	Unit Cost	Quantity	Total (USD)
FUNDED INTERNALLY BY BGS			
Staff time for hosting & maintenance	\$569.32 per day	8 days (4 per year)	4,554.56
Staff time for data entry/upload	\$569.32 per day	8 days (4 per year)	4,554.56
Staff time to package the system for handover	\$569.32 per day	10 days	5,693.20
BGS Total			14,802.32
FUNDED BY SADC-GMI THROUGH THE MAINTENANCE CONTRACT			
Training workshop: staff time	\$569.32 per day	20 days	11,386.40

Training workshop: reimbursables	\$6302.60	1	6,302.60
SADC-GMI Total			17,689.00
TOTAL COST			32,688.72

Table 2: Total cost of hosting, maintenance, data entry and final handover of the SADC-GLA.

BGS Obligations

As outlined above, BGS will commit to hosting and maintaining the web portal, API and database to ensure it continues to operate with its existing functionality. BGS will also upload new data to the database according to the protocols outlined below. At the end of the project, BGS will provide full documentation for the new system and when the time comes to handover the system to SADC-GMI, BGS will provide the necessary support and training to SADC-GMI staff.

SADC-GMI Obligations

Although BGS will host and maintain the SADC-GLA, it is agreed that SADC-GMI are the managers of the GLA as a service for Member States. SADC-GMI therefore have primary responsibility for promoting the use of the SADC-GLA – both by website users looking for documents and those contributing documents – and for the quality of data held within the archive. Encouraging stakeholders within the region to use and continue to contribute to the archive beyond the end of the project will be key to its success in the long-term. As outlined in the Data Entry Protocols below, SADC-GMI will also take responsibility for the QA of new data being provided to BGS for upload.

Data Entry Protocols

Direct Data Entry to the SADC-GLA

There are two routes for data entry into the SADC-GLA – an online form and downloadable excel template – available through the Contribute page of the SADC-GLA web portal (www.sadc-gla.org/SADC/contribute.html). Data entry protocols for each of these routes are described below and illustrated in the Workflow diagrams in Annex 1 and 2.

1. Online form

Any user can enter the metadata details for a single document, which on submission are passed to a private API, generating an email to SADC-GMI (info@sadc-gmi.org). SADC-GMI convert the data from the email into the excel template format (see below) using a simple tool accessed through the web portal. The data can then be copied directly into the excel template for QA by SADC-GMI staff. The QA'd template is then forwarded to BGS for upload into the live database. Users can provide electronic copies of documents by attaching them as PDFs to the email to SADC-GMI, as prompted through the online form. PDFs must be accompanied by a contribution consent form, confirming they have permission to share the document through the SADC-GLA. PDFs and consent forms are forwarded to BGS with the excel template and are stored on our servers alongside the database.

2. Excel Template

Users can download an excel template directly from the web portal, enter metadata details for multiple documents, then email to SADC-GMI (info@sadc-gmi.org). As above, data within the excel template are QA'd by SADC-GMI staff, then forwarded to BGS for upload to the live database. Users can provide electronic copies of documents by attaching them as PDFs to the email or using an online file transfer system, depending on file size. As above, PDFs must be accompanied by a contribution consent form, confirming they have permission to share the document(s) through the SADC-GLA. PDFs and consent forms are forwarded to BGS with the excel template and are stored on our servers alongside the database.

For both data entry routes described above, BGS will perform a further check of the data before loading to the live database (see Workflow diagram in Annex 3).

There is a “report an error” function on the portal that allows users to report mistakes or errors in existing data and provide corrections via a simple form. As for the data entry form, this will be passed through a private API generating an email to the SADC-GMI team to sense-check, before forwarding to the BGS data management team to be updated in the database.

When the GLA is fully handed over to SADC-GMI at the end of the maintenance contract, BGS will provide an in-house tool for direct data entry and management. This will be the same system as used internally within BGS to update the GLA during our hosting period. There will be a bulk upload function and a client-server implementation, most likely in a tool like Microsoft Access or similar, where an administrative user can login and manage the records in the database directly, record by record, and will be able to enter individual new records including document upload. This solution will be available to install on users' terminals, who have connectivity to the hosted database and whom SADC-GMI wish to have administrative access to update the database records directly. It will not be available as a web tool but only as an install on individual computer terminals.

Entry of Data from the SADC-GIP

The new SADC-GIP system provides functionality for document upload. While the primary focus of the SADC-GIP is on map/spatial data, the upload function is not able to differentiate between different types of documents (e.g. an excel document containing data and PDF report) and it is not therefore possible to prevent users from uploading reports etc into the SADC-GIP. The two systems will therefore require a level of maintenance to identify documents in the GIP that should be transferred to the GLA. This could be done monthly by SADC-GMI staff. During the maintenance contract, BGS/IGRAC will look to develop a simple protocol that will allow documents to be easily transferred from the SADC-GIP to the SADC-GLA.

Links to Other Databases

The SADC-GLA should remain linked to the BGS-hosted Africa Groundwater Literature Archive so that updates to both systems are discoverable by all users. This will be the case while both systems are hosted by BGS, but if/when the SADC-GLA is handed over to SADC-GMI, a methodology will have to be put in place to ensure they remain interoperable. This will be possible through the API of the SADC-GLA and AGLA.

Others (tbc by YPs)





- Still unclear what these will be – YPs will be providing this information

- Method for linking will have to be decided on a case-by-case basis
- Time for linking to other databases may be limited depending on input from the YPs, but the API provides functionality for this to happen in future if required

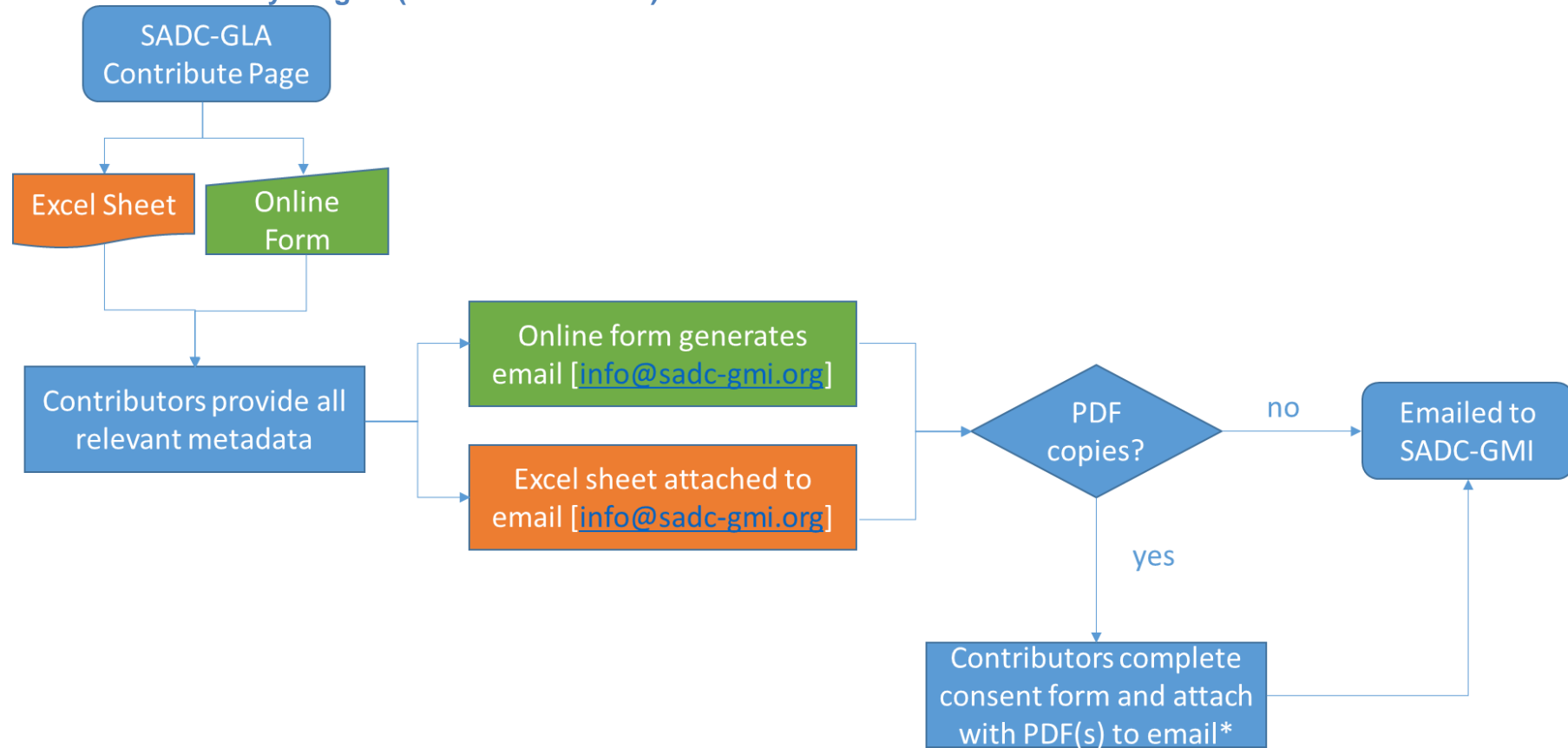
Monitoring Use of the SADC-GLA

BGS/IGRAC/SADC-GMI should monitor the usage of the GIP and GLA over the next 18 months to make recommendations on how they should be maintained in the long-term. The team will make use of Google Analytics and create a dashboard that can be accessed by SADC-GMI, BGS and IGRAC.

Contact persons

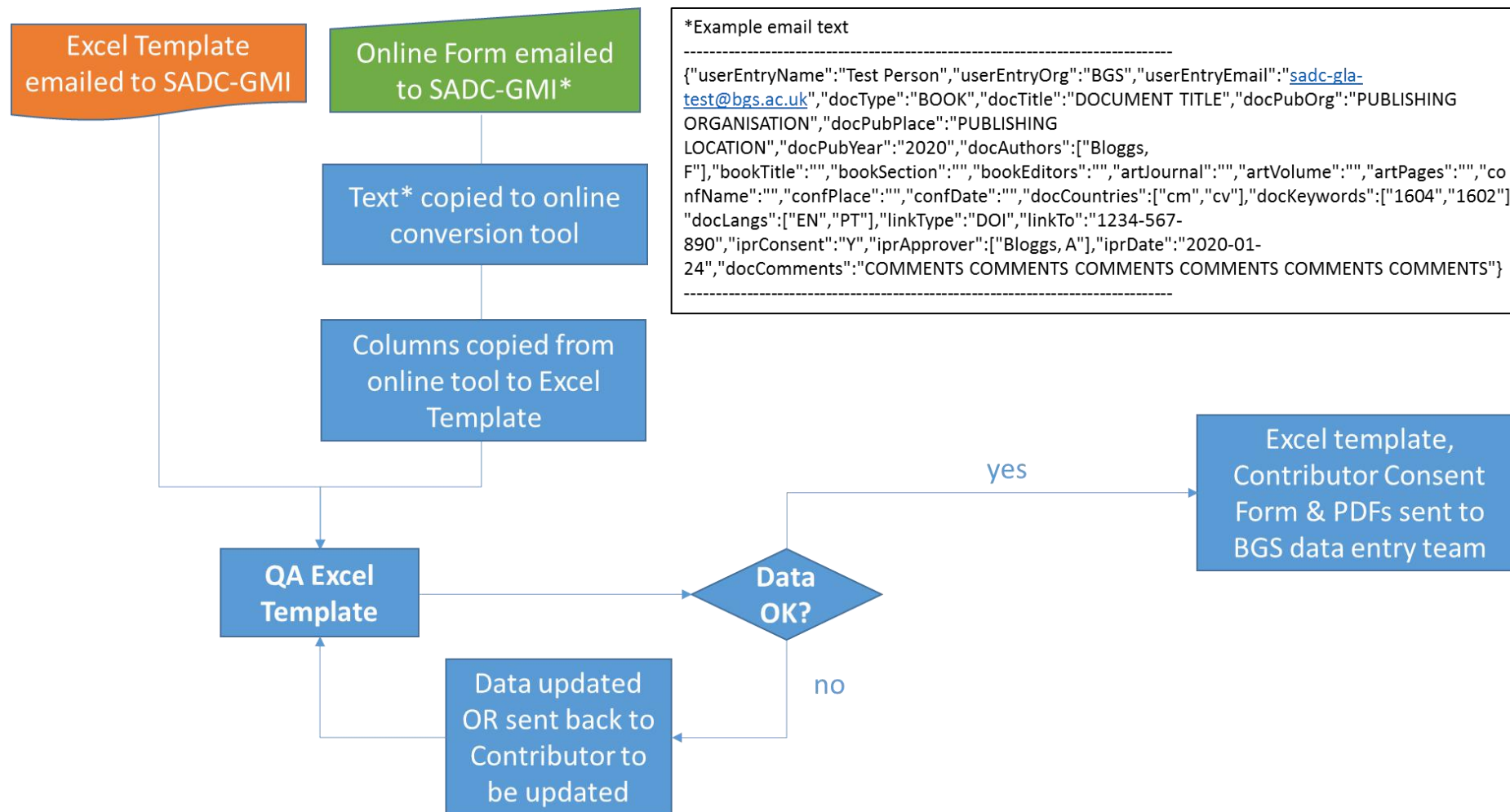
  GROUNDWATER MANAGEMENT INSTITUTE	<p>SADC-Groundwater Management Institute (SADC-GMI)</p> <p>Mr. James Manda (Project Coordinator SADC-GMI)</p> <p>jamesm@sadc-gmi.org</p> <p>+265 991742582</p>
 British Geological Survey Expert Impartial Innovative	<p>British Geological Survey (BGS)</p> <p>Dr. Kirsty Upton (Team leader)</p> <p>kirlto@bgs.ac.uk</p> <p>+ 44 131 6500 210</p>
 International Groundwater Resources Assessment Centre	<p>International Groundwater Resources Assessment Centre (IGRAC)</p> <p>Dr. Arnaud Sterckx</p> <p>arnaud.sterckx@un-igrac.org</p> <p>+31 6 82104674</p>

Annex 1: Data Entry Stage 1 (User to SADC-GMI)



*If a Contributor has too many PDFs to send by email, this should be stated in the email and SADC-GMI should contact the Contributor to arrange online document transfer

Annex 2: Data Entry Stage 2 (SADC-GMI to BGS)



Annex 3: Data Entry Stage 3 (Entry to live database)

