Mise à jour des archives de littérature grise de la SADC Rapport final

Mise à jour des archives de littérature grise de la SADC Rapport final

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

Rapport commercial BGS CR/20/066

Octobre 2020

Référence bibliographique :

Upton, K., Nayembil, M., Burden, S., Lawrie, K., Blackburn, C., Sterckx, A., Sousa, R. & Simister, A. 2020. Mise à jour des archives de littérature grise de la SADC : Final Report. Rapport commercial du British Geological Survey, CR/20/066. 90pp.

Avant-propos

Ce rapport est le produit final d'un projet commercial mené par le British Geological Survey (BGS) et le Centre international d'évaluation des ressources en eaux souterraines (IGRAC), au nom du client, l'Institut de gestion des eaux souterraines de la Communauté de développement de l'Afrique australe (SADC-GMI). Il résume les principales activités et les résultats du projet - principalement le développement d'une nouvelle archive de littérature sur les eaux souterraines pour la région de la SADC - la SADC-GLA. Le projet a été dirigé par le BGS. Le développement de la nouvelle base de données a été dirigé par Martin Nayembil, Ken Lawrie et Andrew Simister ; le développement de l'interface de programmation d'application et de l'interface utilisateur a été dirigé par Patrick Bell, Colin Blackburn et Simon Burden ; le projet a été géré par Kirsty Upton. Les nouvelles données ont été compilées par Arnaud Sterckx et Raquel Sousa de l'IGRAC, en collaboration avec des jeunes professionnels de toute la région.

Remerciements

Outre le personnel du BGS et de l'IGRAC mentionné dans l'avant-propos, l'équipe du projet tient à souligner la contribution du personnel de la SADC-GMI, en particulier le chef de projet James Manda. Les 25 jeunes professionnels suivants, originaires de toute la région SADC, ont joué un rôle crucial dans la réussite de ce projet en collectant de nouvelles données dans leurs États membres respectifs : 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 et Thandilizwe Bengeza. L'équipe tient également à remercier Alan MacDonald (BGS) pour son rôle de conseiller dans le projet, Carl Watson (BGS) pour sa contribution au processus de consultation des utilisateurs, et Antonio Ferreira (BGS), Ahmed Zeggan et Claudinei Nunes da Silva pour leurs services de traduction.

Contenu

Avant-propos3
Remerciements3
1 Introduction5
2 Contexte5
3 Aperçu du projet5
4 Structure du projet, jalons & livrables5
5 Résumé de la consultation des utilisateurs (WP1)7
6 Résumé des données et examen du système (WP1)7
7 Résumé de la mise à jour du système (WP2)8
8 Résumé de la mise à jour des données (WP2)8
9 Résumé de la formation (WP3)9
10 Stratégie à long terme (WP3)9
11 Réflexions et recommandations10
12 Équipe du projet & Personnes de contact11
Annexe 1 : Résumé de la réunion de lancement du projet & Procès-verbal12
Annexe 2 : Rapport d'analyse situationnelle19
Annexe 3 : Mise à jour du projet41
Annexe 4 : Documentation du système49
Annexe 5 : Rapport sur la collecte de données66
Annexe 6 : Résumé de la formation69
Annexe 7 : Stratégie de gouvernance77

FIGURES

Figure 1 Aperçu de l'architecture de la SADC-GLA

1 Introduction

L'Institut de gestion des eaux souterraines de la Communauté de développement de l'Afrique australe (SADC- GMI) est un centre d'excellence régional pour la gestion des eaux souterraines, dont la vision est d'assurer l'utilisation et la protection équitables et durables des eaux souterraines dans la région de la SADC. En 2019, la SADC-GMI a lancé un projet de mise à jour des archives de littérature grise de la SADC, afin d'en faire des archives en ligne entièrement fonctionnelles de la littérature sur les eaux souterraines à l'usage de toutes les parties prenantes des eaux souterraines dans la région de la SADC. Le projet a débuté en février 2020 et se terminera en octobre 2020. Il a été mis en œuvre par le British Geological Survey (BGS), en collaboration avec le Centre international d'évaluation des ressources en eaux souterraines (IGRAC).

Ce rapport fournit un résumé du processus et des résultats du projet.

2 Contexte

En 2010, BGS a développé les archives de littérature grise de la SADC, un catalogue en ligne avec près de 3000 rapports non publiés, qui étaient consultables via un portail web par titre, auteur et pays. En 2014, le BGS a commencé à développer les archives de littérature sur les eaux souterraines en Afrique (AGLA), qui s'appuient sur les SADC-GLA, offrant une fonctionnalité améliorée et près de 1000 documents supplémentaires pour la région SADC. Depuis 2017, l'IGRAC et le SADC-GMI ont développé le portail d'information sur les eaux souterraines de la SADC (SADC-GIP), un portail web pour le partage des données et des informations relatives aux eaux souterraines dans la région de la SADC. Le SADC-GIP est en cours d'amélioration par l'IGRAC dans le cadre d'un projet parallèle avec le SADC-GMI.

3 Aperçu du projet

L'objectif principal de ce projet était de fournir un accès à la littérature sur les eaux souterraines dans la région de la SADC à tous les acteurs concernés, afin de soutenir l'utilisation et la gestion équitables et durables des eaux souterraines dans cette région. Pour atteindre cet objectif, l'équipe du projet a développé un nouveau système - les archives de la littérature sur les eaux souterraines de la SADC (SADC-GLA) - qui peut être facilement mis à jour et recherché, est technologiquement avancé et contient des informations à jour provenant des parties prenantes de la région de la SADC. Le SADC-GLA est accessible via un nouveau portail web :

www.sadc-gla.org

Le SADC-GLA consiste en une nouvelle base de données, contenant environ 4.500 références à la littérature sur les eaux souterraines pour la région SADC. Beaucoup d'entre elles sont disponibles pour les utilisateurs, soit sous forme de PDF téléchargeables, soit par des liens vers des sites web externes. Le portail web (adresse ci-dessus) accède à la base de données par le biais d'une API, ce qui permet une interopérabilité et une flexibilité pour le développement et l'hébergement futurs. Les utilisateurs de la région ont été initiés au nouveau système par le biais d'un atelier de formation en ligne, et une stratégie de gouvernance à long terme pour le SADC-GLA a été développée et lancée avec le SADC-GMI.

4 Structure du projet, jalons & livrables

Le projet a été structuré en 3 paquets de travail :

1. Consultation des utilisateurs et examen des données et du système actuels (WP1)

2. Développement et mise en œuvre d'un nouveau système pour répondre aux besoins identifiés dans le WP1 dans le respect des contraintes de temps et de budget du projet (WP2).
3. Formation des utilisateurs dans la région SADC et développement d'une stratégie de gouvernance à long terme pour assurer la mise à jour et la maintenance du système après la fin du projet (WP3).

Le WP1 a été réalisé en février et début mars 2020. Une réunion de lancement a eu lieu le 5 février, à laquelle ont participé des représentants de SADC-GMI et l'équipe du projet BGS. Lors de cette réunion, le champ d'application du projet a été discuté et convenu, y compris : (1) l'identification des principaux groupes d'utilisateurs de la SADC-GLA (représentés par la SADC-GMI pour les besoins de la consultation des utilisateurs) ; (2) la confirmation que le BGS devra héberger le système à court et moyen terme, avec une stratégie développée à long terme.(2) confirmation que le BGS devra héberger le système à court et moyen terme, et qu'une stratégie sera élaborée pour la gouvernance à long terme du système ; (3) accord sur les termes de référence des jeunes professionnels de la région, qui seront chargés de collecter de nouvelles données pour le SADC-GLA, sous la direction de l'IGRAC. Un résumé et un compte-rendu de la réunion se trouvent à l'annexe 1.

Le processus de consultation des utilisateurs, résumé dans la section 5 ci-dessous, a été mis en œuvre par le biais d'un atelier le 14 février, auquel ont participé la SADC-GMI, l'IGRAC et l'équipe du projet BGS. En parallèle, un examen du système AGLA existant et des données de la SADC a été entrepris, comme résumé dans la section 6. AGLA a été examiné à la place de l'ancienne archive de littérature grise de la SADC, car il s'agissait d'un système plus récent en termes de conception technique et de contenu. Les résultats du WP1 ont été détaillés, ainsi que les recommandations pour la conception de la nouvelle SADC-GLA, dans un rapport d'analyse de situation, remis le 11 mars et fourni en annexe 2.

Le WP2 a démarré fin mars - avec un léger retard dû au développement de la pandémie mondiale - par un atelier interne de deux jours, auquel ont participé l'équipe du projet BGS et l'IGRAC. L'objectif de cet atelier était de : (1) convenir des décisions finales sur la conception du nouveau système SADC-GLA, y compris la base de données, l'API et le portail web ; (2) définir les tâches spécifiques requises, le calendrier et les responsabilités au sein de l'équipe ; (3) convenir des aspects clés du modèle de gouvernance. Les résultats de cet atelier de projet ont été résumés dans un rapport de mise à jour du projet, remis le 31 mars et présenté à l'annexe 3.

Le développement du nouveau système s'est poursuivi en avril et en mai, et le système de démonstration a été mis en ligne le 18 mai pour être examiné par le SADC-GMI. Les commentaires formulés à cette occasion ont été pris en compte et la première version du nouveau système SADC-GLA, décrite à la section 7, a été lancée le 30 juin, en même temps que la documentation technique (fournie à l'annexe 4).

Parallèlement au développement du nouveau système, IGRAC a dirigé l'activité de collecte de données, qui a fourni près de 500 nouveaux documents aux archives, comme résumé dans la section 8 et détaillé dans l'annexe 5.

Le WP3 a consisté à organiser un atelier de formation en ligne de deux jours pour les parties prenantes de toute la région, y compris le personnel de la SADC-GMI. L'atelier de formation a impliqué l'utilisation du système par un plus grand nombre de personnes que celles qui avaient été précédemment impliquées dans les tests, fournissant un retour d'information précieux qui a été traité au cours des semaines

suivantes. Les résultats sont résumés à la section 9, et un rapport complet figure à l'annexe 6.

Une stratégie de gouvernance a été finalisée en septembre, décrivant comment le système serait hébergé, maintenu et mis à jour à l'avenir. Cela a conduit à l'établissement d'un contrat de maintenance entre SADC-GMI et BGS pour une période de deux ans à partir de novembre 2020, après quoi

Le système sera remis à la SADC-GMI. Ceci est résumé dans la section 10 et la stratégie de gouvernance complète est fournie dans l'annexe 7.

Le projet devait initialement se terminer en juin 2020, mais en raison de la pandémie mondiale, une prolongation de contrat de trois mois a été accordée, en grande partie à cause des difficultés rencontrées par les jeunes professionnels pour accéder aux données et en recueillir de nouvelles. Les restrictions aux déplacements dans le monde ont également entraîné une révision des activités de formation, avec une formation en ligne plus courte dans le cadre de ce projet et une formation plus longue dans le pays prévue dans le cadre de la remise finale du système à la fin du contrat de maintenance de deux ans.

5 Résumé de la consultation des utilisateurs (WP1)

Les premières discussions sur le processus de consultation des utilisateurs ont eu lieu lors de la réunion de lancement du projet le 5 février (voir Annexe 1), au cours de laquelle il a été convenu que les partenaires du projet SADC-GMI, IGRAC et BGS pourraient remplir les rôles d'administrateurs du site web ainsi que de représentants des utilisateurs finaux réels du SADC-GLA. Un atelier de recherche sur les utilisateurs a eu lieu le 14 février via skype, auquel ont participé le BGS, l'IGRAC et le SADC-GMI. Miro - un tableau blanc en ligne pour la collaboration visuelle - a été partagé avant l'atelier et utilisé pendant l'atelier pour discuter des différents publics cibles, du parcours de l'utilisateur (en utilisant AGLA comme modèle), de l'administration et du suivi du site, et des questions de DPI.

L'analyse du tableau Miro après l'atelier a permis de définir les principaux besoins des utilisateurs, qui ont été utilisés pour la conception du nouveau système, comme indiqué dans le rapport d'analyse de la situation (annexe 2). Un résumé de ces exigences est fourni ci-dessous :

- Le portail web doit être facile à trouver à l'aide des moteurs de recherche courants.
- La fonctionnalité de recherche doit être simple et intuitive
- Il doit être facile de filtrer les entrées de la base de données pour lesquelles un document est disponible.
- Le plus grand nombre de documents possible devrait être disponible/accessible.
- Tous les documents doivent pouvoir être découverts à l'aide d'une série de critères de recherche.
- Il doit exister un processus simple et bien défini pour la mise à jour ou l'ajout de nouveaux documents aux archives.
- Le système doit être interopérable avec d'autres bases de données documentaires.

6 Résumé de l'examen des données et du système (WP1)

Un examen complet du système AGLA, y compris les données existantes pour la région SADC, la mise en œuvre technique, la fonctionnalité et la présentation a également été entrepris dans le cadre du WP1. Les détails complets de cet examen sont fournis dans le rapport d'analyse de la situation (annexe 2). Certaines des principales conclusions et recommandations de cet examen sont les suivantes :

- Il y avait un nombre important d'entrées dans AGLA sans PDF téléchargeable ou lien vers un site externe ; d'autres métadonnées étaient également manquantes.
    - Toutes les données existantes doivent faire l'objet d'un contrôle qualité général, la base de données doit être normalisée et davantage de contraintes de

dictionnaire doivent être ajoutées pour permettre un meilleur contrôle des métadonnées.

- Le dictionnaire de mots-clés était très granulaire et les termes de recherche les plus utilisés étaient des noms de pays et des termes de recherche très généraux liés à l'eau et à la géologie (par exemple, eau, eaux souterraines, hydrogéologie, géologie).
    - Le dictionnaire des mots-clés doit être rationalisé

- La plupart des termes de recherche sont saisis dans la zone de texte du titre, plutôt que d'utiliser des critères de recherche individuels.
  - Un champ de recherche simple de type Google devrait être mis en place
- Un faible pourcentage de documents a été géoréférencés (mais presque tous avec des informations sur le pays).
  - Cela ne devrait pas être une priorité et la recherche sur la carte devrait être supprimée.
- L'architecture était ancienne et peu flexible, utilisant une combinaison de HTML, CSS, JavaScript et ColdFusion.
  - Ce système devrait être mis à jour pour devenir un système logiciel à trois niveaux, avec des niveaux de présentation, d'application et de données (comme décrit plus en détail ci-dessous).
- Il n'y avait pas de moyen facile pour les utilisateurs de fournir de nouvelles données.
  - Il convient d'élaborer un projet de processus, qui peut évoluer au fil du temps à mesure que le modèle de gouvernance est formalisé.

7 Résumé de la mise à jour du système (WP2)

Le nouveau portail SADC-GLA est mis en œuvre sous la forme d'un système logiciel à trois niveaux comprenant des niveaux de présentation, d'application et de données (figure 1). Dans le système actuel, le niveau de présentation est un portail web et le niveau de données est un système de base de données relationnelle Oracle 11g, bien qu'il puisse facilement être migré vers un système open source. Le niveau d'application, qui sert d'intermédiaire entre le portail Web et la base de données, est un serveur d'interface de programmation d'applications (API) REST (Representational State Transfer).

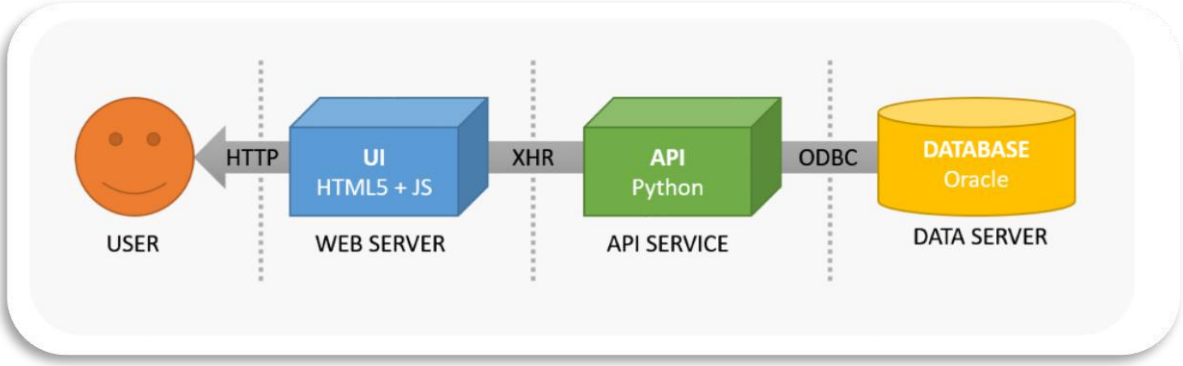


Figure 1 Aperçu de l'architecture de la SADC-GLA

L'avantage de cette approche est que les trois niveaux ne sont que faiblement couplés entre eux. Si la structure de la base de données change, l'application peut gérer ces changements sans affecter le portail web. La conception du portail Web peut être modifiée, ou un nouveau portail ou une nouvelle application peut être développée, sans qu'il soit nécessaire de modifier l'application. L'API peut également être étendue pour fournir des requêtes supplémentaires sans affecter la fonctionnalité existante. La documentation technique du système est fournie à l'annexe 4 (à noter qu'elle est en anglais, mais que des versions française et portugaise sont également disponibles).

8 Résumé de la mise à jour des données (WP2)

Les données existantes dans la SADC-GLA ont été soumises à un contrôle de qualité (QA) par l'équipe du BGS - il s'agit d'un processus continu. Des protocoles pour l'ajout de nouvelles données, initialement à l'aide d'un modèle Excel puis en ajoutant des fonctionnalités à l'aide d'un formulaire en ligne, ont été élaborés comme défini dans la stratégie de gouvernance (voir ci-dessous).

La compilation de nouvelles données pour le SADC-GLA a été dirigée par IGRAC et avait 2 composantes :

1. Migration des données du SADC-GIP et du Ramotswa Information Management System (RIMS)
2. Compilation de nouvelles données par de jeunes professionnels dans les États membres de la SADC

Les données provenant de la SADC-GIP et du RIMS ont été compilées en premier lieu et utilisées pour tester le processus de saisie des données à l'aide du modèle Excel (le formulaire en ligne n'avait pas encore été développé à ce stade). Au total, 69 nouveaux enregistrements ont été ajoutés au SADC-GLA à partir de ces sources.

Après cette première saisie de données, le modèle Excel a été mis à jour et fourni à 25 jeunes professionnels de 16 États membres de la région SADC (la RDC, Madagascar et les Seychelles n'ont pas participé au processus). Au total, 476 nouveaux enregistrements ont été compilés par les jeunes professionnels, avec une quantité importante d'AQ entreprise par le personnel du projet IGRAC (voir l'annexe 5 pour plus de détails).

L'équipe du BGS a utilisé les nouvelles données fournies par l'IGRAC pour développer une méthodologie de saisie des données dans la base de données - celle-ci continuera à être affinée pendant la période de maintenance de deux ans avant la remise finale du système.

9 Résumé de la formation (WP3)

Une formation en ligne pour le SADC-GLA a eu lieu les mercredi 19 et jeudi 20 août 2020. L'objectif était de présenter le portail web SADC-GLA, la base de données et les protocoles de mise à jour des données aux différentes parties prenantes - gestionnaires, utilisateurs et contributeurs - de la région. La formation a été suivie par 26 participants de 12 pays de la région SADC (Malawi, Namibie, Zambie, Botswana, Comores, Eswatini, Lesotho, Maurice, Mozambique, Afrique du Sud, Tanzanie et Zimbabwe), ainsi que par Raquel Sousa de l'IGRAC et James Manda et Brighton Munyai du SADC-GMI. La formation était dirigée par l'équipe du projet BGS. Un résumé complet de la formation est fourni dans l'annexe 6 et tous les matériaux de formation sont disponibles via un lecteur google :

https://drive.google.com/drive/folders/12jPUMcq7A7mzJv9ChO36N9kBYTqUpWcn?usp=sharing

10 Stratégie à long terme (WP3)

Comme mentionné ci-dessus, il a été convenu au début du projet que, en raison du manque de capacité, le BGS devrait héberger et maintenir le SADC-GLA à court et moyen terme avant de le transférer au SADC-GMI. Les termes de cette période de maintenance et de transfert sont définis dans une stratégie de gouvernance co-développée par BGS, SADC-GMI et IGRAC (voir Annexe 7).

La stratégie de gouvernance est décrite :

- Définitions pour l'hébergement et la maintenance
- Protocoles pour la saisie des données pendant la période de maintenance
- Coût de l'entretien

- Transfert à la SADC-GMI
- Obligations de la BGS et de la SADC-GMI

Suite à la stratégie de gouvernance, un contrat de maintenance de deux ans a été établi entre le BGS et SADC-GMI, financé en partie par les fonds de voyage non dépensés du projet actuel et en partie par le BGS dans le cadre de notre mandat d'APD, à partir de novembre 2020.

11 Réflexions et recommandations

Le projet SADC-GLA a été entrepris dans des circonstances difficiles pour toutes les personnes impliquées. Le calendrier du projet était déjà très serré et la pandémie a encore aggravé la situation, la plupart des membres du personnel ayant dû gérer la transition vers le travail à domicile et une réduction des capacités au cours des deux premiers mois du projet. Même en l'absence d'une pandémie mondiale imprévue, un délai plus long aurait permis de consulter davantage les utilisateurs dans le cadre d'un processus de développement plus itératif.

Malgré les défis, un nouveau système SADC-GLA a été développé selon les exigences spécifiées, avec plus de 500 nouvelles entrées, et avec une stratégie en place pour sa gouvernance à long terme. La période de maintenance de deux ans permettra au BGS d'affiner les protocoles de saisie des données avant le transfert final. Elle permettra également à la SADC-GMI de développer la capacité requise pour héberger et maintenir le système à l'avenir.

Les réactions des utilisateurs pendant la formation en ligne ont été très positives, avec une reconnaissance de l'utilité d'une telle ressource pour le développement et la gestion durables des eaux souterraines dans la région. Le succès futur du système dépendra en grande partie de l'adhésion continue et accrue d'une série de parties prenantes, y compris les jeunes professionnels. En tant que propriétaire du système, on s'attend à ce que SADC-GMI continue à engager et à encourager les parties prenantes à utiliser et à contribuer au système afin qu'il reste une ressource actualisée et pertinente dans le futur.

La pandémie était manifestement très prohibitive pour les activités de collecte de données des jeunes professionnels et un engagement continu avec eux pourrait potentiellement permettre d'obtenir davantage de données à mesure que les restrictions sont assouplies. Ils pourraient également être impliqués dans l'atelier de formation de la remise finale afin de les faire participer à nouveau plus tard.

On s'attend à ce que la SADC-GMI prenne des dispositions pour un contrôle de qualité suffisant des nouvelles données, ce qui peut prendre du temps, afin de s'assurer qu'un niveau élevé de métadonnées est maintenu pour que les nouveaux documents soient facilement accessibles via le portail web.

L'une des questions clés pour la SADC-GLA est celle des DPI. En tant que développeurs du système, BGS et IGRAC ont fait tout leur possible pour s'assurer que les droits d'auteur n'ont pas été violés par le partage de documents sans les autorisations requises. Les commentaires des jeunes professionnels ont indiqué que certains fournisseurs de données étaient effrayés par le formulaire de permission DPI ou n'étaient pas sûrs d'être autorisés à le signer. Un engagement plus poussé de la SADC-GMI avec certains des principaux fournisseurs de données de la région pourrait potentiellement aider à surmonter ce problème à l'avenir.

Le nouveau SADC-GLA a été conçu comme un système flexible, ouvert et interopérable, qui a le

potentiel d'être relié à d'autres plateformes de connaissances dans le futur. Lorsque le système sera remis au SADC-GMI à la fin de la période de maintenance, il est prévu qu'un certain niveau de lien soit maintenu avec les Archives Africaines de Littérature sur les Eaux Souterraines afin que les mises à jour pertinentes pour la région de la SADC, qu'elles soient obtenues par le SADC-GLA ou l'AGLA, soient accessibles par les deux systèmes. Il pourrait également être possible de relier le système à d'autres initiatives régionales ou continentales à l'avenir, afin que l'investissement important réalisé au cours de ce projet ait un impact aussi large que possible.

12 Équipe du projet & Personnes de contact

L'équipe de projet était composée de cadres supérieurs et de personnel scientifique et technique de soutien du BGS et de l'IGRAC. Les principaux experts ont de nombreuses années d'expérience dans le développement de systèmes agiles et de bases de données en ligne et dans le travail sur des projets liés aux eaux souterraines avec des partenaires de la région SADC.

	Institut de gestion des eaux souterraines de la SADC (SADC-GMI) M. James Manda (Coordinateur de projet SADC-GMI) jamesm@sadc-gmi.org +265 991742582
	Commission géologique britannique (BGS) Dr. Kirsty Upton (chef d'équipe) kirlto@bgs.ac.uk + 44 131 6500 210
	Centre international d'évaluation des ressources en eaux souterraines (IGRAC) Dr. Arnaud Sterckx arnaud.sterckx@un-igrac.org +31 6 82104674

Annexe 1 : Résumé de la réunion de lancement du projet & Procès-verbal

Mise à jour de la réunion de lancement du projet d'archives de littérature grise de la SADC : Résumé & Procès-verbal

5 février 2020

Résumé du projet

L'Institut de gestion des eaux souterraines de la Communauté de développement de l'Afrique australe (SADC-GMI) est un centre régional d'excellence pour la gestion des eaux souterraines, dont la vision est d'assurer l'utilisation et la protection équitables et durables des eaux souterraines dans la région de la SADC. La SADC-GMI a lancé un projet visant à mettre à jour les archives de littérature grise de la SADC (SADC-GLA), pour en faire des archives en ligne entièrement fonctionnelles de la littérature sur les eaux souterraines pour la région de la SADC. Le projet a débuté en février 2020 et se terminera en juin 2020. Il sera mis en œuvre par le British Geological Survey (BGS), en collaboration avec le Centre international d'évaluation des ressources en eaux souterraines (IGRAC).

En 2010, le BGS a développé les archives de littérature grise de la SADC (SADC-GLA1), un catalogue en ligne avec près de 3000 rapports non publiés, qui étaient consultables via un portail web par titre, auteur et pays. Depuis 2014, le BGS s'est appuyé sur la SADC-GLA pour développer les archives de littérature sur les eaux souterraines en Afrique (AGLA2), dont les fonctionnalités ont été améliorées et qui contiennent près de 1000 documents supplémentaires pour la région SADC. Depuis 2017, l'IGRAC a développé avec le SADC-GMI le portail d'information sur les eaux souterraines de la SADC (SADC-GIP), un portail web pour le partage de données et d'informations relatives aux eaux souterraines dans la région de la SADC. Le SADC- GIP est en cours d'amélioration par l'IGRAC dans le cadre d'un autre projet avec SADC-GMI.

L'équipe du projet améliorera les fonctionnalités et le contenu de la SADC-GLA. Un nouveau système sera développé, qui peut être facilement mis à jour et recherché, qui est technologiquement avancé, et qui contient des informations à jour provenant des parties prenantes de la région SADC. L'objectif principal de ce projet est de fournir un accès à la littérature sur les eaux souterraines dans la région de la SADC à tous les acteurs concernés, afin de soutenir l'utilisation et la gestion équitables et durables des eaux souterraines dans cette région.

Résumé de la réunion de lancement

Une réunion de lancement du projet a eu lieu le 5 février 2020, à laquelle ont participé des représentants de SADC-GMI et l'équipe du projet BGS. Le compte-rendu détaillé de la réunion est disponible ci-dessous.

Au cours de cette réunion, trois principaux groupes d'utilisateurs du SADC-GLA ont été identifiés : les universitaires, les décideurs politiques et les consultants. La SADC-GMI agit en tant que propriétaire du produit et mandataire de ces utilisateurs, il a donc été convenu d'organiser un atelier de suivi de consultation des utilisateurs avec l'équipe de la SADC-GMI, afin de définir les besoins des utilisateurs de manière plus détaillée, de sorte que ceux-ci puissent être reflétés dans la conception du nouvel AGL. L'atelier de consultation des utilisateurs a été programmé pour le 14 février, dont les résultats sont discutés dans le résultat du Work Package 1 : Rapport d'analyse situationnelle.

Il a également été convenu lors de cette réunion que le BGS hébergera le SADC-GLA mis à jour à court et moyen terme, avec l'ambition de le transférer au SADC-GMI à long terme. Ce point sera abordé lors de l'élaboration du modèle de gouvernance au cours du Work Package 2 du projet.

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

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

La collecte de données à l'aide du modèle des jeunes professionnels a fait l'objet d'une première discussion et il a été convenu qu'un mandat devait être rédigé pour les jeunes professionnels avant l'atelier de consultation des utilisateurs du 14 février.

Procès-verbal de la réunion de lancement du projet

Lieu Salle du Conseil de l'IGS et par Skype

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

Participants

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)

Résumé des actions clés

1. **Tous participeront à la prochaine réunion du vendredi 14 février pour poursuivre la consultation sur les besoins des utilisateurs.**
2. **Le processus des jeunes professionnels doit être lancé par la rédaction du cahier des charges.**
3. **Contrats à finaliser**

Ouverture et introduction

- JM a souhaité la bienvenue à tous et les présentations ont été faites

Portée du projet

- JM a souligné la portée de ce projet - mettre à jour l'AGL existant en liaison et en collaboration avec les États membres de la SADC - la visibilité de l'AGL de la SADC est une exigence clé.

Détail de l'archive actuelle de la littérature sur les eaux souterraines en Afrique

- KU a souligné la fonctionnalité existante de l'archive de littérature sur les eaux souterraines en Afrique, qui contient environ 4000 entrées pour la région SADC.
- KU a souligné que le DPI est le principal obstacle à l'autorisation des PDF téléchargeables dans AGLA - ce point sera abordé au cours du projet.
- ACTION : KL vérifiera combien d'entrées de la SADC ont actuellement un PDF ou un lien externe disponible.

Modules de travail et principaux résultats attendus

- KU a présenté les modules de travail et les principaux résultats attendus (voir annexe 1).

- Tous sont d'accord pour dire que le contenu du GLA sera déterminé par la littérature disponible, plutôt que par les besoins des utilisateurs.
- Il existe trois principaux groupes d'utilisateurs, qui peuvent avoir des besoins différents en termes d'accès à la GLA :
    - Academia
    - Décideurs politiques (principalement au sein du gouvernement)
    - Consultants (secteur privé/industrie)
- La GLA sera hébergée par la BGS à court et moyen terme, avec l'ambition d'être transférée à la SADC-GMI à long terme - cette question sera abordée lors du développement du modèle de gouvernance et pourrait impliquer, par exemple, un protocole d'accord entre la SADC-GMI et la BGS.
- SADC-GMI sont les gestionnaires du GLA en tant que service pour les États membres - JM et BM sont les principaux contacts.
- ACTION : TOUS se joindront à la réunion du vendredi 14 février afin de poursuivre la consultation pour définir les besoins des utilisateurs, sous la direction de CW.
- ACTION : CW envoie quelques questions initiales au SADC-GMI avant la prochaine réunion.
- Le modèle des jeunes professionnels (YP) devrait être initié dès que possible ; l'IGRAC devrait être impliqué dans la réunion du 14 février et, avant cela, un cahier des charges pour les YP devrait être rédigé (question : utiliseront-ils les mêmes YP que pour le projet GIP ?)
- ACTION : KU assure le suivi avec IGRAC et SADC-GMI pour commencer à rédiger le mandat et le plan de travail du PJ.

Clarifications et modifications

Contractuel

- ACTION : KU fournira les coordonnées bancaires de BGS pour le contrat, qui pourra ensuite être contresigné par SADC-GMI.
- Tout le monde est d'accord pour que la SADC-GMI gère le budget des Jeunes Parlementaires (6400 USD) - ceci devrait être retiré des Remboursables avant la contre-signature.
- Les remboursements seront effectués sur présentation des factures.
- Tous les participants se sont mis d'accord sur les livrables/étapes, comme indiqué à l'annexe 1.

Termes de référence

- Tous sont satisfaits du cahier des charges, tel que reflété dans les lots de travaux et les produits livrables (annexe 1).

Réunions d'avancement

- Cette réunion aura lieu tous les mois, après la réunion initiale du 14 février.

Toute demande d'information

- Aucune demande d'information supplémentaire de la part du BGS
- SADC-GMI demande un projet de Termes de référence pour les YPs - voir action ci-dessus.

Prochaines étapes immédiates, tâches et résultats attendus

- Résumé de la réunion initiale (ACTION : KU)
- Réunion le vendredi 14 février pour poursuivre le processus de consultation avec SADC-GMI (ACTION : TOUS)
- Lancer le modèle YP (ACTION : KU, IGRAC & SADC-GMI)

Date de la prochaine réunion

- Vendredi 14 février

ANNEXE 1 GROUPES DE TRAVAIL ET PRINCIPAUX PRODUITS À FOURNIR

Dossiers de travail

1. Comprendre les besoins des utilisateurs pour un système mis à jour et évaluer la fonctionnalité du GLA/AGLA existant.
2. Améliorer la fonctionnalité de la GLA pour répondre aux besoins identifiés dans le WP1 dans les limites de temps et de budget du projet.
3. Former les utilisateurs de la région SADC et développer une stratégie à long terme pour assurer la mise à jour et la maintenance du système au-delà de la fin du projet.

Paquet de travail 1

- Consultation avec SADC-GMI pour définir les besoins des utilisateurs et commencer à développer un modèle de gouvernance.
    - Réunion de lancement avec SADC-GMI (en ligne)
    - OUTPUT &gt ; Rapport de synthèse (14 février)
- Examen des données existantes & système
    - OUTPUT &gt ; Rapport d'analyse situationnelle (28 février)
- Lancer un modèle pour les jeunes

professionnels Paquet de travail 2

- Mise à jour des données existantes (AGLA & SADC-GIP)
- Nouvelles données
- Mise à jour de la base de données et migration du contenu
- Définir les protocoles d'ajout et de modification des données
- Développer l'API et le portail web
- Héberger les archives sur une infrastructure convenue
- OUTPUTS
    - Rapport d'avancement (31 mars)
    - Système fonctionnel avec nouveau portail web (15 mai)
    - Rapport final (30 juin)

Work Package 3

- Fournir un manuel d'utilisation et une formation au personnel de la SADC.
- OUTPUTS
    - Manuel de l'utilisateur (30 juin)
    - Formation (30 juin)

Annexe 2 : Rapport d'analyse situationnelle

Mise à jour des archives de littérature grise de la SADC Rapport d'analyse situationnelle

Mars 2020

Les auteurs :

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

Table des matières

Introduction21
Contexte21
Objectifs et portée21
Structure du projet21
Consultation des utilisateurs22
Approche et méthodologie22
Principaux besoins des utilisateurs23
Examen du système actuel24
Examen des données24
Examen du système : Mise en œuvre techniqueErreur ! Signet non défini.
Examen du système : Fonctionnalité et présentation26
Révision du système : Statistiques d'utilisation du site web AGLA pour 201932
Révision du système : Analyse des termes de recherche des utilisateurs33
Recommandations34
Mises à jour des données34
Mises à jour du système logiciel34
Mises à jour liées à la recherche36
Mises à jour liées à l'administration36
Modèle Jeunes Professionnels37
Horaire de travail39
Équipe de projetErreur ! Signet non défini.
Personnes de contactErreur ! Signet non défini.

Introduction

L'Institut de gestion des eaux souterraines de la Communauté de développement de l'Afrique australe (SADC-GMI) est un centre régional d'excellence pour la gestion des eaux souterraines, dont la vision est d'assurer l'utilisation et la protection équitables et durables des eaux souterraines dans la région de la SADC. Le SADC-GMI a lancé un projet de mise à jour des archives de littérature grise de la SADC (SADC-GLA), afin d'en faire des archives en ligne entièrement fonctionnelles de la littérature sur les eaux souterraines pour la région de la SADC.
Le projet a débuté en février 2020 et se terminera en juin 2020. Il sera mis en œuvre par le British Geological Survey (BGS), en collaboration avec le Centre international d'évaluation des ressources en eaux souterraines (IGRAC).

Ce rapport présente les principales conclusions d'un processus de consultation rapide, mené avec la SADC-GMI pour définir les principales exigences des utilisateurs du nouveau SADC-GLA. Il résume un examen du système existant, par rapport à ces exigences, et fait des recommandations pour des améliorations.

Contexte

En 2010, le BGS a développé les archives de littérature grise de la SADC (SADC-GLA3), un catalogue en ligne avec près de 3000 rapports non publiés, qui étaient consultables via un portail web par titre, auteur et pays. Depuis 2014, le BGS s'est appuyé sur la SADC-GLA pour développer les archives de littérature sur les eaux souterraines en Afrique (AGLA4), dont les fonctionnalités ont été améliorées et qui contiennent près de 1000 documents supplémentaires pour la région SADC. Depuis 2017, l'IGRAC a développé avec le SADC-GMI le portail d'information sur les eaux souterraines de la SADC (SADC-GIP), un portail web pour le partage de données et d'informations liées aux eaux souterraines dans la région de la SADC. Le SADC- GIP est en cours d'amélioration par l'IGRAC dans le cadre d'un autre projet avec SADC-GMI.

Objectifs et portée

L'objectif principal de ce projet est de fournir un accès à la littérature sur les eaux souterraines dans la région de la SADC à tous les acteurs concernés, afin de soutenir l'utilisation et la gestion équitables et durables des eaux souterraines dans cette région. Pour y parvenir, l'équipe du projet développera un nouveau système, qui peut être facilement mis à jour et recherché, qui est technologiquement avancé, et qui contient des informations à jour provenant des parties prenantes de la région SADC.

Structure du projet

Le projet est structuré en 3 paquets de travail :

4. Module de travail 1 - Examen des données et du système actuel
5. Work package 2 - Améliorer la fonctionnalité de la SADC-GLA pour répondre aux besoins identifiés dans le WP1 dans les limites de temps et de budget du projet.

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

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

6. Module de travail 3 - Former les utilisateurs de la région SADC et développer une stratégie à long terme pour assurer la mise à jour et la maintenance du système après la fin du projet.

Le présent rapport expose les résultats du premier lot de travaux, qui comporte les éléments suivants :

1. Consultation des utilisateurs : comprendre les besoins et les frustrations des utilisateurs par rapport au système actuel.
2. Examen du système actuel :
    a. Données
    b. Mise en œuvre technique
    c. Fonctionnalité et présentation
3. Recommandations : comment le système devrait être mis à jour pour mieux répondre aux besoins des utilisateurs et pour permettre une maintenance plus facile à l'avenir (le point central du work package 2).

En outre, le module de travail 1 avait pour objectif de lancer le modèle des jeunes professionnels en tant que mécanisme de soutien de la durabilité à long terme de la SADC-GLA (l'objectif du module de travail 3).

Consultation des utilisateurs

Approche et méthodologie

Les activités suivantes ont été entreprises pour comprendre les besoins des utilisateurs et leurs frustrations par rapport au système actuel :

- Les premières discussions sur le processus de consultation des utilisateurs ont eu lieu lors de la réunion de lancement du projet au cours d'une vidéoconférence Skype le 5 février 2020.
- Il a été convenu que, pour les besoins de la consultation des utilisateurs, les partenaires du projet SADC-GMI, IGRAC et BGS pourraient remplir les rôles d'utilisateurs administrateurs du site Web ainsi que de mandataires pour les utilisateurs finaux du monde réel qui recherchent des documents dans les archives documentaires.
- Un atelier de recherche sur les utilisateurs a été programmé pour le 14 février, avec la participation de BGS, IGRAC et SADC-GMI.
- Afin de faciliter l'atelier sur des sites physiques répartis, les partenaires du projet ont convenu d'utiliser Miro, un tableau blanc en ligne pour la collaboration visuelle, et Skype.
- Le tableau Miro a été créé avant la réunion et distribué avec les instructions préalables à l'atelier.
- Pendant l'atelier, au moins une personne par site s'est connectée au tableau Miro et a contribué de manière proactive à la session, voir la figure 1.
- Au cours de l'atelier, les points suivants ont été discutés :
    - Audience(s) cible(s)
    - Le parcours de l'utilisateur, y compris :
        - Optimisation des moteurs de recherche
        - Conception du formulaire de recherche
        - Métadonnées / détails du document

        - Recherche par carte
    - Administration et suivi du site
    - Questions relatives aux DPI

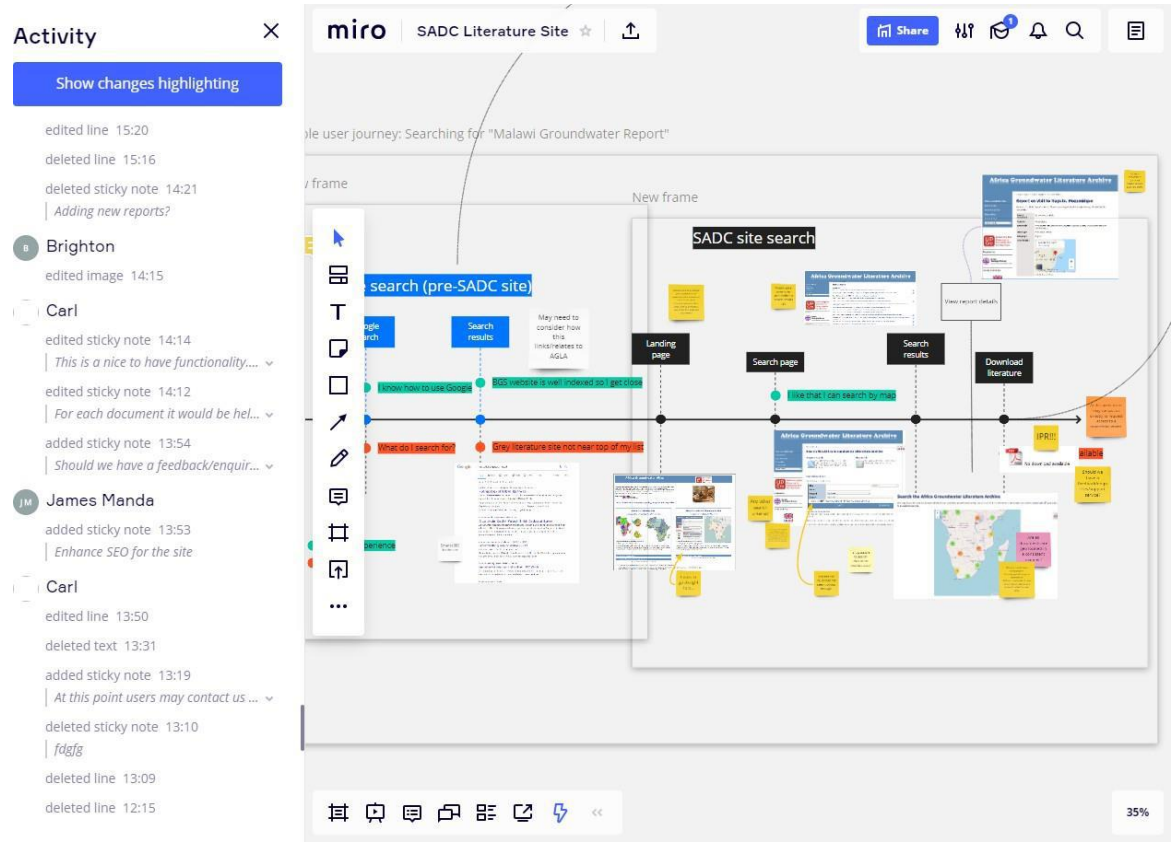


Figure 1 : Tableau Miro utilisé dans l'atelier de recherche sur les utilisateurs.

Dans les jours qui ont suivi l'atelier, le tableau et le contenu de Miro ont été analysés pour définir les principaux besoins des utilisateurs décrits ci-dessous. Enfin, les principales conclusions et recommandations ont été produites afin de guider la conception de la nouvelle version du site Web et de la base de données associée, comme indiqué dans la section Recommandations ci-dessous.

Principaux besoins des utilisateurs

Utilisateurs de sites web à la recherche de documents

Bien qu'il existe une multitude d'utilisateurs externes, notamment des conseillers politiques, des constants commerciaux, des chercheurs universitaires et autres, il est apparu au cours de l'atelier qu'il y avait peu ou pas de différence significative dans leurs exigences.

Objectifs

- Pour trouver les documents relatifs aux eaux souterraines pour la région de l'Afrique australe

Frustrations

- Difficultés à trouver le site web de la SADC
- Vous ne savez pas s'il faut utiliser la carte ou le formulaire de recherche.
- Difficultés à savoir quelles parties du formulaire de recherche utiliser

- La case à cocher qui permet de filtrer les résultats ne contenant aucun document accessible n'est pas signalée.
- Il arrive régulièrement que les métadonnées d'un document indiquent "document non disponible".

Opportunités

- Il y a peut-être beaucoup plus d'utilisateurs potentiels que ceux qui consultent actuellement les sites de la SADC ou de l'AGLA.
- De nombreux documents signalés comme non disponibles sont en fait disponibles ailleurs et nous pourrions aider les utilisateurs à y accéder.
- Si le nouveau site est plus agréable à utiliser, les utilisateurs actuels l'utiliseront peut-être plus souvent.

Administrateurs de sites web

Il serait possible de les subdiviser en catégories plus ciblées, mais au cours de l'atelier, de nombreuses discussions ont porté sur la nécessité d'améliorer la manière dont le contenu du site est maintenu à l'avenir et sur l'intérêt de surveiller l'utilisation du site Web pour informer les améliorations futures.

Objectifs

- Améliorer la manière dont les documents sont décrits afin d'en améliorer la découverte et l'utilité.
- Pour ajouter de nouveaux documents aux archives
- Accroître et contrôler l'impact des archives

Frustrations

- Aucun moyen pour les utilisateurs extérieurs à l'équipe de la base de données BGS d'ajouter de nouveaux documents.
- Aucun moyen pour les utilisateurs autres que l'équipe de la base de données BGS d'améliorer les métadonnées.
- La raison pour laquelle tant de documents sont à accès restreint n'est pas claire
- Il n'est pas clair comment résoudre les problèmes de droits de propriété intellectuelle liés à la mise à disposition d'un plus grand nombre de documents.
- Je ne sais pas combien d'utilisateurs visitent le site ou ce qu'ils recherchent.

Opportunités

- Souhaite rendre les archives de la SADC interopérables avec d'autres bases de données de documents associés.
- Un plus grand nombre de partenaires pourrait ajouter du nouveau contenu, ce qui rendrait les archives plus grandes et plus à jour.

Les recommandations concernant la manière d'améliorer le système actuel pour répondre à ces conclusions sont abordées plus loin dans le rapport.

Examen du système actuel

Afin de comprendre quel travail était nécessaire pour améliorer la mise en œuvre du système, les activités d'examen suivantes ont été entreprises :

1. Examen des données
2. Révision du système : Mise en œuvre technique
3. Examen du système : Fonctionnalité et présentation

Toutes les entrées de la SADC-GLA et de l'AGLA ont été examinées afin de déterminer le niveau d'information qui existe actuellement pour chaque entrée dans les archives. Cet examen identifie le nombre total d'entrées pour la région SADC - à la fois celles de l'archive originale de littérature grise SADC et de l'AGLA - et identifie les principales métadonnées manquantes.

L'examen des données a permis de catégoriser les entrées existantes dans la base de données afin de fournir un résumé des métadonnées comme suit :

	SADC	AGLA
Nombre total d'entrées	2997	988
Entrées avec des mots-clés (%)	89%	71%
Entrées avec auteurs (%)	99%	95%
Entrées avec langue (%)	4%	91%
Entrées avec pays (%)	92%	100%
Entrées avec géoréférencement (%)	2%	35%
Entrées avec lien vers le document Live (%)	14%	44%
Entrées avec lien vers un site externe (%)	18%	41%

Les mots-clés sont semi-hiérarchiques et ont des groupes de parents. Le dictionnaire de mots-clés est étendu et très granulaire. Les étendues géographiques sont classées de deux façons : par pays et par emplacement (avec géoréférencement). La classification par pays dispose d'un dictionnaire et presque toutes les entrées comprennent cette information. Le géoréférencement ne se fait par point ou par boîte de délimitation et est beaucoup moins complet dans l'archive actuelle. Un peu plus de 40 % des entrées ont un lien direct vers un document, et une proportion similaire contient des liens vers un site externe. Le BGS dispose de PDF pour 178 autres documents, qui ne sont pas disponibles pour des raisons de droits de propriété intellectuelle. L'examen initial a permis d'identifier plusieurs doublons dans les archives.

Revue du système : Mise en œuvre technique

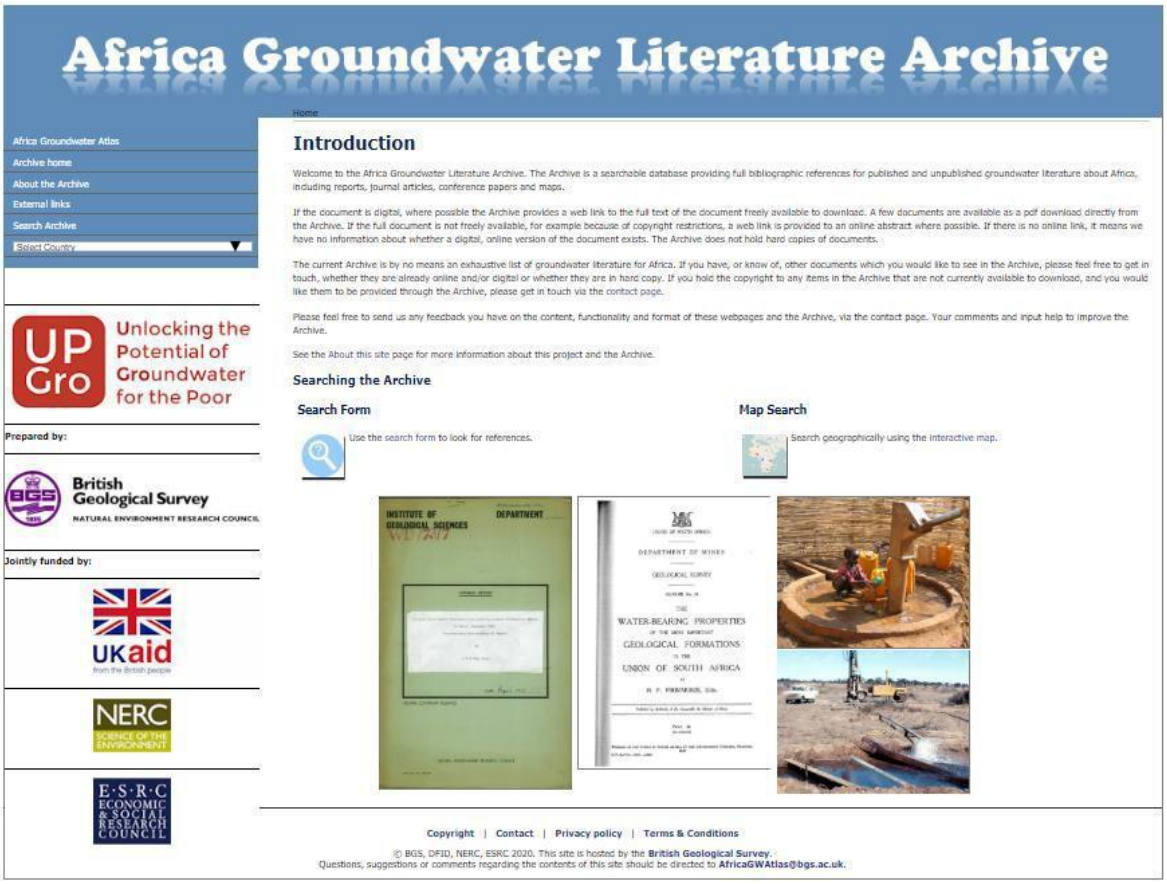
Le portail AGLA a été examiné car il possède la base de code la plus récente et la plus moderne et offre un niveau de fonctionnalité plus élevé. Le portail AGLA est actuellement mis en œuvre en HTML, CSS et JavaScript. L'accès aux données est assuré par des applications ColdFusion lisant une base de données Oracle. La recherche cartographique utilise le plugin JavaScript open source Leaflet Map. Les enregistrements individuels utilisent un plugin Google Map intégré (qui n'est plus supporté gratuitement). Cette pile technologique peut être améliorée pour fournir une architecture plus moderne, plus flexible et plus facile à maintenir.

Elle peut également être mise en œuvre en utilisant des technologies de développement web plus modernes, conformes aux normes industrielles actuelles et aux compétences des développeurs web d'aujourd'hui. Des recommandations pour moderniser la mise en œuvre technique sont fournies plus loin dans ce document.

Examen du système : Fonctionnalité et présentation

Les pages suivantes constituent l'interface utilisateur du système actuel :

Page d'accueil
https://www.bgs.ac.uk/africagroundwateratlas/archive.cfm



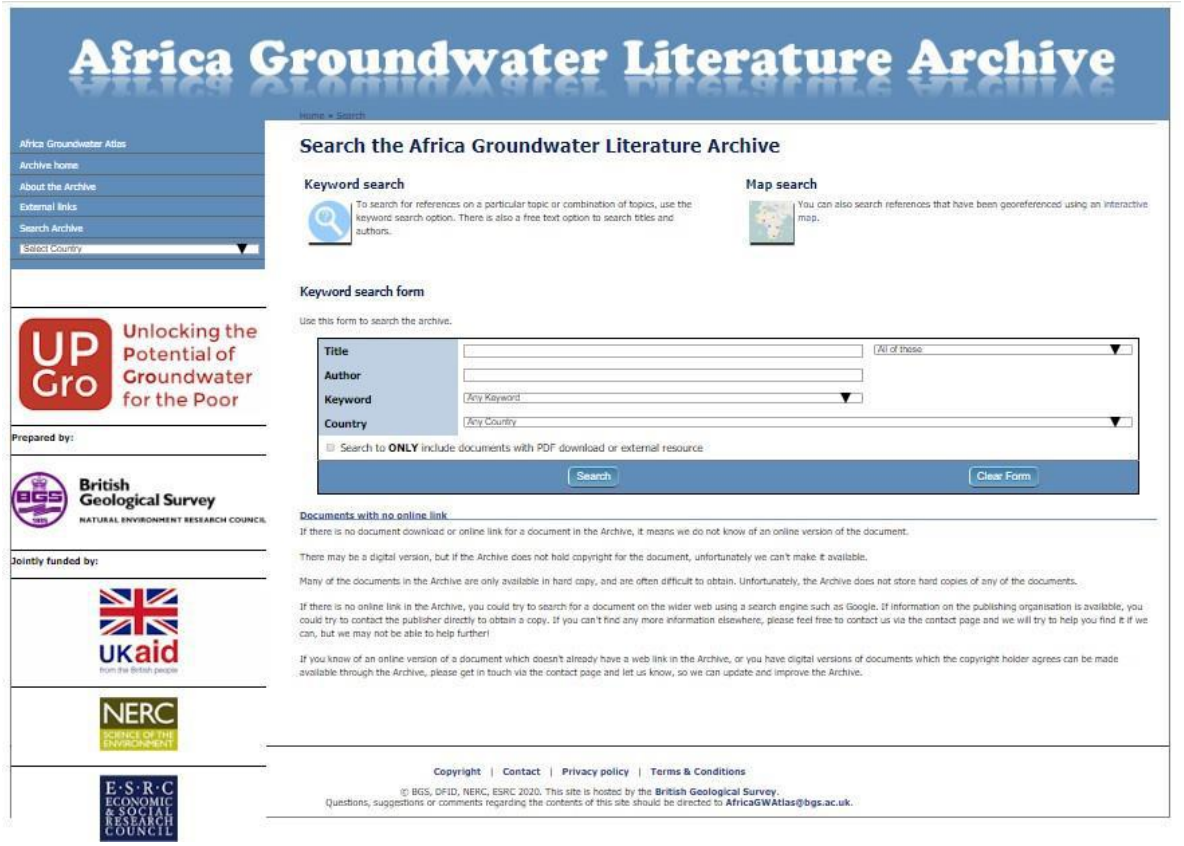
Liens vers :

- Page "Recherche de mots-clés" &gt ; &gt ; &gt ;
- Page "Recherche de carte" &gt ; &gt ; &gt ;

La liste déroulante des pays permet d'accéder aux résultats de la recherche pour le pays sélectionné &gt ; &gt ; &gt ;

Cette page est disponible en anglais et en français.

Recherche de mots-clés
https://www.bgs.ac.uk/africagroundwateratlas/atlas.cfc?method=search



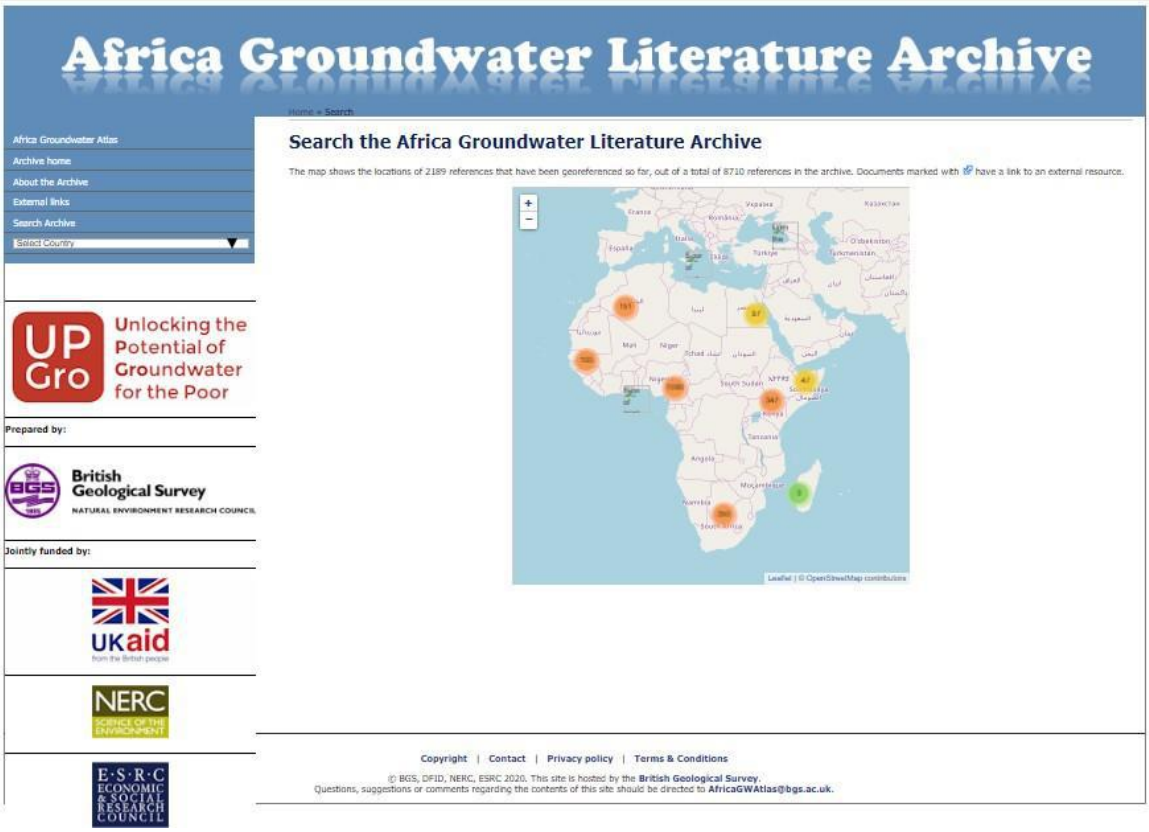
Recherchez par un ou plusieurs éléments :

- Titre (texte libre)
- Auteur (texte libre)
- Mot clé (liste)
- Pays (liste)
- Correspondance ANY/ALL/EXACT (sélectionner)
- Retourner uniquement les correspondances avec des liens vers des fichiers (case à cocher)

Soumet à "Résultats de recherche" >>&gt ;

Cette page est disponible en anglais et en français.

Recherche par carte
https://www.bgs.ac.uk/africagroundwateratlas/leaflet_map.cfm



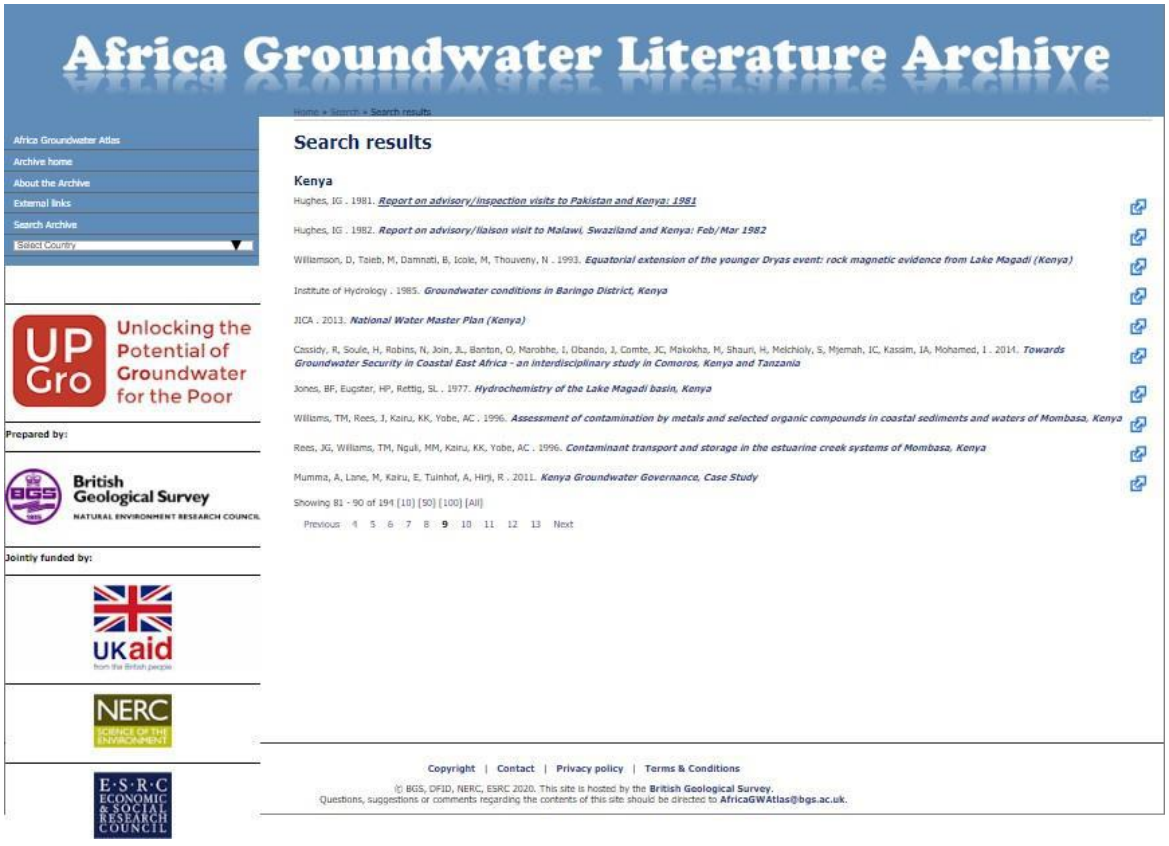
Cliquez sur un élément de

la carte Liens vers "Full

record" >>&gt ;

Cette page est disponible en anglais et en français.

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

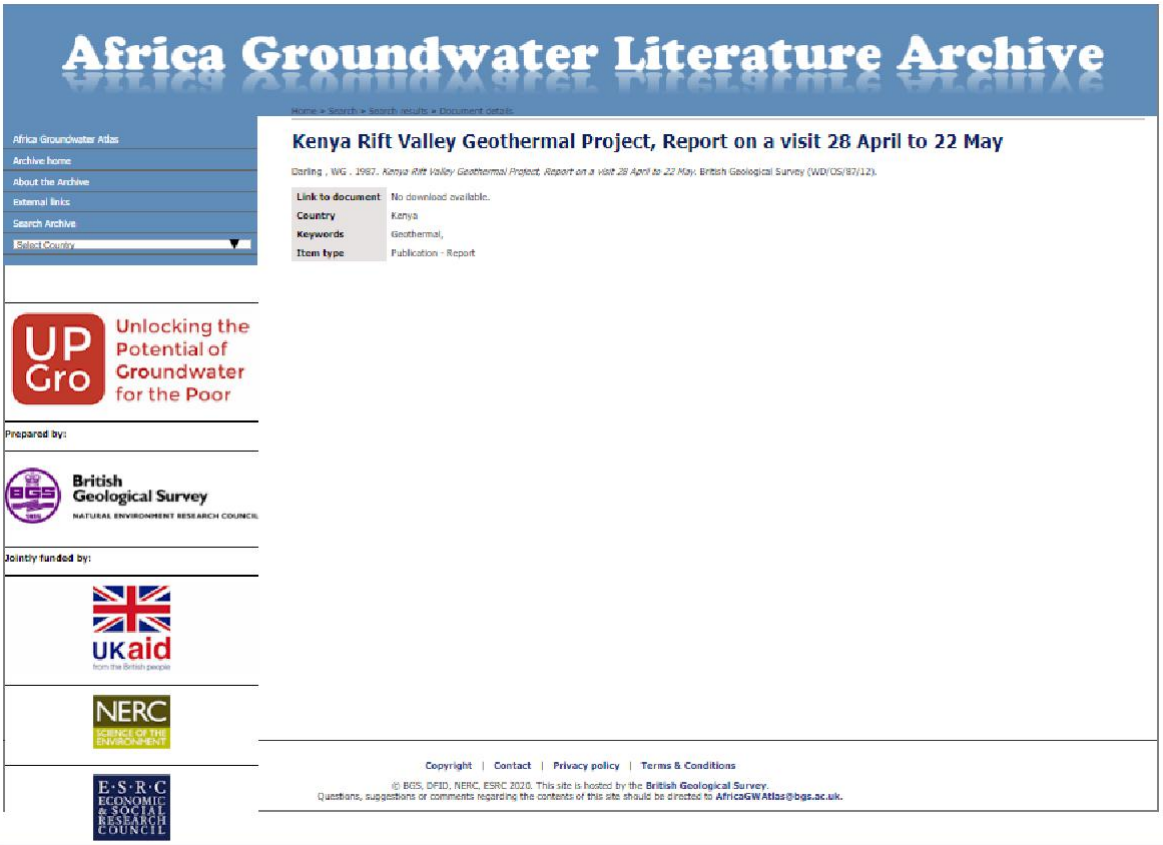


Affiche la liste des rapports correspondants regroupés par pays :

- Auteur(s)
- Année
- Titre
- Icône indiquant la disponibilité du fichier (PAS DE LIEN !)
- Lien vers "Dossier complet" >>&gt ;

Cette page est disponible en anglais et en français.

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



Affiche les détails complets d'un rapport :

- Auteur(s)
- Année
- Titre
- Résumé
- Pays
- Type de rapport
- Langue(s)
- Lien vers le "fichier" si disponible >>&gt ; URL ou DOI
- Couverture (carte Google intégrée - NE FONCTIONNE PAS DE MANIÈRE CONSTANTE POUR TOUS LES UTILISATEURS)

Cette page est uniquement disponible en anglais.

Les recommandations sur la façon dont la fonctionnalité et la présentation du système actuel peuvent être améliorées pour mieux répondre aux besoins des utilisateurs extraits lors de l'exercice de consultation des utilisateurs sont abordées dans la section Recommandations du rapport.

Révision du système : Statistiques d'utilisation du site web d'AGLA pour 2019

Les statistiques d'utilisation du site web de l'ALGA pour l'année 2019 ont été analysées afin de comprendre le niveau d'utilisation du site et d'aider à corroborer toute conclusion résultant du processus de consultation des utilisateurs.

Visiteurs quotidiens

Les visiteurs sont des personnes réelles à la recherche d'informations et comprennent les utilisateurs de navigateurs de bureau et de téléphones portables.

Nombre moyen de visiteurs par jour59

Pages vues par jour84

Nombre total de visiteurs pour 2019

En 2019, 13111 visiteurs ont atteint les différentes pages du site web d'AGLA, comme le montre la répartition ci-dessous :

Page d'accueil (EN ou par défaut)	9167	
Page d'accueil (FR)		429
Page de recherche (EN)	12071	
Page de recherche (FR)	1040	
Page de carte (EN)		839
Page de carte (FR)		67
Page de détails (EN uniquement)		959

Recherches de documents

Il s'agit de requêtes de recherche réellement soumises par les utilisateurs.

279161 recherches de documents PAR 13111 visiteurs = ~21 recherches par visiteur

Visites du site web

Les hits permettent d'accéder à tout fichier faisant partie du site, par exemple des images

individuelles, des fichiers css, etc.

Nombre total de visites	2611476		
Nombre de visiteurs	427128	17% du total	
Spider Hits	2184348	83% du total	(à partir de Google/Bing, etc.)
Nombre moyen de visites par jour	7154		
Nombre moyen de visites par visiteur	20		

Notes sur les statistiques d'utilisation du site web

- La comparaison entre les visiteurs de la page de recherche et ceux de la page d'accueil indique que de nombreuses personnes ont des signets vers la page de recherche.
- Les versions françaises des pages sont demandées en moyenne par 7 % des visiteurs.
- Près de 50 % des visites proviennent de la société de marketing en ligne "AhrefsBot", qui peut être bloquée si on le souhaite (https://botcrawl.com/ahrefsbot/).
- Près de 20 % des résultats proviennent d'une recherche commerciale distribuée au Royaume-Uni et intitulée "MJ12bot" - ces résultats peuvent être bloqués si on le souhaite (https://mj12bot.com/).

Révision du système : Analyse des termes de recherche des utilisateurs

Nous avons également analysé les termes de recherche des utilisateurs entrés sur le portail web afin de mieux comprendre les sujets qui intéressent les utilisateurs.

Ce qui suit est une analyse instantanée détaillée des mots-clés les plus populaires utilisés pour effectuer des recherches sur le portail web de la SADC, dans le but de mieux comprendre les sujets d'intérêt des utilisateurs. Cette analyse instantanée des termes de recherche couvre la période d'avril à septembre 2011. Cette analyse a pris beaucoup de temps et le temps n'était pas disponible pour la répéter pour 2019 dans le cadre de ce rapport d'analyse situationnelle. Une analyse rapide a été entreprise pour confirmer que les résultats de l'analyse de 2011 étaient toujours pertinents pour les recherches effectuées à l'heure actuelle.

Les termes de recherche énumérés dans le tableau 1 sont une combinaison des termes saisis dans tous les champs de recherche disponibles sur le formulaire de recherche (titre, auteur, mot-clé, pays). Toutefois, il convient de noter que 90 % de tous les termes de recherche sont saisis dans la zone de texte "Titre", ce qui suggère qu'un formulaire de recherche plus simple, avec un champ de recherche unique de type Google, pourrait être envisagé lors du redéveloppement du portail.

Termes de recherche par pays	Termes de recherche liés à l'eau	Termes de recherche liés à la géologie	Termes de recherche de l'organisation	Type de rapport Termes de recherche
(231) Mozambique	(337) eau	(296) géologie	(84) ressources	(42) rapport

(115) Malawi (102) Madagascar (65) Zambie (62) Afrique (53) sud (41) Namibie (40) Botswana (35) Zimbabwe (32) Tanzanie (30) Angola (22) Swaziland (20) Lesotho (11) Congo	(230) sol (203) hydro (187) Eaux souterraines (131) Hydrogéologie (42) hydrogéologique	(187) Géologique (167) carte (68) minéral (46) Exploration (42) approvisionnement (24) trou de forage (23) forage	(77) enquête (55) SADC (52) Département (21) britannique	(26) Littérature (22) revue (20) Télécharger (14) étude (8) maître

Recommandations

Les recommandations suivantes sont faites pour améliorer le portail SADC-GLA en fonction des besoins des utilisateurs et pour améliorer sa maintenabilité et sa durabilité à l'avenir.

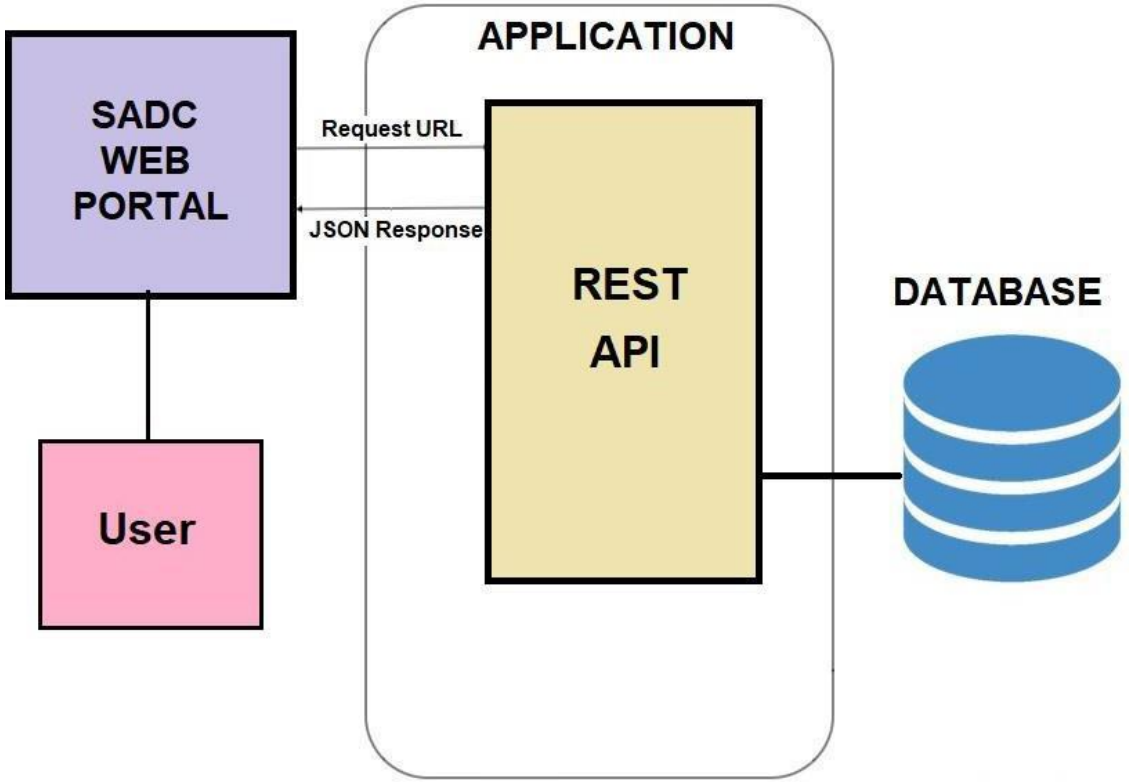
Mises à jour des données

Un contrôle de qualité général de toutes les entrées existantes de la SADC sera entrepris pour combler les lacunes de données identifiées, si possible, et supprimer les doublons. L'accent sera mis sur la mise à jour des liens vers des documents vivants et vers des sites externes afin d'améliorer la disponibilité de la littérature via les archives.
Le géoréférencement ne sera pas entrepris pour les entrées existantes dans la base de données, mais les informations sur les pays seront complétées. Le dictionnaire des mots-clés sera rationalisé.

La base de données sera normalisée afin que toute manipulation des données fasse l'objet d'un audit précis et complet. Des contraintes de dictionnaire supplémentaires seront ajoutées à la base de données pour permettre un meilleur contrôle des métadonnées.

Mises à jour du système logiciel

Afin de moderniser la mise en œuvre technique du portail SADC-GLA, il est recommandé d'adopter un système logiciel à trois niveaux comprenant un niveau de présentation, un niveau d'application et un niveau de données, comme indiqué dans le diagramme ci-dessous.



Dans ce cas, le niveau de présentation sera un portail web, tandis que le niveau de données sera le système de base de données Oracle révisé. Tous les développements utiliseront le système de gestion du code source GitLab et seront déployés par des pipelines CI/CD pilotés par des tests.

Portail web

Le portail web fournira l'interface utilisateur et sera responsable de la présentation de l'interface permettant aux utilisateurs de définir leurs besoins de recherche. Le portail web interagira avec l'application (décrite ci-dessous) afin de retourner et de présenter les informations demandées par l'utilisateur. Le portail web sera mis en œuvre en utilisant html, javascript (en utilisant des bibliothèques telles que jQuery) et css et sera réactif afin de s'assurer qu'il fonctionne sur une gamme de tailles d'écran.

Application

L'application se situe entre le portail web et la base de données. Elle fournit une interface de programmation d'applications (API) qui reçoit les demandes du portail web, interroge la base de données et renvoie ensuite les données au portail web. L'application mettra en œuvre une API standard REST (Representational State Transfer), en acceptant les demandes sous la forme de localisateurs de ressources uniformes (URL) et en renvoyant les données sous forme de texte standard JSON (JavaScript Object Notation).

Base de données

D'une manière générale, la base de données sera rationalisée en une structure robuste avec une architecture qui nous permettra de continuer à l'héberger comme aujourd'hui, mais qui se prêtera également à une exploitation sur une autre plateforme si nécessaire.

Avantages

L'avantage de cette approche est que les trois niveaux ne sont que faiblement couplés entre eux. Si la structure de la base de données change, l'application peut gérer ces changements sans affecter le portail web. La conception du portail Web peut être modifiée, ou un nouveau portail ou une nouvelle application peut être développée, sans qu'il soit nécessaire de modifier l'application. L'API peut également être étendue pour fournir des requêtes supplémentaires sans affecter la fonctionnalité existante.

Mises à jour liées à la recherche

Les recommandations suivantes sont classées dans l'ordre du parcours de l'utilisateur saisi lors de l'atelier de recherche sur les utilisateurs et ne sont pas classées par ordre de priorité :

i. Améliorer l'optimisation des moteurs de recherche
ii. Pensez à vérifier l'orthographe des critères de recherche
iii. Simplifier le formulaire de recherche. Il pourrait s'agir d'une seule zone de texte pour les critères de recherche, qui serait ensuite utilisée dans les requêtes par rapport aux champs de métadonnées clés.
iv. Rendre plus évidente l'option "exclure les résultats lorsque le document n'est pas disponible".
v. Supprimer ou simplifier éventuellement l'interface cartographique. La plupart des documents concernent un pays plutôt qu'un lieu spécifique et la plupart des utilisateurs semblent effectuer des recherches par nom de pays plutôt que par lieu géographique plus ciblé.
vi. Rafraîchir la page des résultats de la recherche et les pages de détails du document en tenant compte des champs les plus remplis selon les informations fournies par le groupe de travail sur la révision des données.
vii. Si un document n'est pas disponible en téléchargement, offrez aux utilisateurs un moyen facile de le rechercher ailleurs. Par exemple, de nombreux documents sont accompagnés de DOI ou de références académiques qui sont très susceptibles de permettre de trouver le document via des moteurs de recherche populaires.
viii. L'examen des utilisateurs a omis de prendre en compte la nécessité pour le système d'être à la fois en anglais et en français. Nous reviendrons sur les procurations des utilisateurs pour demander cette exigence.

Mises à jour relatives à l'administration

i. Utiliser Google Analytics et créer un tableau de bord accessible à la SADC-GMI, au BGS et à l'IGRAC.
ii. Créer un projet de processus pour la saisie de nouveaux documents qui implique la SADC-GMI et ses partenaires de confiance, mais s'attendre à ce que ce processus évolue au fil du temps à mesure que le modèle de gouvernance est formalisé.

iii. Définir le modèle de gouvernance pour l'ajout, l'édition et le contrôle de la qualité du contenu du nouveau site.
iv. Étudier comment surmonter les problèmes de DPI et, en fin de compte, rendre davantage de documents accessibles au public.

Modèle des jeunes professionnels

Depuis septembre 2019, la SADC-GMI a lancé un projet visant à étendre le portail d'information sur les eaux souterraines de la SADC (SADC-GIP), qui est mis en œuvre par l'IGRAC. Le SADC-GIP, lancé en 2017, est un portail web, composé d'un visualiseur de cartes et d'une base de données documentaire pour le stockage et le partage de données et d'informations relatives aux eaux souterraines dans la région de la SADC. Il est géré par la SADC-GMI avec le soutien technique de l'IGRAC. Il a été créé pour rendre les informations facilement disponibles pour les États membres de la SADC, les institutions de recherche et d'autres organisations internationales. Il contient, entre autres, des cartes compilées au cours du projet de cartographie hydrogéologique de la SADC, une base de données sur les forages, les cartes des risques liés aux eaux souterraines de la SADC, etc.

Dans le cadre du projet, de jeunes professionnels des membres de la SADC ont été engagés pour être formés sur les aspects de stockage et de partage des données sur les eaux souterraines. Ils travaillent dans des ministères ou des agences responsables de la gestion des eaux souterraines et sont impliqués dans des sujets tels que la collecte de données sur les eaux souterraines, la gestion des données, le développement ou la mise en œuvre de politiques. La plupart d'entre eux ont moins de 35 ans et ont un diplôme de l'enseignement supérieur. Les jeunes professionnels sont impliqués dans les activités suivantes du projet :

- Contribuer à l'analyse situationnelle des bases de données sur les eaux souterraines dans la SADC.
- Suivre un cours de formation en ligne sur QGIS.
- Participer à un atelier de gestion des connaissances ;
- Participer à un atelier de formation sur les systèmes d'information géographique (SIG), l'infrastructure de données spatiales (IDS) et le SADC-GIP.

Les personnes focales des Etats membres ont été invitées à nommer 2 jeunes professionnels par pays. A ce jour, 25 jeunes professionnels font partie du projet, provenant de 13 Etats membres de la SADC (sur 16).

État membre	Nombre de jeunes professionnels
Angola	2
Botswana	2
Comores	2

RDC	\
eSwatini	2

Lesotho	2
Madagascar	\
Malawi	2
Maurice	1
Mozambique	2
Namibie	2
Seychelles	\
Afrique du Sud	2
Tanzanie	2
Zambie	2
Zimbabwe	2
Total	25

Il sera proposé au jeune professionnel de prolonger son contrat afin de contribuer également au projet SADC-GLA. Il sera chargé de collecter des documents et des notices bibliographiques relatifs aux eaux souterraines dans son pays. Il peut s'agir de rapports, de documents politiques, de manuels, d'articles scientifiques, d'études, de livres, etc. provenant d'organisations nationales et internationales. L'IGRAC validera les documents collectés et les téléchargera dans le SADC-GLA. Afin de suivre l'engagement des jeunes professionnels, ceux-ci feront un rapport sur leurs recherches bibliographiques. Leurs rapports comprendront : les organisations consultées ; la liste des données collectées à partir de chaque source ; les principaux défis pour collecter les documents ; les futures sources potentielles d'information et les suggestions pour le SADC-GLA.

Les jeunes professionnels recevront une indemnité de 200 USD à la fin de toutes les tâches, qui devront être accomplies avant le 30/04/2020. Le jeune professionnel sera supervisé par l'équipe du projet et les personnes focales.

Horaire de travail

	Description de la tâche	Jan	Février	Mar	Avril	Mai	Jun
Paquet de travail 1	**WP1.1 Examen des données**						
	1. Examiner les données existantes 2. Lancer le modèle des jeunes professionnels						
	WP1.2 Revue du système						
	3. Identifier les exigences et les fonctionnalités clés 4. Développer un modèle de gouvernance 5. Revoir le système actuel (base de données et portail)						
	Rapport WP1.3						
	6. Rapport sur le WP1, y compris le plan pour le WP2						
Paquet de travail 2	**WP2.1 Mise à jour des données**						
	7. Mise à jour des données GLA existantes 8. Préparer les nouvelles données de l'AGLA et de la SADC-GIP 9. Compiler les nouvelles données des États membres						
	WP2.2 Mise à jour du système						
	10. Mettre à jour et alimenter le système de base de données 11. Développer des protocoles pour l'ajout, la modification et la mise à jour des données. 12. Développer une API sur le web 13. Développer un nouveau portail web 14. Accueillir le système GLA actualisé de la SADC						
	Rapport WP2.3						
	15. Rapport sur le WP2, y compris la stratégie à long terme						
Travail Paquet 3	16. Manuel de l'utilisateur 17. Entreprendre une formation approfondie des principaux utilisateurs du système						
Autre	Gestion continue du projet						

Annexe 3 : Mise à jour du projet

Mise à jour de l'archive de littérature grise de la SADC Mise à jour du projet

31 mars 2020

Actions et mises à jour de la réunion précédente (17 mars)

ACTION 1 : Nom de domaine : l'équipe du BGS envoie les options de domaine et les implications financières au SADC-GMI ; le SADC-GMI finalise le nom des archives et l'URL.

UPDATE 1 : De IONOS, que BGS a utilisé précédemment pour enregistrer des noms de domaine :

- .co.uk - 10 £ par an
- .org.uk - 10 £ par an
- .com - £13 par an
- .org - £16 par an
- .info - £16 par an
- Les noms de domaine existants peuvent être vérifiés sur le site web de IONOS.

ACTION 2 : Rapport d'avancement : Le BGS/IGRAC enverra un bref rapport d'avancement avant le 31 mars.

MISE À JOUR 2 : Terminé

ACTION 3 : Document de gouvernance : Le BGS/IGRAC enverra un projet de document pour commentaires avant le 15 avril, pour discussion lors de la prochaine réunion du 17 avril.

UPDATE 3 : En cours

ACTION 4 : Jeunes Professionnels : La SADC-GMI/IGRAC finalisera les contrats et les instructions pour les jeunes professionnels, y compris l'approche des pays actuellement absents du projet GIP.

MISE À JOUR 4 : Les PJ ont maintenant les contrats et les instructions pour commencer le travail

Atelier de projet interne

L'équipe du projet a organisé un atelier de deux jours les 25 et 26 mars. L'atelier devait se tenir au siège du BGS à Keyworth, mais s'est déroulé en ligne via ZOOM. L'objectif de l'atelier était le suivant

- Sur la base des résultats du rapport d'analyse de la situation, prendre les décisions finales sur la conception du nouveau système SADC-GLA, y compris la base de données, l'API et le portail web.
- Définir les tâches spécifiques requises, le calendrier et les responsabilités au sein de l'équipe.
- Convenir des aspects clés du modèle de gouvernance

Le premier jour n'a été suivi que par le personnel du BGS, tandis que le deuxième jour a également impliqué le personnel d'IGRAC.

Base de données

La structure de la nouvelle base de données a été finalisée (annexe 1) et est maintenant en cours de mise en œuvre. Le travail a commencé sur la mise à jour des données existantes dans la base de données

La base de données AGLA afin qu'elle puisse être migrée vers la nouvelle structure de la base de données. Il s'agira de réviser et de mettre à jour le dictionnaire de mots-clés en une structure rationalisée "parent" et "enfant", puis de mettre en correspondance tous les mots-clés existants dans la base de données AGLA avec le nouveau dictionnaire de mots-clés SADC-GLA. La mise à jour de la base de données devrait être terminée le 17 avril (figure 1).

Ensuite, l'équipe de la base de données se concentrera sur l'AQ des données existantes dans la SADC-GLA - principalement en complétant les principales données manquantes - et l'ajout de nouvelles données - d'abord de la SADC-GIP, puis des PJ et de l'IGRAC.

L'équipe de la base de données est également en train de finaliser un nouveau modèle de métadonnées qui sera utilisé et testé par l'IGRAC et les PJ pour rassembler de nouvelles données pour le SADC-GLA.

Portail web et API

La structure du nouveau portail web et de l'API a été finalisée (annexe 2) et les travaux préparatoires sont en cours afin que la phase principale de développement puisse commencer dès que le développement de la base de données sera terminé. La majeure partie du développement du portail web et de l'API se déroulera sous la forme d'un sprint de développement au cours des deux dernières semaines d'avril (Figure 1). Au cours de ce sprint, le personnel chargé du web et des applications travaillera intensivement sur le projet, avec la participation de l'équipe chargée de la base de données si nécessaire.

Auparavant, l'équipe web produira les mises en page finales et les maquettes pour examen par la SADC-GMI. Le texte des pages web sera également rédigé pour être revu par la SADC-GMI.

Processus de révision

Après le sprint de développement, le nouveau système sera soumis à un processus de révision interne au sein du BGS au cours des deux premières semaines de mai (Figure 1). Le nouveau système fonctionnel sera ensuite mis à disposition pour une révision externe le 15 mai, ce qui permettra de répondre aux commentaires de la SADC-GMI avant l'achèvement du projet à la fin du mois de juin.

Nouvelles données de la SADC-GIP et des jeunes professionnels

Des instructions ont été envoyées aux PJ le 25 mars avec un modèle provisoire de métadonnées, qui sera mis à jour avec le nouveau modèle au début du mois d'avril. Le BGS travaille également sur un modèle d'autorisation des DPI, qui sera complété par les PJ pour s'assurer que les autorisations requises ont été données pour les documents à partager dans la GLA.

La date limite du 22 mai a été fixée pour que les PJ renvoient les modèles et les

rapports complétés à IGRAC. IGRAC assurera un suivi avec les PJ pendant cette période pour s'assurer que la date limite est respectée. IGRAC assurera ensuite l'assurance qualité des données fournies par les PJ afin qu'elles puissent être chargées directement dans le GLA par le personnel du BGS.

Avant l'échéance du PJ, IGRAC utilisera le nouveau modèle de métadonnées pour fournir des données pour tous les documents pertinents actuellement stockés dans le SADC-GIP. Ces métadonnées seront téléchargées dans le GLA (Figure 1).

Gouvernance

La gouvernance a été discutée lors de l'atelier interne du projet et le travail a commencé sur le développement d'un document de gouvernance, qui sera distribué à SADC-GMI le 15 avril.

Autre

Un traducteur français est confirmé et l'équipe du projet est en train de rechercher un traducteur portugais.

L'impact potentiel de Covid-19 sur la réalisation du projet a été discuté lors de l'atelier du projet. L'équipe du projet ne prévoit pas de retard dans le développement du système. Il a été suggéré qu'il pourrait y avoir des retards dans la collecte de données par les PJ, mais cela devrait être minime, car la majorité de la collecte de données sera effectuée en ligne/à distance. L'atelier de formation, qui devait avoir lieu en Afrique du Sud en juin, ne sera pas possible. Ce point doit être discuté entre la SADC-GMI et le BGS pour déterminer s'il peut avoir lieu à une date ultérieure ou en ligne.

FIGURE 1 : TÂCHES DU PROJET & ; CALENDRIER

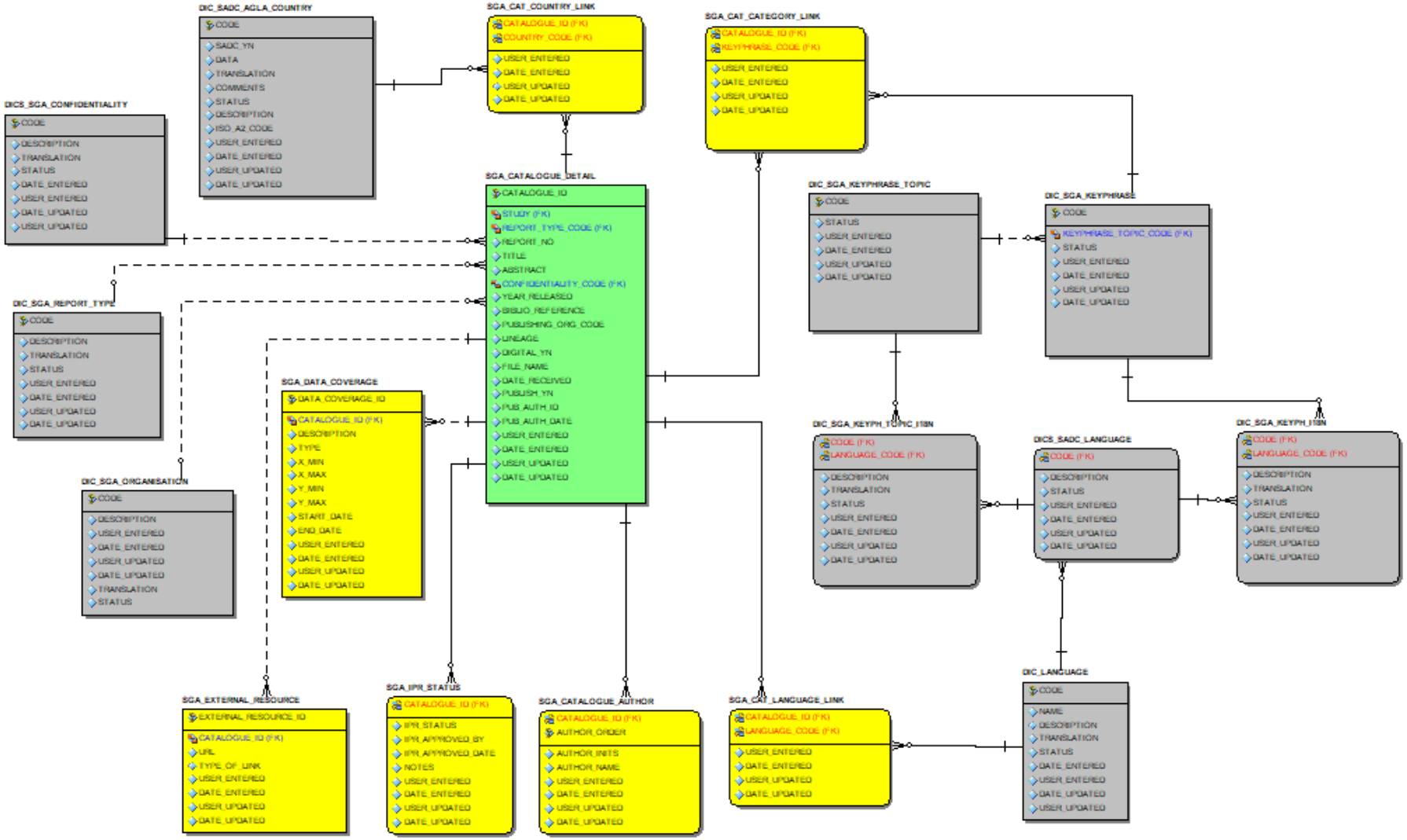
	30 mars. 3 Avril	6−10 Avril	13−17 Avril	20−25 Avril	27 avril. 1 mai	4−8 Mai	11−15 Mai	18−22 Mai	25−29 Mai	1−5 Jun	8−12 Jun	15−19 Jun	22−26 Jun
BASE DE DONNEES													
Mise à jour de la structure de la base de données													
Corriger et migrer les données existantes													
AQ des données existantes													
Ajouter les données existantes de la SADC-GIP													
Contribuer à SPRINT													
WEB & API													
Travail de préparation													
SPRINT													
Examen interne													
Examen externe													
SYSTÈME FINAL													
Finalisation sur la base de l'examen externe													

Rapport final & Manuel d'utilisation													
GOUVERNANCE													
Projet de document de gouvernance pour la SADC-GMI													
JEUNES PROFESSIONNELS													
Recevoir des modèles de métadonnées													
Modèles de métadonnées AQ													
Ajouter des données													
DELIVRABLES	**1**		**2**				**3**						**4**

DELIVRABLES

1. Rapport d'activité - 31/3
2. Projet de document de gouvernance (15/4) et réunion d'avancement (17/4)
3. Système fonctionnel - 15/5
4. Système final, rapport et manuel d'utilisation - 30/6

Annexe 1 Conception du modèle de base de données relationnelle de la SADC-GLA


DIC_SADC_AGLA_COUNTRY
CODE
SADC_YN
DATA
TRANSLATION
COMMENTS
STATUS
DESCRIPTION
ISO_A2_CODE
USER_ENTERED
DATE_ENTERED
USER_UPDATED
DATE_UPDATED
SGA_CAT_COUNTRY_LINK
CATALOGUE_ID (FK)
COUNTRY_CODE (FK)
USER_ENTERED
DATE_ENTERED
USER_UPDATED
DATE_UPDATED
SGA_CAT_CATEGORY_LINK
CATALOGUE_ID (FK)
KEYPHRASE_CODE (FK)
USER_ENTERED
DATE_ENTERED
USER_UPDATED
DATE_UPDATED
DICS_SGA_CONFIDENTIALITY
CODE
DESCRIPTION
TRANSLATION
STATUS
DATE_ENTERED
USER_ENTERED
DATE_UPDATED
USER_UPDATED
SGA_CATALOGUE_DETAIL
CATALOGUE_ID
STUDY (FK)
REPORT_TYPE_CODE (FK)
REPORT_NO
TITLE
ABSTRACT
CONFIDENTIALITY_CODE (FK)
YEAR_RELEASED
BIBLIO_REFERENCE
PUBLISHING_ORG_CODE
LINEAGE
DIGITAL_YN
FILE_NAME
DATE_RECEIVED
PUBLISH_YN
PUB_AUTH_ID
PUB_AUTH_DATE
USER_ENTERED
DATE_ENTERED
USER_UPDATED
DATE_UPDATED
DIC_SGA_KEYPHRASE_TOPIC
CODE
STATUS
USER_ENTERED
DATE_ENTERED
USER_UPDATED
DATE_UPDATED
DIC_SGA_KEYPHRASE
CODE
KEYPHRASE_TOPIC_CODE (FK)
STATUS
USER_ENTERED
DATE_ENTERED
USER_UPDATED
DATE_UPDATED
DIC_SGA_REPORT_TYPE
CODE
DESCRIPTION
TRANSLATION
STATUS
USER_ENTERED
DATE_ENTERED
USER_UPDATED
DATE_UPDATED
SGA_DATA_COVERAGE
DATA_COVERAGE_ID
CATALOGUE_ID (FK)
DESCRIPTION
TYPE
X_MIN
X_MAX
Y_MIN
Y_MAX
START_DATE
END_DATE
USER_ENTERED
DATE_ENTERED
USER_UPDATED
DATE_UPDATED
DIC_SGA_KEYPH_TOPIC_I18N
CODE (FK)
LANGUAGE_CODE (FK)
DESCRIPTION
TRANSLATION
STATUS
USER_ENTERED
DATE_ENTERED
USER_UPDATED
DATE_UPDATED
DICS_SADC_LANGUAGE
CODE (FK)
DESCRIPTION
STATUS
USER_ENTERED
DATE_ENTERED
USER_UPDATED
DATE_UPDATED
DIC_SGA_KEYPH_I18N
CODE (FK)
LANGUAGE_CODE (FK)
DESCRIPTION
TRANSLATION
STATUS
USER_ENTERED
DATE_ENTERED
USER_UPDATED
DATE_UPDATED
DIC_SGA_ORGANISATION
CODE
DESCRIPTION
USER_ENTERED
DATE_ENTERED
USER_UPDATED
DATE_UPDATED
TRANSLATION
STATUS
SGA_EXTERNAL_RESOURCE
EXTERNAL_RESOURCE_ID
CATALOGUE_ID (FK)
URL
TYPE_OF_LINK
USER_ENTERED
DATE_ENTERED
USER_UPDATED
DATE_UPDATED
SGA_IPR_STATUS
CATALOGUE_ID (FK)
IPR_STATUS
IPR_APPROVED_BY
IPR_APPROVED_DATE
NOTES
USER_ENTERED
DATE_ENTERED
USER_UPDATED
DATE_UPDATED
SGA_CATALOGUE_AUTHOR
CATALOGUE_ID (FK)
AUTHOR_ORDER
AUTHOR_INITS
AUTHOR_NAME
USER_ENTERED
DATE_ENTERED
USER_UPDATED
DATE_UPDATED
SGA_CAT_LANGUAGE_LINK
CATALOGUE_ID (FK)
LANGUAGE_CODE (FK)
USER_ENTERED
DATE_ENTERED
USER_UPDATED
DATE_UPDATED
DIC_LANGUAGE
CODE
NAME
DESCRIPTION
TRANSLATION
STATUS
DATE_ENTERED
USER_ENTERED
DATE_UPDATED
USER_UPDATED

Annexe 2 Conception du portail Web de la SADC-GLA

- **Accueil**
  - Texte d'introduction
  - Recherche par expression {TEXT} >&gt ; Résultats de la recherche
  - Liens vers les pages des pays (x16) >&gt ; Liste des pays
  - Lien vers >&gt ; Recherche avancée
  - *Coordonnées de contact*
- **Recherche avancée**
  - Titre {TEXT}
  - Auteur {TEXT}
  - Pays / Pays {LISTE DE CODES}
  - Hiérarchique (2 niveaux) Mot clé / Mots-clés {LISTE DE CODES}
  - Uniquement lié {BOOLEAN}
  - {GO BUTTON} >&gt ; Résultats de la recherche
  - *Autoriser plusieurs pays et mots-clés*
- **Résultats de la recherche**
  - Table des citations avec options de tri par colonne et liens de pagination
  - Les citations affichent des liens en direct lorsqu'elles sont disponibles
  - *Mettre en évidence les critères de recherche (c'est-à-dire le mot dans le titre, le mot-clé, etc.).*
  - *Capacité de filtrage*
  - Les rangées sont liées à >&gt ; Détails de la citation
- ACTION : TOUS doivent décider de la liste des champs à afficher dans les résultats de recherche.
- **Détails de la citation**
  - Tableau de toutes les métadonnées disponibles pour une citation
  - Les citations affichent des liens en direct lorsqu'elles sont disponibles
  - Si des données spatiales sont présentes, l'enrichissement progressif JavaScript est utilisé pour ajouter la carte.
- **Liste des pays**
  - Liste sur une seule page de toutes les citations pour un seul pays
  - Destiné à être facilement indexé par les moteurs de recherche
  - Les rangées sont liées à >&gt ; Détails de la citation
- **À propos de**
- **Modalités, confidentialité et droits d'auteur**

Documentation du système d'archivage de la littérature sur les eaux souterraines de la SADC

Juin 2020

Les auteurs :

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

Table des matières

Introduction51
Aperçu du système51
Interface utilisateur51
  Design51
  Descriptions des dossiers52
  Bibliothèques JavaScript utilisées52
  Application JavaScript53
  Gestion des données53
  Service de cartographie53
API53
  Design53
  Utilisation de l'API54
  Mise en œuvre du logiciel54
  Déploiement du serveur55
  Glossaire55
Base de données55
  Aperçu de la base de données55
  Structure de la base de données56
  Dictionnaires - Vocabulaires56
    SGA_CATALOGUE_DETAIL58
  Environnement de la base de données60
  Résumé des données de la base de données61
Personnes de contactErreur ! Signet non défini.

Introduction

L'Institut de gestion des eaux souterraines de la Communauté de développement de l'Afrique australe (SADC-GMI) est un centre d'excellence régional pour la gestion des eaux souterraines, dont la vision est d'assurer l'utilisation et la protection équitables et durables des eaux souterraines dans la région de la SADC. En 2020, le SADC-GMI a lancé un projet de mise à jour des archives de littérature grise de la SADC (SADC-GLA), afin d'en faire des archives en ligne entièrement fonctionnelles de la littérature sur les eaux souterraines pour la région de la SADC. Le projet a débuté en février 2020 et a été mis en œuvre par le British Geological Survey (BGS), en collaboration avec le Centre international d'évaluation des ressources en eaux souterraines (IGRAC).

Le nouveau SADC-GLA est mis en œuvre comme un système flexible et interopérable, comprenant une base de données, une application et une interface utilisateur en ligne. Le système sera hébergé par le British Geological Survey pendant deux ans, avant d'être transféré au SADC-GMI. Ce rapport fournit les spécifications techniques du système.

Aperçu du système

Le portail SADC-GLA est mis en œuvre sous la forme d'un système logiciel à trois niveaux comprenant des niveaux de présentation, d'application et de données, comme le montre la figure 1. Dans le système actuel, le niveau de présentation est un portail Web et le niveau de données est un système de base de données relationnelle Oracle 11g. Le niveau d'application, qui sert d'intermédiaire entre le portail Web et la base de données, est un serveur d'interface de programmation d'applications (API) REST (Representational State Transfer).

L'avantage de cette approche est que les trois niveaux ne sont que faiblement couplés entre eux. Si la structure de la base de données change, l'application peut gérer ces changements sans affecter le portail web. La conception du portail Web peut être modifiée, ou un nouveau portail ou une nouvelle application peut être développée, sans qu'il soit nécessaire de modifier l'application. L'API peut également être étendue pour fournir des requêtes supplémentaires sans affecter la fonctionnalité existante.

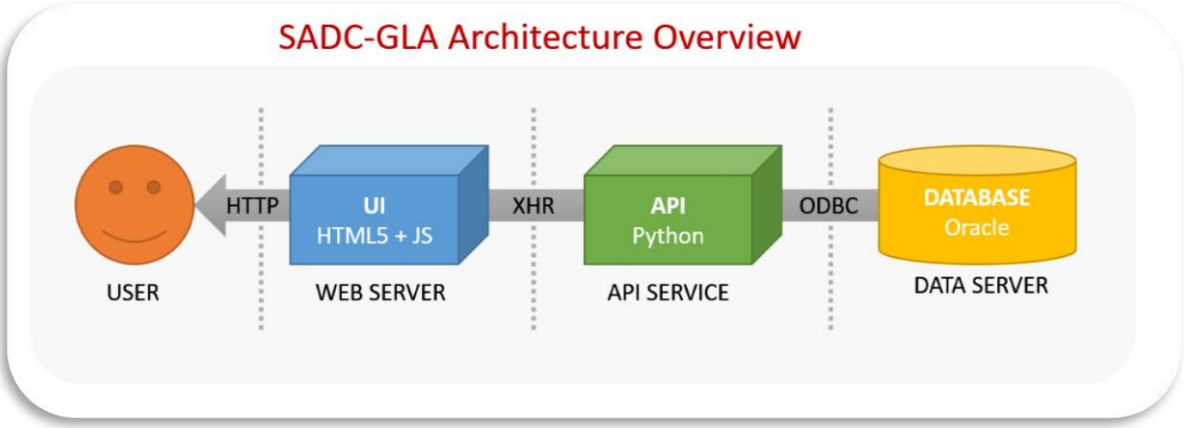


Figure 1 : Aperçu de l'architecture de la SADC-GLA

Interface utilisateur

Design

L'interface utilisateur de la SADC-GLA est implémentée en HTML et JavaScript et accède aux données en utilisant des requêtes XHR vers l'API.

- L'interface utilisateur et la base de données ne sont pas directement liées - toutes les demandes de données utilisent XHR pour appeler l'API.
- Tous les accès aux données sont en lecture seule et toutes les demandes utilisent le protocole HTTPS.
- Le code de l'interface utilisateur est conçu pour être autonome et ne renvoie à aucune ressource externe - aucun service tiers n'est utilisé.
- L'interface utilisateur n'utilise pas de cookies - aucune fenêtre de "consentement" n'est donc nécessaire.
- Dans la mesure du possible, les normes du W3C seront respectées.
- Les bibliothèques JavaScript utilisées (jQuery et Leaflet) seront stockées localement.
- Toutes les ressources (CSS, images, etc.) seront stockées localement.
- Le serveur web hôte est utilisé pour fournir des inclusions côté serveur pour l'en-tête, le pied de page et la navigation. Ce serveur Web est actuellement IIS, mais Apache et NGINX pourraient fournir la même fonctionnalité.
- Aucun autre code (Java, PHP, etc.) ne devra être installé sur le serveur web pour déployer l'interface utilisateur.

Cela devrait permettre de déployer l'interface utilisateur sur n'importe quel serveur Web, à condition qu'il dispose d'un accès réseau HTTP à l'API, ce qui peut être une machine locale ou un serveur distant n'importe où dans le monde.

Descriptions des dossiers

Dossier	Type(s) de fichier	Contenu
/resources/images/...	PNG, SVG, JPG & GIF	Logos, drapeaux de pays, icônes d'interface utilisateur
/ressources/CSS/...	CSS	Principaux styles de mise en page
/resources/cache/...	JSON	Frontières de pays mises en cache sous forme de GeoJSON
/ressources/JS/...	JS	Bibliothèques JavaScript et code JavaScript de base
/SADC/...	HTML5	Modèles de pages Web
/SADC/includes/...	HTM	En-tête, navigation et pied de page HTML5 inclus
/SADC/CSS/...	CSS	Police et couleurs du site web
/SADC/docs/...	XLSX, PDF	Téléchargements, par exemple un fichier Excel de saisie de données
/SADC/js/...	JS	Code d'initialisation de la page Web

Bibliothèques JavaScript utilisées

- **jQuery 3.4.1 - traitement des événements et des demandes XHR https://jquery.com/**

- **Leaflet 1.5.1 - montre la carte sur la page de détail https://leafletjs.com/**
- **Leaflet Lat/Lng Graticule Plugin - ajoute une grille en superposition à la carte https://github.com/cloudybay/leaflet.latlng-graticule**

Application JavaScript

Fichier	Description
agw.core.js	Enveloppe pour les composants modulaires
agw.api.js	Gère tous les appels API
agw.ln.js	Gestion des fonctionnalités multilingues
agw.detail.js	Gère l'affichage d'une fiche détaillée
agw.detail.map.js	Traite la carte affichée pour un enregistrement détaillé
agw.searchForm.js	Gère l'interaction avec le formulaire de recherche
agw.phraseSearch.js	Gère la fonctionnalité de recherche de phrases
agw.searchResults.js	Gère l'affichage des résultats de la recherche
agw.memberCountries.js	Gère l'affichage des listes de pays

Gestion des données

Un téléchargement de la feuille Excel (telle qu'envoyée aux jeunes professionnels) sera disponible pour ceux qui souhaitent saisir plusieurs enregistrements. Un formulaire de saisie (avec les mêmes champs que le fichier Excel) est fourni - les données saisies sont sérialisées dans un format facile à "couper-coller" qui peut être copié directement dans la feuille Excel. Les nouvelles entrées saisies de cette manière seront envoyées par courriel à SADC-GMI pour l'AQ. Une fenêtre contextuelle " signaler une erreur " est disponible sur la page de détail de chaque enregistrement. Elle permet aux utilisateurs de signaler des erreurs dans les données existantes. Les rapports d'erreur seront envoyés par courriel à SADC-GMI pour l'AQ.

Service de cartographie

La carte Leaflet utilisée sur la page de détails du document utilise un service de carte WMS ouvert fourni par le BRGM pour le compte de OneGeology (https://mapsref.brgm.fr/wxs/refcom- brgm/refign). Les données topographiques indiquées proviennent d'Open Street Map.

API

Design

Le serveur API met en œuvre une API REST standard :

- Recevoir des demandes HTTP sous la forme de localisateurs de ressources

uniformes (URL) de la part d'un client,

- En utilisant les paramètres de ces URL pour former des requêtes de base de données,
- L’interrogation de la base de données,
- Renvoyer les données au client sous forme de texte standard en notation d'objets JavaScript (JSON).

Utilisation de l'API

L'API répond à un certain nombre de requêtes des clients. Une requête est formée en ajoutant du texte structuré à l'URL de base de l'API.

Les requêtes suivantes sont disponibles :

- obtenir une liste des pays de la SADC
- obtenir une liste de mots-clés
- obtenir les détails du résumé d'un document en utilisant son ID unique
- obtenir les détails complets d'un document à l'aide de son identifiant unique
- obtenir tous les détails du résumé pour un seul pays de la SADC
    - rechercher dans la base de données les résumés qui correspondent à l'un ou à l'ensemble d'une liste de termes de recherche :
    - pays, mot-clé, auteur, titre, type de rapport
    - recherche générale par expression
    - réduire par une gamme d'années de publication + choisir seulement les documents avec des liens disponibles
    - obtenir les résultats par lots gérables

Pour chacune de ces requêtes, l'API renvoie une réponse JSON structurée contenant les données et toute métadonnée pertinente ou toute information d'erreur, si la requête a échoué. La documentation complète de toutes les requêtes et de leurs réponses, y compris des exemples, se trouve sur la page de documentation de l'API du serveur. Une documentation supplémentaire est disponible sous la forme d'un document OpenAPI JSON, bien que celui-ci soit plus utile aux systèmes automatisés pour faciliter la création de requêtes et l'analyse des réponses.

Mise en œuvre du logiciel

L'API est mise en œuvre à l'aide du langage Python 3 en utilisant un certain nombre de paquets standard. Le paquet principal est FastAPI qui est utilisé pour définir des routes (requêtes) et retourner les modèles définis (réponses) en JSON. FastAPI utilise uvicorn pour exécuter un service API. FastAPI valide aussi automatiquement les requêtes entrantes, retournant des erreurs si les requêtes sont incorrectes d'une certaine manière. Le paquetage etlhelper est utilisé pour gérer les connexions aux bases de données et les requêtes SQL, renvoyant des lignes de la base de données en réponse aux requêtes SQL.

Le logiciel est structuré en un certain nombre de couches :

- les fonctions d'itinéraire et les fonctions de données de soutien
- définitions des modèles
- les gestionnaires de bases de données
- Requêtes SQL

Pour une requête typique, l'API transmettra la requête à la route appropriée. La route analysera et validera automatiquement l'URL à l'aide de FastAPI. Si la requête est correcte, la fonction de route structure les paramètres de la requête et demande une réponse à sa fonction de données. La fonction de données transmet la requête structurée à un gestionnaire de base de données. Le gestionnaire utilise les paramètres de la requête pour construire la requête SQL appropriée et l'envoyer à la base de données en utilisant etlhelper.

Une fois que la base de données a renvoyé un certain nombre de lignes en réponse à une requête, le gestionnaire effectue quelques modifications structurelles et renvoie la liste des lignes à la fonction de données. Le site

La fonction data utilise les lignes et construit des objets de modèle à partir des données de ces lignes. Ces éléments de données sont ajoutés à un objet modèle de réponse et renvoyés à la route. La route renvoie alors cette réponse avec FastAPI transformant automatiquement le modèle en texte JSON avant de le renvoyer au client.

Déploiement du serveur

Le serveur peut être déployé de plusieurs façons. Une fois que tous les paquets Python sont installés, il peut simplement être exécuté en ligne de commande à l'aide de uvicorn. Cependant, il est plus pratique de déployer le serveur en utilisant un conteneur Docker. Un tel conteneur est défini dans l'implémentation et peut être exécuté de plusieurs manières.

Il existe une certaine configuration concernant principalement la base de données cible. Bien que le déploiement actuel utilise Oracle, il serait facile de passer à un autre système de base de données, PostgreSQL par exemple, en modifiant la configuration et une partie du SQL. Le reste du code n'aurait pas besoin d'être modifié.

Glossaire

API - Application Programming Interface

(interface de programmation

d'applications) HTTP - HyperText

Transfer Protocol (protocole de transfert

hypertexte) JSON - JavaScript Object

Notation (notation d'objet JavaScript)

REST - Representational State Transfer

(transfert d'état représentationnel) SQL -

Structured Query Language (langage de

requête structuré)

URL - Localisateur de ressources uniformes

Base de données

Aperçu de la base de données

La base de données est mise en œuvre dans un système de gestion de base de données relationnelle Oracle 11gR2 et catalogue la littérature sur les eaux souterraines concernant la Communauté de développement de l'Afrique australe (SADC). Elle contient des métadonnées générales sur la littérature, y compris les références de milliers de rapports, articles, livres, documents de conférence, documents politiques et cartes. La base de données est l'un des composants d'un système d'architecture à trois niveaux qui comprend une couche API et une interface GUI qui offre une bibliothèque cataloguée et consultable, y compris des téléchargements, soit directement dans le système, soit par un lien Web

externe, de la littérature sur les eaux souterraines concernant la SADC. La base de données offre des métadonnées supplémentaires sur la littérature qui permettent au système SADC-GLA d'être facilement consultable par auteur, titre, pays et mots-clés.

Le contenu de la base de données provient principalement des archives Grey Literature Archive (GLA) existantes, gérées par le BGS. Les métadonnées des documents sont conservées dans leur langue d'origine dans le catalogue, mais un ensemble internationalisé de dictionnaires de mots-clés (vocabulaires contrôlés) en anglais, français et portugais est également mis en œuvre. Ces dictionnaires sont utilisés pour marquer les articles du catalogue afin de faciliter la recherche et l'extraction à partir du portail. Les nouvelles données acquises seront soumises à des procédures d'AQ et de CQ et seront incluses dans la base de données le cas échéant.

Structure de la base de données

La base de données est mise en œuvre dans un système de base de données relationnelle Oracle 11gR2 et est compatible avec d'autres systèmes de base de données relationnelles. Elle est compatible avec d'autres systèmes de bases de données relationnelles. La base de données est de conception normalisée et contient une table d'index principale (SGA_CATALOGUE_DETAIL) qui contient les métadonnées d'une instance unique d'un article de catalogue, puis des liens vers d'autres tables qui capturent les métadonnées où il peut y avoir plusieurs instances de divers attributs relatifs à l'article de catalogue. Ces autres attributs d'un catalogue qui peuvent avoir plus d'une instance incluent : Les auteurs, les pays, les mots-clés-keyphrases, la langue du document de catalogue, les liens vers des ressources externes, le statut IPR et la couverture des données.

La base de données utilise également des techniques de dénormalisation pour regrouper les sorties de données dans des vues/objets de requête qui servent de couche intermédiaire de base de données pour faciliter l'accès par l'API et optimiser les performances des requêtes. Cette abstraction permet également de mettre à jour la structure principale de la base de données sans affecter le système réel jusqu'à ce qu'il soit approprié de déployer les changements sur le système réel.

La base de données se compose de 19 tables (entités) dont 11 tables sont des vocabulaires contrôlés avec des sous-ensembles. Il existe également 11 objets d'interrogation qui font partie de la couche intermédiaire d'interrogation de la base de données. La base de données utilise également les types d'objets suivants dans le cadre de sa conception : séquences, déclencheurs et vues. Les objets de la base de données sont nommés selon une norme de dénomination cohérente dans l'ensemble du modèle de données.

La base de données intègre également une piste d'audit complète dans le cadre de sa mise en œuvre avec des tables d'historique et des déclencheurs. Toutes les modifications apportées aux données sont enregistrées, y compris le nom de l'utilisateur et la date à laquelle l'action a été effectuée. Pour toutes les nouvelles entrées de données, l'utilisateur et la date de cette action sont également enregistrés.

Dictionnaires - Vocabulaires

La base de données utilise un certain nombre de vocabulaires contrôlés (termes et leurs définitions) afin de maintenir un niveau de terminologie cohérent dans toute la base de données pour le marquage des données. Ces termes sont utilisés pour organiser les informations afin de faciliter la recherche et/ou la saisie des données. Il s'agit de termes prédéfinis et autorisés, de listes de mots ou de phrases soigneusement sélectionnées, qui sont utilisés pour étiqueter les données afin qu'elles puissent être facilement récupérées par diverses techniques de recherche.

Ils réduisent l'ambiguïté inhérente au balisage des données en texte libre, le même terme étant appliqué et réutilisé de manière cohérente dans toute la base de données. Ils sont aussi généralement utilisés pour alimenter des listes déroulantes dans des applications ou des listes de contrôle que l'utilisateur peut sélectionner pour la saisie et l'extraction de données.

La base de données SADC-GLA a mis en place 11 dictionnaires pour couvrir les domaines suivants : Mots-clés - phrases clés, pays, langue, confidentialité, statut des DPI, types de documents et types de liens entre documents.

Les dictionnaires d'expressions clés ont été internationalisés pour trois langues convenues,

à savoir l'anglais, le français et le portugais, afin de permettre le marquage approprié des articles du catalogue et leur présentation sur le portail web pour la recherche et la récupération dans chacune de ces langues.

Tableau 1 : Entités/Tables et leurs définitions

TABLEAU	DÉFINITION
SGA_CATALOGUE_DETAIL	Table qui contient les principales métadonnées d'un document du catalogue qui est pertinent ou associé à un ou plusieurs pays de la SADC.
SGA_CAT_COUNTRY_LINK	Table qui contient la liste des pays qui se rapportent à un document. C'est l'association entre un document et un pays.
SGA_CAT_CATEGORY_LINK	Table qui contient la liste des phrases clés pertinentes associées au document.
SGA_CATALOGUE_AUTHOR	Table qui contient la liste des auteurs (par ordre de publication) pour chaque document.
SGA_CAT_LANGUAGE_LINK	Table qui contient la ou les langues originales d'un document.
SGA_RESSOURCE_EXTERNE	Tableau qui contient un lien vers l'emplacement d'un document ou vers une page où ce lien peut être accessible.
SGA_IPR_STATUS	Table qui contient une référence au DPI d'un document.
SGA_DATA_COVERAGE	Table qui contient la ou les étendues géographiques qui définissent la zone d'influence d'un document.
Dictionnaires - Vocabulaires contrôlés	
DIC_SADC_AGLA_COUNTRY	Table de dictionnaire des noms de pays avec leurs codes ISO
DIC_SGA_KEYPHRASE	Tableau de dictionnaire des phrases clés des eaux souterraines. Utilisé pour étiqueter les éléments du catalogue (documents) pour la recherche et la récupération.
DIC_SGA_KEYPHRASE_TOPIC	Tableau de dictionnaire des sujets de phrases clés. Il s'agit de la catégorisation des phrases clés en termes de plus haut niveau pour le filtrage et le sous-classement dans les recherches.
DICS_SADC_LANGUAGE	Table de dictionnaire sous-ensemble de langues.

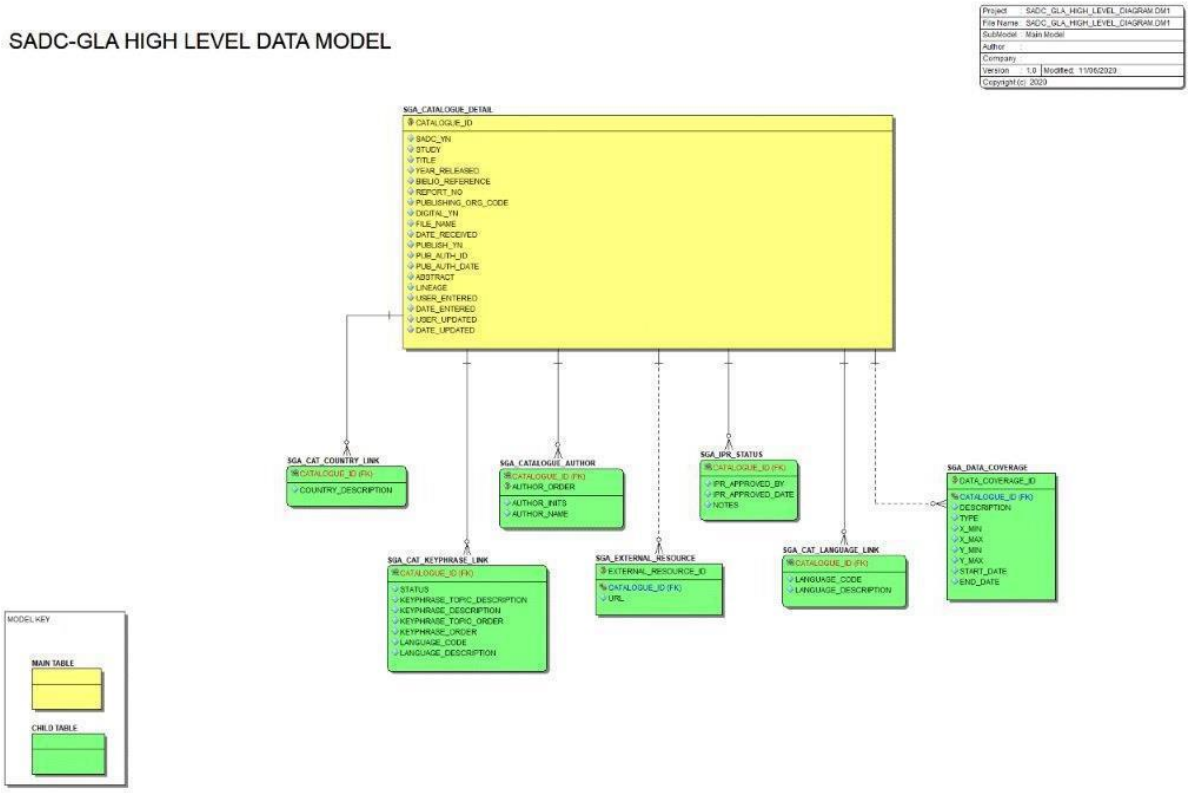


Figure 2 : Modèle de données de haut niveau (montrant uniquement les entités de haut niveau).

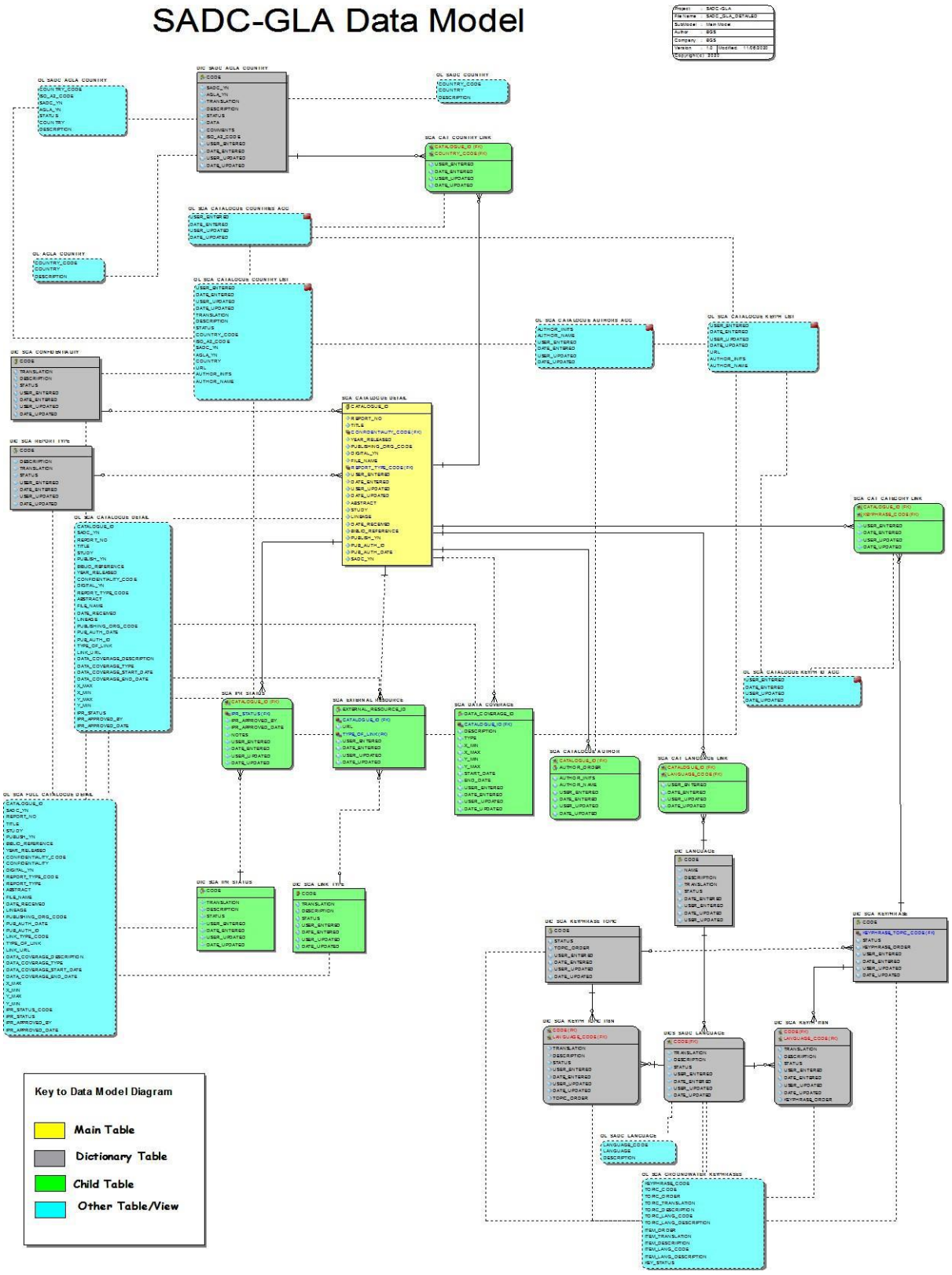


Figure 3 : Modèle de données détaillé de la base de données

Environnement de la base de données

La base de données est basée sur le modèle relationnel et mise en œuvre dans un système de gestion de base de données relationnelle Enterprise Oracle 11gR2 hébergé au BGS. L'implémentation de la base de données utilise la norme ISO/ANSI SQL et est compatible avec la plupart des systèmes de gestion de bases de données relationnelles. Des unités de programme ont également été définies dans Oracle PL/SQL afin d'appliquer l'audit dans le cadre de la mise en œuvre d'une piste d'audit complète.

La base de données comprend l'utilisation des types d'objets de base de données suivants : Tables, vues, séquences et déclencheurs.

Le jeu de caractères de la base de données Oracle est AL32UTF8 et fonctionne sur une plate-forme Linux x86 64 bits. Également compatible sur les plateformes Windows.

Les détails du SQL de création d'objet et du code PL/SQL sont tous inclus dans l'annexe.

Résumé des données dans la base de données

Les données contenues dans le nouveau catalogue SADC-GLA proviennent en grande partie des archives GLA existantes, qui ont été enrichies par le BGS au cours de nombreuses années pour inclure un grand nombre de documents pertinents. Toute nouvelle donnée acquise dans le cadre des travaux de la SADC-GLA sera également soumise à des procédures d'AQ et de CQ avant d'être incluse dans la base de données afin d'accroître cette ressource pour la SADC.

Les données consistent en une liste de documents relatifs aux pays de la SADC et sont stockées dans une série de tables de base de données normalisées ainsi que d'autres tables. Ces tables contiennent des métadonnées relatives à ces documents et sont nécessaires pour les définir dans un système de catalogue qui est à la fois consultable et actualisable.

Un résumé des principales données est présenté dans les tableaux ci-dessous à titre d'illustration. Il y a un peu plus de 4000 documents dans la base de données (Tableau 2), principalement pour les pays de la SADC (Tableau 3), en 7 langues (Tableau 4) bien que la majorité soit en anglais.

Tableau 2 : Résumé des types de documents

Types de documents	Nombre de documents
Rapports	2124
Articles	1031
Sections du livre	238
Cartes	182
Documents de la conférence	176
A catégoriser	110
Livres	45

Thèses	39
Autres documents	24

Dissertations	19
Sortie de données numériques	10
Documents de politique générale	6
Total	4004

Tableau 3 : Documents par pays

Pays	Nombre de documents
Botswana	731
Afrique du Sud	707
Zimbabwe	637
Malawi	431
Tanzanie, République-Unie de	352
Zambie	283
Madagascar	257
Namibie	235
Lesotho	157
Mozambique	122
Congo, La République Démocratique du Congo	99
eSwatini	98
Angola	88
Kenya	36
Maurice	32
Ouganda	31
Comores	30
Seychelles	29
Éthiopie	28

Ghana	22
Soudan	15
Érythrée	13
Rwanda	12
Nigeria	12
Égypte	11
Burundi	11
Sénégal	10
Djibouti	9
Niger	7
Mali	7
Sud-Soudan	6
Arabie Saoudite	4
Sierra Leone	4
Algérie	4
Cameroun	4
Burkina Faso	4
Maroc	3
Bénin	3
Congo	3
Chad	3
Gabon	3
Tunisie	3
Côte d'Ivoire	3
Sri Lanka	3
Libye	2

Somalie	2
Gambie	1
Papouasie-Nouvelle-Guinée	1
Réunion	1
République centrafricaine	1
Antilles néerlandaises	1
Guinée	1
Mauritanie	1
Togo	1
Liberia	1

Tableau 4 : Documents par langue

NOM	Nombre de documents
Anglais	3824
Français	136
Allemand	21
Portugais	20
Afrikaans	2
Chichewa;Chewa ; Nyanja	2
Espagnol ; Castillan	1

Les données sont passées par plusieurs étapes d'AQ et de CQ pour supprimer les doublons, corriger les erreurs, développer les métadonnées et ajouter de la valeur aux données. Le nouveau système SADC-GLA a été mis à jour pour refléter ce travail considérable.

Parmi les processus et procédures mis en œuvre, on peut citer :

1. Contrôles textuels

    - Langue spécifique si possible
    - Structure des titres
    - Orthographe/Rationalisation des noms et initiales des auteurs
2. Détection et rationalisation des doublons (en cours)
    - Duplication en plusieurs langues
    - Duplication structurelle (espaces supplémentaires, utilisation d'acronymes, ponctuation, etc.)
3. Attribution et vérification des métadonnées
    - Pays
    - Mots/phrases clés
    - Correction de la langue originale
4. Liens vers les documents
    - Validité
    - Type

Annexe 5 : Rapport sur la collecte de données

Mise à jour des archives de documents sur les eaux souterraines de l'ADC

Rapport sur la collecte de données, 28 septembre 2020, IGRAC

Introduction

L'Institut de gestion des eaux souterraines de la Communauté de développement de l'Afrique australe (SADC-GMI) est un centre régional d'excellence pour la gestion des eaux souterraines, dont la vision est d'assurer l'utilisation équitable et durable et la protection des eaux souterraines dans la région de la SADC. Le SADC-GMI a lancé un projet de mise à jour des archives de la littérature sur les eaux souterraines de la SADC (SADC-GLA), afin d'en faire des archives en ligne entièrement fonctionnelles de la littérature sur les eaux souterraines pour la région de la SADC. Le projet a débuté en février 2020 et devait se terminer en juin 2020, mais la fin du projet a été reportée à octobre 2020 en raison des pandémies Covid19. Le projet a été mis en œuvre par le British Geological Survey (BGS), en collaboration avec le Centre international d'évaluation des ressources en eaux souterraines (IGRAC). Le rôle de l'IGRAC était de faire migrer les notices bibliographiques du SADC-GIP et d'inciter les jeunes professionnels des États membres de la SADC à collecter des notices supplémentaires dans leurs pays. Ce rapport fournit une vue d'ensemble de ces activités de projet. Il souligne les réalisations et les défis rencontrés et fournit des recommandations pour le développement de la SADC-GLA dans les années à venir.

Migration des données du système de gestion de l'information de la SADC-GIP et de Ramotswa

Les notices bibliographiques ont été migrées du portail d'information sur les eaux souterraines de la SADC (SADC-GIP) et du système de gestion de l'information Ramotswa (RIMS). Ces deux systèmes étaient hébergés dans le système mondial d'information sur les eaux souterraines (GGIS), géré par IGRAC. En 2020, il a été décidé de migrer les couches cartographiques contenues dans ces visionneuses vers une nouvelle version du SADC-GIP (https://sadc-gip.org/) et les notices bibliographiques vers le SADC- GLA. En mai-juin 2020, 44 notices du SADC-GIP et 25 notices du RIMS ont été migrées vers le SADC-GLA. IGRAC a rempli les modèles de téléchargement préparés par BGS avec les 69 enregistrements. Ce travail a été effectué par deux membres du personnel, afin de tester l'exhaustivité, la robustesse et la convivialité du modèle, ainsi que la cohérence des entrées. Quelques défauts mineurs ont été identifiés dans le modèle. Certaines incohérences ont également été identifiées entre les contributions des deux équipes, notamment en ce qui concerne l'utilisation de mots-clés, de rédacteurs et d'éditeurs, ainsi que l'utilisation de noms d'organisations comme auteurs. Le modèle a été mis à jour en conséquence, dans la mesure du possible, mais plusieurs incohérences résultent de choix subjectifs des deux équipes.

Programme pour les jeunes professionnels

Dans le cadre d'un autre projet SADC-GMI (Expansion du portail d'information sur les eaux souterraines de la SADC, SADC-GIP), de jeunes professionnels des départements nationaux de l'eau des États membres ont été engagés pour participer à plusieurs missions

et activités de renforcement des capacités. 25 jeunes professionnels ont été engagés, provenant de 13 Etats membres (Tableau 0-1). Les mêmes jeunes professionnels ont été engagés dans le projet actuel. La RDC, Madagascar et les Seychelles ont été invités à nouveau à nommer quelqu'un (ils ne l'avaient pas fait auparavant), mais ils n'ont pas réagi. Les jeunes professionnels ont été chargés de collecter des documents dans leur pays pour les télécharger dans le SADC-GLA. Il leur a été demandé de remplir les modèles de téléchargement et de rédiger un court rapport expliquant leur approche et leurs efforts de collecte de documents. Il leur a également été demandé de collecter plus de 50 documents dans leur pays. Un formulaire DPI a été partagé afin de recueillir l'autorisation des fournisseurs de données pour le partage des documents dans le SADC-GLA. La date limite initiale pour cette activité était mai 2020 mais, en raison du verrouillage, elle a été repoussée à août 2020

Tableau 0-1 Nombre de jeunes professionnels engagés par État membre.

#	État membre	Nombre de jeunes professionnels
1	Angola	2
2	Botswana	2
3	Comores	2
4	RDC	- −
5	Eswatini	2
6	Lesotho	2
7	Madagascar	- −
8	Malawi	2
9	Maurice	1
10	Mozambique	2
11	Namibie	2
12	Seychelles	- −
13	Afrique du Sud	2
14	Tanzanie	2
15	Zambie	2
16	Zimbabwe	2
	Total	*25*

Au total, 476 nouvelles fiches ont été collectées (tableau 0-2). 59% d'entre eux étaient déjà disponibles en ligne, tandis que 9% d'entre eux ont été mis à disposition pour téléchargement dans le SADC-GLA (formulaire DPI rempli). 19 % des documents ont été partagés sans autorisation DPI et 13 % des enregistrements ont été partagés sans hyperlien ou copie numérique. Plusieurs jeunes professionnels ont signalé des problèmes avec les formulaires DPI : ils ont déclaré que les fournisseurs de données étaient soit effrayés par le formulaire, soit incertains quant à leur droit de le signer. Bien que les jeunes professionnels de Tanzanie aient abandonné et que plusieurs jeunes professionnels n'aient pas atteint le seuil de 50 enregistrements, le nombre d'enregistrements collectés dans le cadre de cette activité est satisfaisant.

Il est important de prendre en compte les difficultés rencontrées par les jeunes professionnels en raison des pandémies. La plupart d'entre eux étaient enfermés chez eux, avec une mauvaise connexion internet. Ils n'ont pas été en mesure de visiter les bureaux où les archives sont stockées et d'interviewer les collègues d'autres organisations. Ces organisations pourraient être engagées dans les années à venir pour contribuer à la SADC-GLA.

En outre, le SADC-GLA n'a été mis en ligne que vers la fin du projet. Les jeunes professionnels ont été initiés au nouveau système lors d'un atelier virtuel en août 2020. Cet atelier a été très utile pour expliquer l'objectif de la collecte de données et répondre aux questions des jeunes professionnels, comme l'utilisation du formulaire DPI. Il aurait été plus facile pour les jeunes professionnels si la formation avait eu lieu avant le début de leur mission, et la qualité et la cohérence globales des contributions auraient pu être meilleures. Les jeunes professionnels ont également noté que le formulaire de téléchargement en ligne

est plus convivial que le modèle de téléchargement dans Excel. Il est probable que le formulaire en ligne devienne le principal canal de collecte de nouveaux enregistrements à l'avenir.

Pour l'équipe du projet, il a été difficile de fournir des instructions claires aux jeunes professionnels avant que la base de données ne soit entièrement développée (par exemple, deux versions du modèle de téléchargement ont été partagées, le formulaire DPI a été partagé après le modèle de téléchargement), ce qui était la conséquence d'un plan de travail serré.

Outre ces difficultés de collecte de données, l'IGRAC a eu beaucoup de mal à examiner les contributions des jeunes professionnels (plusieurs semaines au total). Les principaux problèmes concernaient la qualité des données saisies dans les modèles de téléchargement, en particulier l'identification des éditeurs et le nom de fichier des copies numériques soumises (souvent trop long et difficile à suivre). Les enregistrements incomplets ou non valides ont été corrigés. Ces contrôles de qualité ont également permis d'améliorer la cohérence de l'ensemble des contributions. Dans les années à venir, on s'attend à ce que les administrateurs de la SADC-GLA doivent consacrer une quantité importante de temps à l'examen des nouvelles entrées.

Tableau 0-2 Nombre d'enregistrements par État membre.

Pays	Nouveaux records	Enregistrements avec lien hypertexte	Enregistrements numériques copie		Documents sans hyperlien ou copie numérique
			avec le consentement du DPI	Sans DPI consentement	
Angola	29	- –	19	4	6
Botswana	29	- –	- –	- –	29
Comores	20	- –	18	2	- –
eSwatini	13	9	- –	3	1
Lesotho	26	10	6	9	1
Malawi	43	31	- –	12	- –
Maurice	57	32	- –	- –	25
Mozambique e	43	35	- –	9	- –
Namibie	44	44	- –	38	- –
Afrique du Sud	59	59	- –	- –	- –
Tanzanie	- –	- –	- –	- –	- –
Zambie	26	14	- –	12	- –
Zimbabwe	48	47	- –	- –	1
Total	*476*	*281*	*43*	*89*	*63*
Total (%)	*100%*	*59%*	*9%*	*19%*	*13%*

Les modèles de téléchargement final, les copies numériques, les rapports et les formulaires DPI sont disponibles dans Google Drive :

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

Résumé de la formation sur les archives documentaires sur les eaux souterraines de la SADC

Aperçu

Une formation en ligne pour la SADC-GLA a eu lieu les mercredi 19 et jeudi 20 août 2020. L'objectif était de présenter le portail web SADC-GLA, la base de données et les protocoles de mise à jour des données aux différentes parties prenantes - gestionnaires, utilisateurs et contributeurs - de la région. La formation a été suivie par 26 participants de 12 pays de la région SADC (Malawi, Namibie, Zambie, Botswana, Comores, Eswatini, Lesotho, Maurice, Mozambique, Afrique du Sud, Tanzanie et Zimbabwe), ainsi que par Raquel Sousa (RS) d'IGRAC et James Manda et Brighton Munyai de SADC-GMI. La formation était dirigée par Kirsty Upton (KU), Ken Lawrie (KL), Simon Burden (SB) et Martin Nayembil (MN) du BGS. Le programme complet de la formation figure à l'annexe 1.

Résumé des commentaires/problèmes soulevés

Saisie générale des données

- La plupart des gens, mais pas tous, ont préféré le formulaire en ligne - il est important de conserver les deux méthodes de saisie des données.
- Clarifier le processus de saisie des données pour le modèle Excel et le formulaire en ligne.
    - Est-ce que la SADC-GMI traite les données provenant des emails, ou est-ce qu'elle se contente de vérifier et de transmettre les emails au BGS ?
    - Les courriels provenant du formulaire en ligne doivent-ils être traités dans un format différent qui permet une entrée directe dans les tableaux, plutôt que d'utiliser le modèle Excel comme étape intermédiaire ?
    - Clarifier la façon dont les utilisateurs soumettent les PDF lorsqu'ils utilisent le formulaire en ligne - joindre aux courriels ? - et fournir le modèle de DPI en téléchargement (à joindre également au courriel).
    - ACTION : développer le flux de travail montrant les étapes du processus de téléchargement/saisie des données, à partager avec la SADC-GMI.
- Clarifier la différence entre contribution et consentement DPI sur le formulaire en ligne et le modèle Excel.

Éditeurs

- La base de données et le formulaire en ligne n'en autorisent qu'un seul, alors que le modèle Excel en autorise plusieurs.

Mots clés

- Ces derniers sont difficiles à saisir dans le modèle Excel.
- ACTION : KL doit inclure un menu déroulant pour TOUS les

  mots-clés.
- Il se peut qu'il en manque certains (les participants donneront

  leur avis).
- ACTION : KL ajoutera des indications sur le nombre de mots

  à saisir Formulaire en ligne

Online Form

- Besoin de plus de conseils/de notes à côté de chaque champ.
    - ACTION : SB ajoutera des conseils basés sur les notes contextuelles dans le modèle Excel.
- ACTION : SB fournira une fonctionnalité permettant de supprimer un auteur/mot-clé/pays de la liste de sélection.

Modèle Excel

- Les participants ont des difficultés à modifier le modèle en raison du verrouillage
    - ACTION : KL/KU font un suivi avec RS pour discuter du type de modifications nécessaires et voir si elles peuvent être activées.
- Formats de date - problèmes si les formats de date ne correspondent pas (par exemple, JJ/MM/AAAA ou MM/JJ/AAAA).
- Les boîtes d'orientation contextuelles sont gênantes
    - ACTION : KL va étudier la possibilité de les déplacer (on pourrait également ajouter une note sur la page de garde indiquant que la touche ESC ferme la boîte).
- De nombreux participants ont essayé de remplacer les lignes d'exemple
    - ACTION : KL ajoutera des conseils pour commencer en dessous de celles-ci, ou les déplacera sur une feuille séparée ?

Qualité des données

- Problèmes soulignés par l'IGRAC :
    - Confusion entre les noms d'auteurs et les noms de famille
    - Remplir les mauvais champs pour le type de document
    - Mots clés limités
    - Duplicata (seulement environ 10%)
- Une liste complète des problèmes mis en évidence par l'exercice des participants est fournie à l'annexe 2.
- Les problèmes les plus courants sont les suivants :
    - Orthographe/ponctuation, etc.
    - Données manquantes
    - Données dans des champs non pertinents pour le type de document
    - Inclure les initiales de l'auteur pour une organisation
    - Trop peu de mots-clés
    - Nom du contributeur manquant
    - Consentement au DPI lorsqu'aucun PDF n'est requis ou que le contributeur n'est pas le propriétaire.
    - Subjectivité des mots-clés

Résumé des réflexions et des commentaires des participants

- La SADC-GLA est une ressource très précieuse pour la région pour le partage et l'accès à l'information - elle a été considérablement améliorée par rapport aux précédentes archives de littérature grise de la SADC et aux archives de littérature sur les eaux souterraines en Afrique.
- La formation a permis de mieux comprendre le nouveau système.
- Les participants aimeraient jouer le rôle de formateurs pour leurs collègues, en utilisant les ressources de la session de formation - peuvent-elles être traduites ?
- L'exercice des participants a été utile

Résumé des réflexions des formateurs

- La formation a été extrêmement utile du point de vue des tests utilisateurs - nous avons obtenu de nombreux commentaires très utiles sur la façon dont le formulaire en ligne et le modèle peuvent être améliorés.

Annex 1: Training Schedule

Heure (SAST)	Heure (BST)	Contenu de la session	Responsable de la Session
Mercredi 19 Aout			
10 :30-10:50	09 :30-09:50	**Introduction** Objectifs de la SADC-GLA, besoins des utilisateurs et caractéristiques générales du nouveau système.	Kirsty
10 :50-11:30	09 :50-10:30	**Introduction au portail web** Visite du portail web, y compris la fonctionnalité de recherche, les résultats de la recherche et la fonctionnalité de saisie de l'utilisateur.	Simon
PAUSE			
11 :45-12:30	10 :45-11:30	**Introduction à la base de données** Introduction au modèle de données, aux métadonnées et aux questions d'assurance qualité.	Martin & Ken
Déjeuner			
14 :00-15:00	13 :00-14:00	**Mise à jour de la SADC-GLA** Introduction aux protocoles de saisie des données. Introduction à la tâche des participants : chaque participant reçoit un ensemble de documents à saisir via le formulaire en ligne et un modèle Excel. formulaire en ligne et un modèle excel	Simon & Kirsty
PAUSE			
15 :15-16:45	14 :15-15:45	**Mise à jour de la SADC-GLA (suite)** Tâche des participants : les participants font l'exercice individuellement ; la réunion ZOOM reste ouverte pour que les participants puissent poser des questions ; les participants envoient les modèles Excel complétés lorsqu'ils ont terminé.	Kirsty & Ken
Jeudi 20 Aout			
11 :00-11:45	10 :00-10:45	**Mise à jour de la SADC-GLA (suite)** Feedback sur l'exercice des participants	Kirsty & Ken
11 :45-12:30	10 :45-11:30	**Mise à jour de la SADC-GLA (suite)** Du formulaire en ligne et du modèle excel à la base de données – qu'est-ce qui est nécessaire ?	Ken
12 :30-13:00	11 :30-12:00	**Clôture**	Kirsty

DOCUMENT 1 : Gestion intégrée des ressources en eau et plan d'eau du Botswana

- Type de document
    - Rapport (lllllllllll)
    - Document de politique générale
- Éditeur
    - Gouvernement du Botswana (lll)
    - PNUD, GEF
    - Département des affaires de l'eau Ministère des minéraux, de l'énergie et des ressources en eau (lllll)
    - Ministère des minéraux, de l'énergie et des ressources en eau (lll)
    - Centre de recherche appliquée et Aurecon(l)
- Auteur
    - Aucun !
    - Ministère des affaires de l'eau... (lll)
    - Centre pour la recherche appliquée et Aurecon Botswana (llllll)
    - Aurecon Botswana (l)
- Mots-clés
    - Approvisionnement général en eau
    - IWRM (lll)
    - Protection générale des eaux souterraines (ll)
    - Eau potable
    - Changement climatique
    - Assainissement général
    - Qualité générale des eaux souterraines
    - Écologie générale
    - Moyens de subsistance
    - Gestion des ressources en eaux souterraines (llll)
    - Politique(l)
    - Caractérisation générale des aquifères
    - Approvisionnement général en eau (l)
    - Législation

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

- Mots-clés

- Hydrologie
- Hydrogéologie générale d'ingénierie
- Salinité
- Climat

DOCUMENT 3 : Les défis de la gestion des ressources en eau souterraine dans les aquifères côtiers d'Afrique de l'Est : Enquêtes et leçons apprises dans les îles Comores, au Kenya et en Tanzanie.

- **Mots-clés**

- Gestion des ressources en eaux souterraines (llllll)
- Evaluation des ressources en eaux souterraines
- Eau potable
- Eaux souterraines urbaines
- Protection générale des eaux souterraines

- Qualité générale des eaux souterraines (ll)
- Caractérisation générale des aquifères (lll)
- Changement environnemental général (ll)
- Vulnérabilité
- Population
- Approvisionnement général en eau
- Géophysique
- Salinité
- Gouvernance (ll)
- Géologie générale ou type d'aquifère
- Hydrogéologie générale d'ingénierie

- Auteurs
- Auteurs manquants
- Détails du journal
- Titre de la revue manquant
- Mélange du titre de l'article et du titre de la revue

DOCUMENT 4 : ENQUÊTE SUR LES EAUX SOUTERRAINES DE CUVELAI-ETOSHA

- Éditeur
  - Ministère de l'Agriculture, des Eaux et des Forêts (lllll)
  - Département des affaires de l'eau, Ministère de l'agriculture... (lll)
  - Bittner Water Consult (ll)
  - Manquant
  - République de Namibie
- Auteur :
  - Bittner Water Consult (lllllll)
  - Département des affaires de l'eau, Foresterie, Division de l'hydrologie (llll)
  - Département des Affaires de l'Eau, Ministère de l'Agriculture...
- Mots-clés
  - Qualité générale des eaux souterraines (ll)
  - Développement général
  - Approvisionnement général en eau (ll)
  - Eaux souterraines urbaines
  - Données sur les eaux souterraines
  - Niveaux des eaux souterraines
  - Gestion des ressources en eaux souterraines
  - Protection générale des eaux souterraines
  - Développement général des forages
  - Géophysique
  - Géologie générale ou type d'aquifère
  - Propriétés de l'aquifère
  - Surveillance des eaux souterraines (ll)

- Évaluation des ressources en eaux souterraines (ll)
- Exploration générale des eaux souterraines (ll)
- Caractérisation générale des aquifères
- Géologie
- Pluviométrie
- Hydrologie

DOCUMENT 6 : GESTION DES DONNÉES SUR LES EAUX SOUTERRAINES PAR LE MINISTÈRE DE L'IRRIGATION ET DU DÉVELOPPEMENT DE L'EAU DU MALAWI - UNE ÉTUDE DE CAS

- Éditeur
  - Manquant
- Mot clé :
  - Données sur les eaux souterraines (ll)
  - Approvisionnement général en eau
  - Gestion des ressources en eaux souterraines (llllll)
  - Exploration générale des eaux souterraines
  - Environnement général
  - Gouvernance générale (ll)
  - Protection générale des eaux souterraines (ll)
  - Géologie générale ou type d'aquifère

GÉNÉRALITÉS

- Questions pour lesquelles les lignes d'exemples ont été remplacées, plutôt que de commencer de nouvelles lignes sous les exemples.
- Dates de contribution
- Numéros d'entrée de document manquants
- Mots clés de sous-thème manquants
- Fautes d'orthographe
  - Majuscules modifiées (les données doivent être saisies exactement, sans modification)
    - Raison - peut créer des problèmes de duplication
    - Rend la recherche plus difficile en créant des différences
  - Substitution (remplacement de "&" par "et")
- Ponctuation/caractères supplémentaires lorsqu'ils sont copiés d'un document.
- Données dans les colonnes non pertinentes pour le type de document
- Pour une ressource externe - URL - le document et le DPI ne sont pas nécessaires.
- Ne pas approuver les DPI sans le consentement du propriétaire.
- Copie numérique - oui seulement si vous fournissez un PDF - si vous entrez une URL vers un autre site web où le document est disponible, ceci (et le nom du document) n'est pas requis.
- Pour tous les champs comportant une liste déroulante, vous devez uniquement choisir dans cette liste (pas de modifications ni d'ajouts).
- N'oubliez pas d'indiquer le nom et la date du contributeur
- Pays manquant
- Partie de l'URL manquante
  - Doit donner des instructions sur la façon de copier le lien à partir de la barre de fenêtre du navigateur !
- Les noms des organisations doivent figurer dans la colonne J uniquement - pas dans la colonne des initiales (I).

- Il est important d'obtenir des commentaires sur les documents !

- URLs manquants !
  - Doit donner des instructions sur la façon de copier le lien à partir de la barre de la fenêtre du navigateur !
- Remplissez toutes les colonnes pertinentes !!!!!
- Visez plus d'un mot-clé - la plupart des publications peuvent en contenir plus d'un.
- S'il y en a plusieurs (auteurs, mots-clés, langues, pays, etc.), chacun doit être sur une nouvelle ligne.
- Ne laissez pas de caractères parasites dans le tableur (si vous mettez quelque chose dans un champ, temporairement ou par erreur, supprimez-le).
- N'UTILISEZ PAS la fonction de remplissage automatique, surtout dans les champs contrôlés par des listes déroulantes, car elle peut permettre la saisie de valeurs non répertoriées.
- N'UTILISEZ PAS le copier-coller lorsque vous saisissez des données dans les champs de la liste déroulante, car cela interrompt le contrôle.

Annexe 7 : Stratégie de gouvernance

Stratégie de gouvernance des archives de littérature grise de la SADC

Septembre 2020

Auteurs :

Patrick Bell, Martin Nayembil & Kirsty Upton (BGS)

James Manda (SADC-GMI)

Table des matières

Objectif de ce document 79
Contexte du projet 79
Discussions précédentes sur la gouvernance 79
Composantes clés de la gouvernance 80
Hébergement et maintenance 80
Téléchargement des données 80
Transfert à la SADC-GMI 81
Contrat de maintenance 81
Coût de la maintenance 82
Obligations du BGS 83
Obligations de la SADC-GMI 83
Protocoles de saisie des données 83
Saisie directe des données dans la SADC-GLA 83
Entrée de données à partir de la SADC-GIP 84
Liens vers d'autres bases de données 84
Autres (à confirmer par les PJ) 84
Suivi de l'utilisation de la SADC-GLA 85
Personnes de contact 86
Annexe 1 : Phase 1 de la saisie des données (utilisateur de la SADC-GMI) 87
Annexe 2 : Phase 2 de la saisie des données (SADC-GMI à BGS) 87
Annexe 3 : Étape 3 de la saisie des données (entrée dans la base de données active) 88

Contexte du projet

Cette stratégie de gouvernance décrit comment les archives documentaires sur les eaux souterraines de la SADC (SADC-GLA) seront hébergées, entretenues et mises à jour après la fin du projet de développement (à partir du 31 octobre 2020). Un contrat de maintenance sera mis en place entre le British Geological Survey (BGS) et l'Institut de gestion des eaux souterraines de la SADC (SADC-GMI) pour une période de deux ans, du 1er novembre 2020 au 31 octobre 2022. Cette stratégie définit les rôles et responsabilités du BGS et du SADC-GMI pendant la période du contrat de maintenance, y compris les mécanismes par lesquels les nouvelles données sont entrées dans le SADC-GLA. Elle décrit également le processus de transfert au SADC-GMI à la fin de la période de maintenance. La stratégie a été préparée par le BGS et le SADC-GMI avec la contribution des partenaires du projet, le Centre international d'évaluation des ressources en eau souterraine (IGRAC).

Contexte du projet

SADC-GMI est un centre d'excellence régional pour la gestion des eaux souterraines, dont la vision est d'assurer l'utilisation et la protection équitables et durables des eaux souterraines dans la région de la SADC. En 2019, SADC-GMI a lancé un projet de mise à jour des archives de littérature grise de la SADC, afin d'en faire des archives en ligne entièrement fonctionnelles de la littérature sur les eaux souterraines pour la région de la SADC. Le projet a débuté en février 2020 et se terminera en octobre 2020. Il est mis en œuvre par le BGS en collaboration avec l'IGRAC.

En 2010, le BGS a développé les archives de littérature grise de la SADC5, un catalogue en ligne avec près de 3000 rapports non publiés, qui étaient consultables via un portail web par titre, auteur et pays. Depuis 2014, le BGS s'est appuyé sur ce catalogue pour développer les archives de littérature sur les eaux souterraines en Afrique6 (AGLA), dont les fonctionnalités ont été améliorées et qui contiennent près de 1000 documents supplémentaires pour la région SADC. Depuis 2017, l'IGRAC a développé, avec SADC-GMI, le portail d'information sur les eaux souterraines de la SADC (SADC-GIP), un portail web pour le partage de données et d'informations relatives aux eaux souterraines dans la région de la SADC. Le SADC-GIP est en cours d'amélioration par l'IGRAC dans le cadre d'un autre projet avec le SADC-GMI.

L'objectif principal du projet SADC-GLA est de fournir un accès à la littérature sur les eaux souterraines dans la région de la SADC à tous les acteurs concernés, afin de soutenir l'utilisation et la gestion équitables et durables des eaux souterraines dans cette région. Pour atteindre cet objectif, l'équipe du projet a développé un nouveau système, qui peut être facilement mis à jour et recherché, qui est technologiquement avancé et qui contient des informations à jour provenant des parties prenantes de la région SADC.

Discussions antérieures sur la gouvernance

Les termes de référence du projet, tels que définis par SADC-GMI, stipulaient que l'objectif clé était de "créer des archives entièrement fonctionnelles et bien approvisionnées qui deviennent une ressource pour les chercheurs, les acteurs du secteur et toutes les autres parties prenantes de la région SADC". Suite aux tâches clés de mise à jour de la GLA, les consultants devaient : "prévoir des rôles d'utilisateur à appliquer pour les différents niveaux d'administration des archives" et "étudier la migration possible des archives vers un serveur SADC-GMI approprié".

[5] https://www.bgs.ac.uk/sadc/index.cfm

[6] https://www.bgs.ac.uk/africagroundwateratlas/archive.cfm

En réponse à cela, la proposition du consultant a décrit un plan pour développer un système de base de données moderne, flexible et portable, avec des protocoles pour ajouter, éditer et mettre à jour les données, en tenant compte des différents rôles des utilisateurs, des DPI et du contrôle de qualité. Les consultants ont accepté d'héberger le système mis à jour sur l'infrastructure convenue, le système étant initialement hébergé par BGS, avec la possibilité de le transférer à SADC-GMI si/quand leur capacité le permet.

Ceci a été consolidé lors de la réunion de lancement du projet (février 2020), où il a été convenu que BGS hébergerait le SADC-GLA à court et moyen terme, avec l'ambition de le transférer au SADC-GMI à long terme. Il a également été convenu que la SADC-GMI est le gestionnaire de la GLA en tant que service pour les États membres, James Manda et Brighton Munyai étant les premiers responsables des archives.

La gouvernance a été étudiée plus en détail au cours du processus de consultation des utilisateurs (Work Package 1), dont les résultats ont été résumés dans le rapport d'analyse de la situation (mars 2020) et discutés lors de la réunion du projet le 17 mars 2020. Au cours de ce processus, les utilisateurs du système (représentés par le personnel de la SADC-GMI) ont souligné la nécessité de :

1. Améliorer la façon dont le contenu du SADC-GLA est maintenu, y compris l'ajout de nouveaux documents et la mise à jour des entrées existantes dans la base de données.
2. Contrôler la façon dont le site web est utilisé afin d'informer les améliorations futures Il a également été convenu que :

1. Il est au-delà de la portée de ce projet de développer une interface web qui fournit aux utilisateurs un login et la possibilité de télécharger directement dans la base de données hébergée par le BGS.
2. Le BGS et l'IGRAC devraient explorer les options pour relier la SADC-GLA avec la SADC-GIP, qui est potentiellement une source continue de documents/publications qui pourraient être inclus dans la GLA.
3. L'API permet à d'autres systèmes de se lier au système et peut également se lier à d'autres qui sont découvrables par le biais d'un API.

Composantes clés de la gouvernance

Hébergement et maintenance

Comme indiqué ci-dessus, il a été convenu que le BGS continuera à héberger le système SADC-GLA pendant une période de deux ans, conformément à un contrat de maintenance (décrit ci-dessous). A la fin de cette période de deux ans, le SADC-GLA sera soit transféré au SADC-GMI, si leur capacité le permet, soit le contrat sera renouvelé.

L'hébergement et la maintenance continus de l'interface web, de l'API et de la base de données impliqueront de répondre aux mises à jour nécessaires pour remédier à toute vulnérabilité en matière de sécurité qui serait découverte et de passer périodiquement à une nouvelle infrastructure en fonction de tout changement d'étape dans les protocoles sous-jacents du BGS. Cela n'affectera pas la fonctionnalité du système.

Téléchargement des données

Comme indiqué ci-dessus, il n'est pas possible, dans le cadre du projet de développement, de fournir un accès à distance permettant aux utilisateurs de télécharger des données directement dans la base de données SADC-GLA lorsqu'elle est hébergée par le BGS.

BGS en raison du temps considérable nécessaire pour développer une méthode de saisie des données basée sur un serveur client ou sur le web, avec la sécurité, l'assurance qualité et le contrôle qualité requis.

Le BGS continuera donc à télécharger de nouvelles données et de nouveaux documents, qui ont été entièrement vérifiés et dont la qualité a été assurée par la SADC-GMI, et qui sont fournis par le biais des protocoles de saisie de données décrits ci-dessous. Le téléchargement des données sera inclus dans le contrat de maintenance, comme décrit ci-dessous.

Remise au SADC-GMI

Un transfert initial aura lieu à la fin du contrat du projet de développement en octobre 2020, avec une documentation technique complète fournie à SADC-GMI. Il était initialement prévu d'organiser un atelier de formation en Afrique du Sud, mais cela n'a pas été possible en raison de la pandémie mondiale de Covid-19. Au lieu de cela, une formation en ligne d'un jour et demi a eu lieu en août 2020, à laquelle ont participé 30 parties prenantes de 12 pays de la région SADC, y compris SADC-GMI. La formation en ligne a permis de présenter le portail web, la base de données et les méthodes de saisie des données, y compris des instructions détaillées sur l'utilisation du formulaire en ligne et du modèle excel, ainsi qu'une discussion approfondie sur les métadonnées et la qualité des données.

La deuxième phase du transfert aura lieu lorsque le système complet sera transféré à SADC-GMI à la fin du contrat de maintenance. Dans le cadre de ce transfert, BGS examinera les capacités techniques de SADC-GMI, conditionnera la base de données et l'application en conséquence, et fournira des directives et une formation sur l'installation, la maintenance continue et le chargement des données dans le système. Un budget est inclus dans le contrat de maintenance pour une formation technique plus longue, dans le pays, dans le cadre de ce transfert.

Contrat de maintenance

Le BGS hébergera, entretiendra et téléchargera les données vers le SADC-GLA selon un contrat de maintenance. Le contrat de maintenance sera basé sur les éléments suivants :

- BGS hébergera le portail web, l'API et la base de données sur nos propres serveurs pendant deux ans, après quoi le contrat sera soit revu et renouvelé, soit résilié ;
- Si le contrat est résilié après deux ans, la SADC-GMI prendra l'entière responsabilité du système ;
- Le contrat peut être résilié plus tôt si la SADC-GMI est en mesure d'assumer la responsabilité de l'hébergement et de la maintenance du GLA ;
- La résiliation nécessitera un préavis d'au moins 3 mois afin de laisser le temps à la base de données et à l'application d'être emballées avec des directives actualisées pour l'installation et la préparation de la formation dans le pays ;
- Pendant la période du contrat de maintenance, l'hébergement et la maintenance de l'interface web, de l'API et de la base de données consisteront à répondre aux mises à jour nécessaires pour remédier à toute vulnérabilité en matière de sécurité et à passer périodiquement à une nouvelle infrastructure en fonction de tout changement d'étape dans les protocoles sous-jacents du BGS ; cela n'impliquera aucune modification de la fonctionnalité du système ;
- Pendant la période du contrat de maintenance, le téléchargement des données sera également effectué par le BGS ; les données seront fournies au BGS conformément aux protocoles de saisie des données décrits ci-dessous ; ces données devront être entièrement vérifiées et leur qualité assurée afin de contribuer à maintenir l'intégrité de la base de données ; en fonction de la fréquence des nouvelles données fournies, le BGS téléchargera de nouvelles données soit sur une base ad hoc, soit à intervalles fixes de 2 mois si cela devient ingérable ;

- Lorsque la SADC-GMI assumera l'entière responsabilité de l'hébergement et de la maintenance de la SADC-GLA à la fin du contrat de maintenance, la BGS apportera son soutien au transfert du système ; les exigences techniques du transfert seront déterminées par les capacités internes de la SADC-GMI ; la base de données et l'application, avec des directives d'installation actualisées, seront emballées en conséquence.

Coût of Maintenance

Les remboursements non dépensés du contrat de développement (Tableau 1) seront transférés dans le contrat de maintenance pour couvrir une partie des coûts d'hébergement, de maintenance, de saisie des données et de transfert final, le reste étant financé par le BGS. Les coûts complets associés à l'hébergement, à la maintenance, à la saisie des données et au transfert final sont présentés dans le tableau 2, les coûts pour la SADC-GMI et la BGS étant clairement définis.

Article	**Coût unitaire**	**Quantité**	**Total (USD)**
Indemnités journalières	266	22	5,852
Vols internationaux	1463	4	5,852
Transport aller/retour à l'aéroport	322.5	4	1,330
Infrastructure informatique			4,655
TOTAL			**17,689**

Tableau 1 : Remboursements non dépensés dans le cadre du contrat actuel, en grande partie en raison des restrictions de voyage imposées par Covid-19.

Item	**Coût unitaire**	**Quantité**	**Total (USD)**
FINANCÉ À L'INTERNE PAR LE BGS			
Temps du personnel pour l'hébergement et la maintenance	$569.32 par jour	8 jours (4 par an)	4,554.56
Temps de travail du personnel pour la saisie/le téléchargement des données	$569.32 par jour	8 jours(4 par an)	4,554.56
Temps passé par le personnel à préparer le système pour le transfert	$569.32 par jour	10 jours	5,693.20
Total BGS			**14,802.32**
FINANCÉ PAR LA SADC-GMI PAR LE BIAIS DU CONTRAT DE MAINTENANCE			
Atelier de formation : temps du personnel	$569.32 par jour	20 jours	11,386.40

Atelier de formation : remboursements	$6302.60	1	6,302.60
Total SADC-GMI			**17,689.00**
COÛT TOTAL			**32,688.72**

Tableau 2 : Coût total de l'hébergement, de la maintenance, de la saisie des données et de la remise finale de la SADC- GLA.

Obligations du BGS

Comme indiqué ci-dessus, le BGS s'engagera à héberger et à entretenir le portail web, l'API et la base de données pour s'assurer qu'ils continuent à fonctionner avec leurs fonctionnalités existantes. Le BGS téléchargera également de nouvelles données dans la base de données conformément aux protocoles décrits ci-dessous. A la fin du projet, BGS fournira une documentation complète pour le nouveau système et lorsque le moment sera venu de remettre le système à SADC-GMI, BGS fournira le soutien et la formation nécessaires au personnel de SADC-GMI.

Obligations de la SADC-GMI

Bien que le BGS hébergera et maintiendra le SADC-GLA, il est convenu que le SADC-GMI est le gestionnaire du GLA en tant que service pour les Etats membres. La SADC-GMI est donc la première responsable de la promotion de l'utilisation de la SADC-GLA - à la fois par les utilisateurs du site web à la recherche de documents et par ceux qui contribuent aux documents - et de la qualité des données contenues dans les archives. Encourager les parties prenantes de la région à utiliser et à continuer à contribuer aux archives après la fin du projet sera la clé de son succès à long terme. Comme indiqué dans les protocoles de saisie des données ci-dessous, la SADC-GMI sera également responsable de l'assurance qualité des nouvelles données fournies au BGS pour le téléchargement.

Protocoles de saisie des données

Saisie directe des données dans le SADC-GLA

Il existe deux voies pour la saisie des données dans le SADC-GLA - un formulaire en ligne et un modèle excel téléchargeable - disponibles sur la page Contribuer du portail web du SADC-GLA (www.sadc-gla.org/SADC/contribute.html). Les protocoles de saisie des données pour chacune de ces voies sont décrits ci-dessous et illustrés dans les diagrammes de flux de travail des annexes 1 et 2.

1. **Formulaire en ligne**

Tout utilisateur peut saisir les détails des métadonnées d'un seul document, qui, une fois soumis, sont transmis à une API privée, générant un courriel à la SADC-GMI (info@sadc-gmi.org). La SADC-GMI convertit les données de l'e-mail au format Excel (voir ci-dessous) à l'aide d'un outil simple accessible via le portail web. Les données peuvent ensuite être copiées directement dans le modèle excel pour que le personnel du SADC-GMI procède à l'AQ. Le modèle AQ est ensuite transmis au BGS pour être téléchargé dans la base de données en ligne. Les utilisateurs peuvent fournir des copies électroniques de documents en les joignant sous forme de PDF à l'e-mail adressé à SADC-GMI, comme demandé dans le formulaire en ligne. Les PDF doivent être accompagnés d'un formulaire de consentement de contribution, confirmant qu'ils ont la permission de partager le document par le biais du SADC-GLA. Les PDF et les formulaires de consentement sont envoyés au BGS avec le modèle excel et sont stockés sur nos serveurs avec la base de

données.

2. Modèle Excel

Les utilisateurs peuvent télécharger un modèle excel directement à partir du portail web, saisir les détails des métadonnées pour plusieurs documents, puis envoyer un courriel à la SADC-GMI (info@sadc-gmi.org). Comme ci-dessus, les données contenues dans le modèle Excel sont contrôlées par le personnel du SADC-GMI, puis transmises au BGS pour être téléchargées dans la base de données en ligne. Les utilisateurs peuvent fournir des copies électroniques des documents en les joignant en tant que PDF à l'e-mail ou en utilisant un système de transfert de fichiers en ligne, selon la taille du fichier. Comme ci-dessus, les PDF doivent être accompagnés d'un formulaire de consentement de contribution, confirmant qu'ils ont la permission de partager le(s) document(s) par le biais de la SADC-GLA. Les PDF et les formulaires de consentement sont envoyés au BGS avec le modèle excel et sont stockés sur nos serveurs avec la base de données.

Pour les deux voies de saisie des données décrites ci-dessus, le BGS effectuera un contrôle supplémentaire des données avant de les charger dans la base de données active (voir le diagramme de flux de travail à l'annexe 3).

Il existe une fonction "signaler une erreur" sur le portail qui permet aux utilisateurs de signaler des erreurs ou des fautes dans les données existantes et de fournir des corrections via un simple formulaire. Quant au formulaire de saisie des données, il passera par une API privée générant un courriel à l'équipe SADC-GMI pour un contrôle de sens, avant d'être transmis à l'équipe de gestion des données du BGS pour être mis à jour dans la base de données.

Lorsque la GLA sera entièrement remise à SADC-GMI à la fin du contrat de maintenance, le BGS fournira un outil interne pour la saisie et la gestion directe des données. Il s'agira du même système que celui utilisé en interne par le BGS pour mettre à jour la GLA pendant notre période d'hébergement. Il y aura une fonction de téléchargement en masse et une implémentation client-serveur, très probablement dans un outil comme Microsoft Access ou similaire, où un utilisateur administratif pourra se connecter et gérer directement les enregistrements dans la base de données, enregistrement par enregistrement, et pourra entrer de nouveaux enregistrements individuels, y compris le téléchargement de documents. Cette solution sera disponible pour être installée sur les terminaux des utilisateurs, qui ont une connectivité à la base de données hébergée et à qui la SADC-GMI souhaite donner un accès administratif pour mettre à jour directement les enregistrements de la base de données. Elle ne sera pas disponible en tant qu'outil web mais seulement en tant qu'installation sur des terminaux informatiques individuels.

Rubrique de données à partir du SADC-GIP

Le nouveau système SADC-GIP offre une fonctionnalité de téléchargement de documents. Bien que le SADC-GIP soit principalement axé sur les données cartographiques/spatiales, la fonction de téléchargement n'est pas en mesure de différencier les différents types de documents (par exemple un document Excel contenant des données et un rapport PDF) et il n'est donc pas possible d'empêcher les utilisateurs de télécharger des rapports, etc. dans le SADC-GIP. Les deux systèmes nécessiteront donc un certain niveau de maintenance pour identifier les documents du GIP qui doivent être transférés dans le GLA. Cette opération pourrait être effectuée mensuellement par le personnel du SADC-GMI. Pendant le contrat de maintenance, le BGS/IGRAC cherchera à développer un protocole simple qui permettra de transférer facilement les documents du SADC-GIP au SADC-GLA.

Liens vers d'autres bases de données

La SADC-GLA doit rester liée aux archives de la littérature sur les eaux souterraines en Afrique, hébergées par le BGS, afin que les mises à jour des deux systèmes soient accessibles à tous les utilisateurs. Ce sera le cas tant que les deux systèmes seront hébergés par le BGS, mais si/quand le SADC-GLA sera transféré au SADC-GMI, une méthodologie devra être mise en place pour assurer qu'ils restent interopérables. Cela sera possible grâce à l'API de la SADC-GLA et de l'AGLA.

Autres (à déterminer par les PJ)

- - On ne sait pas encore très bien de quoi il s'agit - les PJ fourniront ces informations.
-
- - La méthode de liaison devra être décidée au cas par cas.
- - Le temps nécessaire à l'établissement de liens avec d'autres bases de données peut être limité en fonction de la contribution des PJ, mais l'API offre une fonctionnalité permettant de le faire à l'avenir si nécessaire.

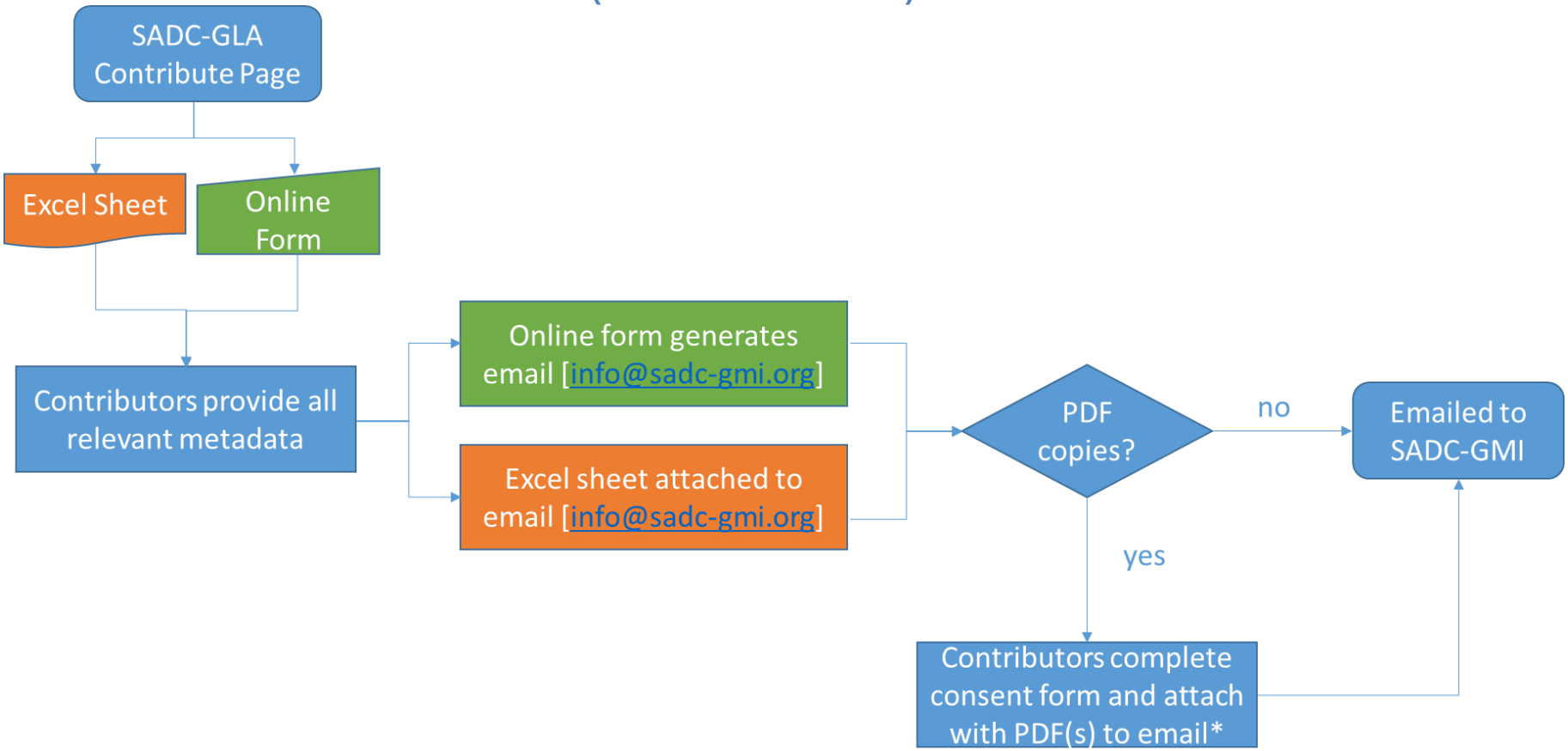
Suivi de l'utilisation de la SADC-GLA

BGS/IGRAC/SADC-GMI devrait surveiller l'utilisation du GIP et du GLA au cours des 18 prochains mois afin de faire des recommandations sur la façon dont ils devraient être maintenus à long terme. L'équipe utilisera Google Analytics et créera un tableau de bord accessible à la SADC-GMI, la BGS et l'IGRAC.

Personnes à contacter

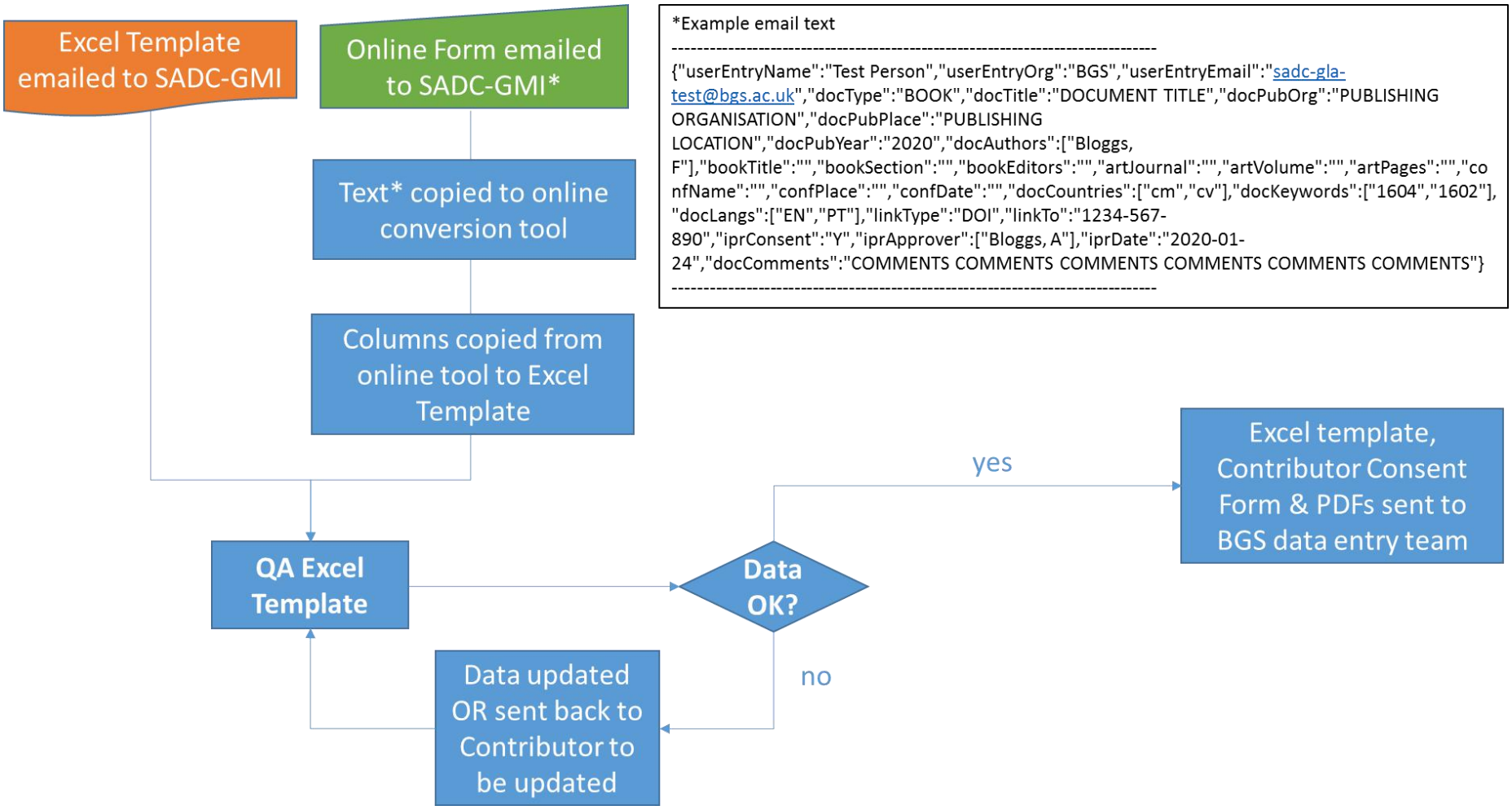
	SADC-Groundwater Management Institute (SADC-GMI) Mr. James Manda (Project Coordinator SADC-GMI) jamesm@sadc-gmi.org +265 991742582
	British Geological Survey (BGS) Dr. Kirsty Upton (Team leader) kirlto@bgs.ac.uk + 44 131 6500 210
	International Groundwater Resources Assessment Centre (IGRAC) Dr. Arnaud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

Annexe 1 : Phase 1 de la saisie des données (Utilisateur à SADC-GMI)

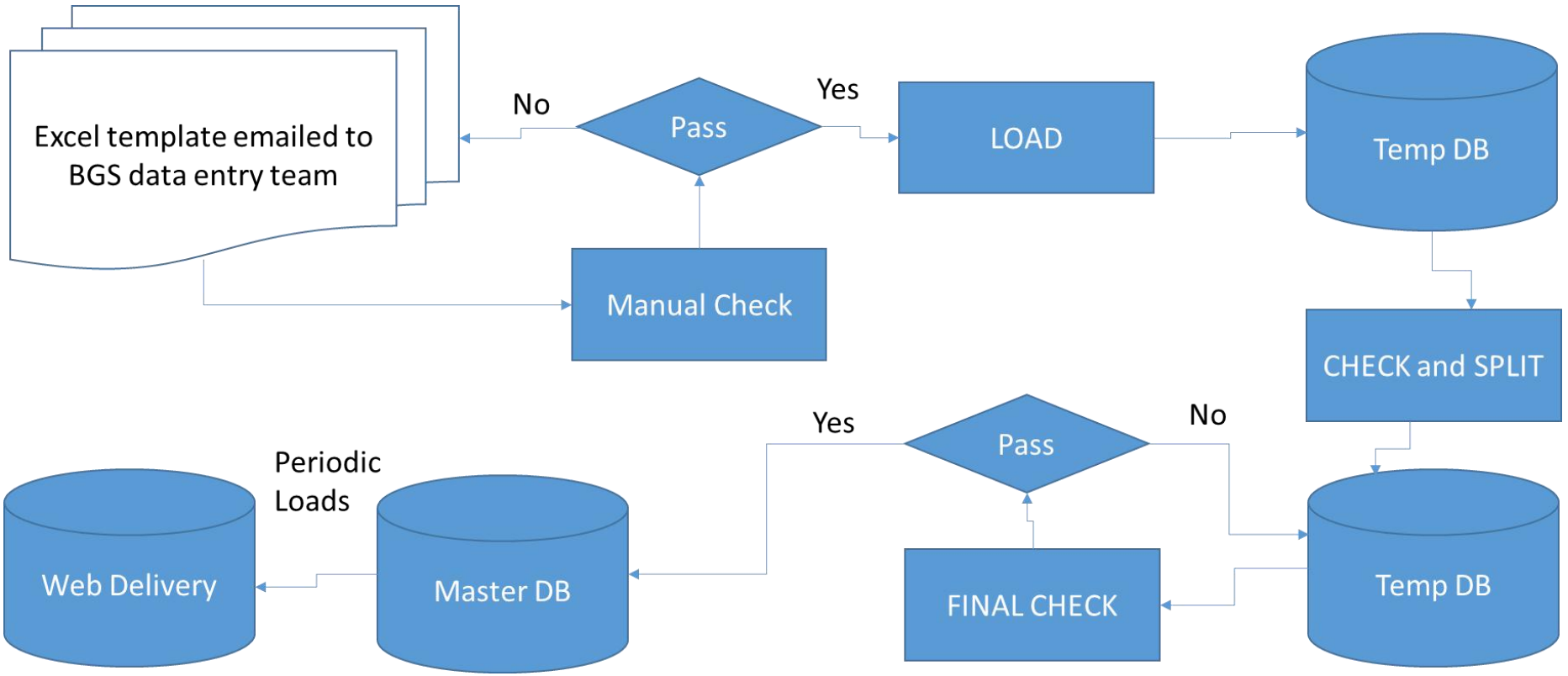


*If a Contributor has too many PDFs to send by email, this should be stated in the email and SADC-GMI should contact the Contributor to arrange online document transfer

Annexe 2 : Phase 2 de la saisie des données (SADC-GMI à BGS)



Annexe 3 : Étape 3 de la saisie des données (entrée dans la base de données active)


Excel template emailed to BGS data entry team
No
Pass
Yes
LOAD
Temp DB
Manual Check
CHECK and SPLIT
Yes
Pass
No
Periodic Loads
Web Delivery
Master DB
FINAL CHECK
Temp DB