2ND REGIONAL TRAINING WORKSHOP IN BIODIVERSITY INFORMATICS AND BIODIVERSITY WEB-PORTAL ADMINISTRATION

Organized by the Albertine Rift Conservation Society with funding support from the JRS Biodiversity Foundation

Chez Lando Hotel, Kigali 3 - 5 September, 2012

ARCOS, September 2012

Contents

Front Photo: Participants photo, with ARCOS Team and the Guest of Honour, Dr Marie Christine Gasingirwa, Director General for Science, Technology and Research, Ministry of Education, Rwanda (Centre),

Introduction

The Albertine Rift montane forests and freshwater ecosystems provide important services to the people of the region. However, due to high population and various development pressures, the threat to the biodiversity and ecosystems in the Albertine Rift region is alarming and very little is being done to monitor this trend in order to guide future actions and decision-making despite the recognized role of information in decision-making. One of the problems in the Albertine Rift region is not only the availability of information but also the lack of mechanisms to access, interpret and analyze the existing information and present it in a simple format to guide the policy-makers and the public in general.

In view of this, ARCOS has developed a program termed ***"Albertine Rift Biodiversity Monitoring and Information Systems (ARBMIS)*** and to achieve the objectives of this program, ARCOS with funding from JRS Biodiversity Foundation, is implementing a project called "*Albertine Rift Biodiversity Portal - Building competence for biodiversity Information Systems in the Albertine Rift region (July 2011-June 2013)".* The project aims at developing a regional biodiversity data portal, hub for information sharing in the region through a database driven website and an informative and educational web portal.

In the framework of this project, ARCOS has secured partnerships with different institutions holding data in the Albertine Rift region and MoUs have been signed whereby the parties agreed to partner and collaborate in collecting biodiversity data and information and make them readily available to the public in general and decision makers in particular, through the web-portal ARCOS was developing.

Another component of the project also concerns building capacity among the institutions' staff in collecting, analysing and publishing the correct and accurate data. In line with this, ARCOS regularly organizes trainings, visits and workshops in this strive to raise data centres' staff skills in the above-mentioned fields.

In this regard, ARCOS organized a 3-day training workshop on 3 -5 September, 2012 at Chez Lando Hotel, Kigali; a workshop whose theme was: **"Delivering Practical Regional Biodiversity Information Systems in the Albertine Rift"**

This workshop was the 2nd of its kind following the one held last December 2011 and in addition to the usual topics relating to bioinformatics, GIS, Remote sensing and Database Management Systems, participants were presented the prototype of the portal and taken through steps to be followed when uploading and publishing data through it.

Moreover, as part of working closely with BGIF Network, some sessions facilitated by TanBIF (GBIF Node in Tanzania) looked at the GBIF's Integrated Publishing Toolkit; a very convenient tool that eases the task of uploading data and allows for some processing of the data.

The present report gives a summary of the content of the sessions as well as a brief outline of the outcomes of the training.

1. Attendance

All data centres partners of ARCOS in this venture were invited to this training. Depending on its size, each institution was represented by 1 or 2 of its staff.
In addition to data centres, other partner institutions in the Albertine Rift as well as GBIF Nodes in the region were also invited as a way to promote collaboration and inter-institutional exchange throughout the region. Out of 23 invited, 18 attended the training: 3 people from DRC, 4 from Burundi, 1 from Uganda, 4 from Tanzania, 1 from Kenya and 5 from Rwanda (see annex 1 for a complete list of participants to the workshop).

The following table lists the facilitators for the different sessions of the training:

Names	Institution	Sessions facilitated	e-mail
Dr Sam Kanyamibwa	Albertine Rift Conservation Society	Overall facilitator	skanyamibwa@yahoo.com
Dr Marie Christine Gasingirwa	Ministry of Education, Rwanda	Opening ceremony	cgasingirwa@nur.ac.rw
Mrs Zachary Kimutai Maritim	Albertine Rift Conservation Society	Albertine Rift Regional Biodiversity Portal: presentation of the prototype and its features	Kimutaimz@gmail.com
Mr Apollinaire William	Antioch University (PhD Candidate)	GIS and Remote Sensing: Geo-referenced Data capture, management and analysis; Image enhancement, geo-referencing and classification.	williappollo2005@gmail.com
Mr Faith Nshimba	Tanzania Biodiversity Information Facility	Integrated Publishing Toolkit	fshimba@costech.or.tz
Mr Hafidh Fikiri	Tanzania Biodiversity Information Facility	Integrated Publishing Toolkit	

2. Training Report

3.1. Welcome Note and Opening speech

The training opened in a cheerful mood whereby new participants (those who didn't attend the previous training) introduced themselves and shared their expectations. Some of these expectations are: "…at the end of this training, I expect to leave with enough knowledge and skills to accurately collect, manage and analyse biodiversity data and be able to publish it through the web portal ARCOS has developed".

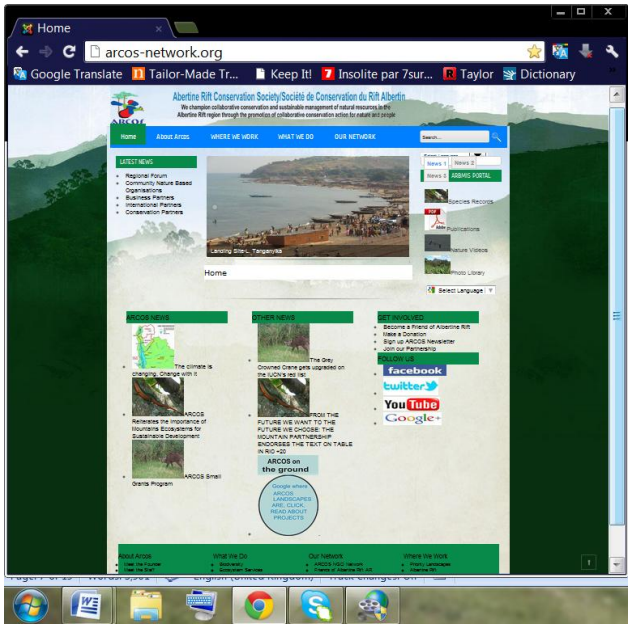
Welcoming the participants, Dr Sam Kanyamibwa; Executive Secretary of ARCOS; expressed his gratitude to everybody for having responded to the invitation. He explained that this training is not only a follow up of the one held last December 2011 but also a test field for the web portal prototype ARCOS has developed. He urged the participants to grasp this opportunity to build up a solid network of professionals throughout the region and at the same time become better focal persons for their respective institutions. He also thanked JRS Biodiversity Foundation for the funding support.

Dr Gasingirwa Marie Christine (left) and Dr Sam Kanyamibwa (right) during the opening ceremony

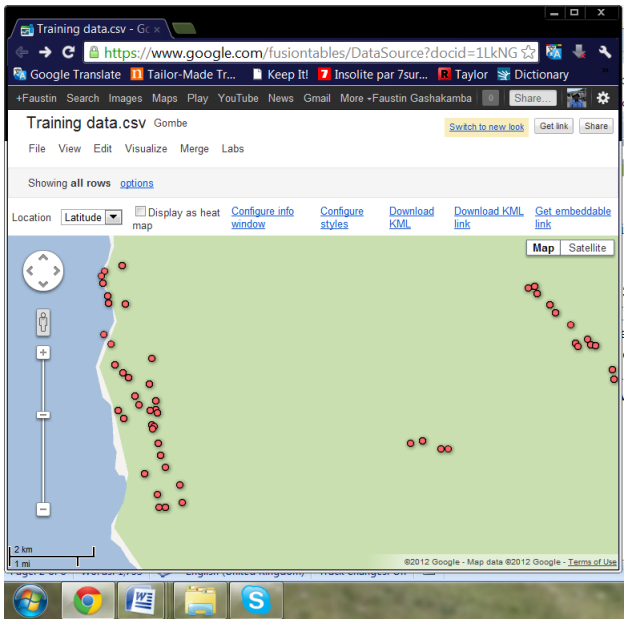
In her opening speech, Dr Marie Christine Gasingirwa; Director General for Science, Technology and Research in the Ministry of Education, Rwanda; congratulated ARCOS for having taken this initiative to establish this regional biodiversity data sharing and information exchange platform. She noted the urgency of fighting ignorance and poverty through "home-made" solutions and recognized that this is one step forward in this direction. She stressed on the fact that "accurate" biodiversity data and information is key to better decisions in the field of natural resource management and emphasized on this initiative as "highly commendable and noble".

3.2. Day 1: introducing the Albertine Rift Biodiversity Portal

Sessions under this sub-theme consisted of presenting the prototype of the web portal ARCOS has developed, explaining its features and exploring its functionality through various practical exercises including uploading data (tabular, reports, pictures, etc) on the portal and using dedicated tools to prepare data before uploading.
Participants were given different portable software to for instance compress pictures without losing their display quality and converting shape files to Keyhole Markup Language (KML) which is the format compatible with Google maps (same maps used on the portal).
During one session facilitated by Zachary Maritim Kimutai, participants showed a lot of enthusiasm when they were explained how geo-referenced data (species occurrence data for instance) can be displayed using the Google map application.



The Albertine Rift Biodiversity Data Portal is accessed from the ARCOS revamped website



The portal uses the Google map application to display geo-referenced data. Hence, one has to have a Gmail account in order to be able to access the application

The Albertine Rift Biodiversity Portal is integrated into the revamped ARCOS site. The "home" interface gives access to the main menus used to navigate the website (regional conservation news, ARCOS and its activities, contacting us, etc.). Moreover, from the homepage one is provided with a window to the data portal. To enter, one has to provide valid login credentials (username and password) and will then be presented with a different interface with links to all the tools necessary to navigate the portal (upload, download, filter, summarize, visualize data, etc.).

As for any other website, speed in loading and refreshing requires a lightweight structure. Therefore any data publisher is encouraged to take advantage of the tools available to compress and structure the data for more portability before uploading to the portal.

Since we don't have a server of our own yet, we have hosted the spatial component of the portal on Google servers. This means that a Google account is required to access some of the features of the portal such as browsing the maps produced from loaded geo-referenced data.

3.3. Day 2: Spatial Data Mobilisation, Remote Sensing and Landuse Planning

Sessions under this sub-theme underpinned topics like collecting data using GPS, how to project data and how to geo-reference an image.

a. GPS data collection, projections

Outside, the group was explained the nuts and bolts of using effectively GPS and split into 4 groups to practice collecting data around hotel. Back in the hall, the class was taken through steps to get data from the GPS, display and project it using the ESRI's ArcGIS 10.

b. Image geo-referencing

GPS is only one of the many sources of GIS data or geo-spatial data. Images (Satellite images, aerial photographs, scanned maps, etc) form also a big bulk of the data used in GIS. However, these images have to be processed in one way or another before they can be of any use for GIS analysis. Image geo-referencing and classification is the most important processing that have to be done to go from the image to meaningful data.

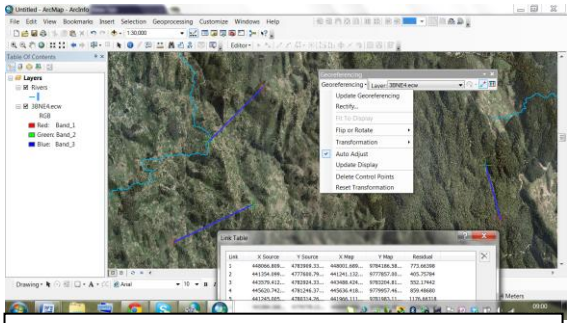


Image of Mukura landscape during a georeferencing session in ArcMap 10

Geo-referencing is the process of scaling, rotating, translating and deskewing an image to match a particular size and position. Not only satellite images or orthophotos are geo-referenced; scanned maps also have to be geo-referenced prior to their being exploited usefully in a GIS.

To successfully geo-reference an image, one has to recognize at least 3 "control points" that will make possible the transformation of the image to the right size and position. ArcGIS provides all necessary tools to geo-reference images and participants were taken through the process taking an example to the non-georeferenced orthophoto of Mukura (Western of Rwanda).

c. Image classification

The intent of the classification process is to categorize all pixels in a digital image into one of several land cover classes, or "themes" (Lillesand and Kiefer, 1994). Normally, images are conceptually a table of values arranged in a column x row layout whereas the values represent either the reflectance of emittance of the different features in a particular region of the electromagnetic spectrum. Therefore, the values stored in an unclassified image could be of little use for everyday applications.

For instance, if one has to derive land use map from the image, then the different DN statistics have to be matched to actual land use classes they represent. This can be done through "supervised" or "unsupervised" image classification.

During this session, the participants were taken through the process of supervised classification of a QuickBird image of Kigali using ArcGIS 10.
First of all, "training samples" were created following visual interpretation and a "signature file" was created and saved. Secondly, the image was classified using the maximum likelihood classification method. Finally, the resulting raster has been filtered, smoothed, generalized and vectorized to generate landuse polygons.

d. Introduction to remote sensing

Remote Sensing is the science and art of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area, or phenomenon under investigation. (Lillesand and Kiefer, 1994).
During this session, the participants were given a brief overview of what remote sensing is, how it relates to GIS and how RS tools and methods can be harnessed during the analysis of biodiversity data collected on ground. Concepts like image enhancement, image visualization and interpretation, image resolution (temporal, spectral, spatial and radiometric) were explained and the group took the opportunity to experience the feel of performing some of these operations.

3.4. Day 3: Data Management and Publishing

Sessions under this sub-theme sought to instill good data management skills and attitudes to participants; skills relating to data cleaning, quality control and structuring. Most importantly, participants were introduced to the Integrated Publishing Toolkit (IPT), a free tool developed and deployed by GBIF to facilitate the process of publishing biodiversity data through a web-portal environment.

a. Database Management Systems (DBMS)

This session facilitated by Mr Zachary Maritim Kimutai looked at the benefits of using a DBMS when dealing with large databases. In fact, DBMSs support very large datasets, maintain data correctness, support concurrent use, provide a high-level query language, use a data model, provide backup and recovery functions and reduce data redundancy. For this reason, a DBMS is a good tool to use when one wants to set up, maintain and possibly share data stored into a database.
Different DBMSs are used for different databases. For instance, ArcGIS comes with its own package called ArcCatalog which allows creation and maintenance of different types of geodatabases.
Non-complex text or even tabular data can be stored in a non-database format like an Excel sheet but it's always a good practice to store data in a database using a DBMS.

b. Integrated Publishing Toolkit; overview

The Integrated Publishing Tool has been developed by GBIF in 2009 to help parties publish biodiversity data and information on internet. This session aimed to bring the participants to understand what IPT is and what it is used for.
The IPT is a new among the big family of GBIF tools to make possible biodiversity data shared and exchanged in order to achieve the institution's mission to make biodiversity data freely available to benefit science and society.

The IPT is a Java web application developed to connect and serve 3 kinds of biodiversity data: Primary occurrence (distribution of organisms), Species names information and General resource metadata.
It is multilingual and comes with a mapping server to help display geo-referenced occurrence data in a map format. Most importantly, IPT is a very customizable and extensible tool. The best way to do it is through CSS files but for maximum flexibility one can even directly change the HTML code though this increases chances that the code becomes bugged or will not be compatible with future upgrades.
IPT supports the Darwin Core schema. However, the tool is so extensible through the use XML files so other data types can be availed to users and the GBIF secretariat are very approachable if one wants to make some.
In this session facilitated by Mr Faith Shimba and Hafidh Fikiri of TanBIF, participants were given the minimum skills of an average IPT user. The session didn't intend to cover all aspects of IPT, it only summarized the most relevant concepts like the tool's interface and how it can be customized, the software and hardware platform requirements to deploy it and different options to host it and pros and cons of each alternative.
TanBIF already uses IPT on its portal. Hence, participants opened accounts on the portal and practiced uploading, checking and analysing data using the tool on the TanBIF portal.

3. Training evaluation

At the end of the training, participants had some time to give their impressions on the training and reflect on ways for improvements in further similar events. Most of the bulk of the feedback went to praise the idea of developing such regional biodiversity data sharing and exchange platform and supporting this institutional empowerment exercise ARCOS has undertaken to make the idea workable.
The following are some of the thoughts expressed by participants caught in verbatim:

- *All topics were so interesting indeed that it would have been a good idea to increase the time for the training from 3 to 10 days.*
- *The GIS component of the training was the most helpful to me. I was already dealing with biodiversity data but wasn't getting all from it without the power or GIS tools to analyse spatially the data.*
- *It was a big challenge to work with different people, from different backgrounds and with different levels of technical skills in the topics covered. Thanks to the skilled trainers, we've managed to get along pretty well.*
- *The training went smoothly overall. However, it was a considerable challenge for some of us who didn't participate to the previous training. Institutions should be consistent in choosing representatives to send for such events.*

- *Logistically, the training ran awesomely. The topics covered are equally important and helpful to me. Nevertheless, would have appreciated more if the period was increased.*
- *More practice sessions should be planned. ARCOS should set up a mailing list to allow participants stay in touch*
- *A regular exchange platform in the form of an alumni community should be developed to help participants keep practicing the skill they have learnt.*
- *Further step should involve in-service trainings where ARCOS would facilitate such capacity building exercise within the institutions themselves.*

5. Planned follow-up activities

A special session was facilitated by ARCOS Executive Secretary, to give opportunity to partners to brainstorm on priority needs and suggestions for possible follow up project activities. This discussion was motivated by the fact that the current JRS project funding will end in June 2013. Participants suggested some interesting ideas, including reinforcing outreach efforts for general public, further institutional capacity development, education, and knowledge management. ARCOS and partner institutions will continue seeking further funding opportunities to ensure sustainability and possibly scale-up of this undertaking.

After this training, participants agreed to keep in touch from their respective institutions. An active online forum has been agreed upon whereby different issues related to biodiversity data management and GIS will be discussed.

Moreover, ARCOS is planning to continue building capacity of these institutions to empower them logistically to collect and share biodiversity data for supporting informed decision-making in the region for the benefit of our nature and people.

6. Conclusions and Closing ceremony

Participants were equipped with advanced knowledge and skills necessary to collect biodiversity data, analyse it using different tools and publish it through the ARCOS Regional Biodiversity Portal using the GBIF's Integrated Publishing Tool.
Special attention was accorded to thoroughly explaining the features of the portal and its administration. Hence, participants left with a feeling of familiarity and ease using the different tools integrated into the portal.
In addition to this, the workshop served as a platform to promote collaboration and exchange between professionals from diverse fields of knowledge in the region which will benefit the health of the Albertine Rift Information System ARCOS is developing.

The closing ceremony involved the short closing speech by Dr Sam Kanyamibwa who praised the enthusiasm and interest which participants have shown during the training. He noted that ARCOS being a network, the training should be a key to further networking and collaboration between present institutions and recommended that an alumni program should be developed to allow further exchange between the participants.
All successful participants were then awarded a certificate which attested their participation and the various skills gained during the training.
The event was capped with a convivial farewell cocktail with invitees and all ARCOS' staff.

Certificate copy (Left) and Issuance of certificates (Right)

Annex I: List of Participants

Institution	In full	Names of participant	e-mail
IGEBU	Institut Géographique de Burundi	Ms Karire Anita	anigkari@yahoo.fr
OVG	Observatoire Volcanologique de Goma	Mr Patrick Habakaramo	patrickhabak@yahoo.fr
LTA	Lake Tanganyika Authority	Mr Pacifique Ndoricimpa	pacifique.ndoricimpa@lta-alt.org
TAWIRI	Tanzani Wildlife Research Institute	Mr Chediel K. Mrisha	cmrisha@yahoo.com
		Mr Mwita Machoke	machoke2000@hotmail.com
Makerere University/NBDB	National Biodiversity Data Bank	Dr Mathias Behangana	behangana@yahoo.com
INECN	Institut National pour	Mr Alain Charles Kakunze	akakunze@yahoo.fr

	l'Environnement et la Conservation de la Nature		
UNIBU	Université de Burundi	Dr Joël Ndayishimiye	ndayishimiye_joel@yahoo.fr
CRSN / Lwiro	Centre de Recherche en Sciences Naturelles	Mr Robert Kizungu	robertkizungu@yahoo.fr
		Mr Mwanga Milinganyo	jacques.mwanga@gmail.com
CGIS	Centre for Geographic Informations Systems	Mr Jean Damascene Mazimpaka	jdmazimpaka@nur.ac.rw
RDB	Rwanda Development Board	Mr Kwizera Janvier	kjanvier2010@gmail.com
ACNR	Association des Ecologistes Nationale du Rwanda	Mr Prosper Karame	karapros@gmail.com

Annex II. Workshop Program

Day 1: Introducing the Albertine Rift Biodiversity Portal		
Time	**Activity**	**Responsible**
08:30 - 08:40	Welcome from ARCOS and Introductions	Dr Sam Kanyamibwa (Executive Secretary ARCOS)
08:40 - 09:00	Opening Remarks	Representative of GoR from MINEDUC
09:00 - 09:15	Summary and Planning for the course	Faustin Gashakamba (ARCOS)
09:15 – 10:00	Introducing the Albertine Rift Biodiversity Portal	Zac (ARCOS)
10:00 – 10:30	Discussions on the features and functionalities of the Data Portal Prototype (Noting Comments for validation)	Faustin Gashakamba (ARCOS)
10:30 – 11:00	**Coffee/tea break**	Claudien (ARCOS)
11:00 – 12:00	Practical Exercise: How to use the Data Portal Prototype (Creating an Account, Uploading Data, Searches, Downloads, etc)	Zac Maritim & Faustin Gashakamba (ARCOS)
12:00 -13:00	Overview of data held in data centres	Data centres representatives
13:00 – 14:00	**Lunch break**	Claudien (ARCOS)
14:00 – 16:00	Individual Project Exercise	Working Groups
16:00 – 16:30	Coffee/Tea break	Claudien (ARCOS)

16:30 – 17:30	Questions and answers	All Participants
17:30	Close of the session	All

Day 2: Spatial Data Mobilisation, Remote Sensing and Landuse Planning		
08:30 – 09:30	Data collection: GPS and georeferencing	Apollinaire (ARCOS)
09:30 – 10:30	Image classification using ArcGIS 10	Apollinaire (ARCOS)
10:30 – 11:00	**Coffee/tea break**	**Claudien**
11:00 – 12:00	Practical Exercise: Classification using Arc GIS 10 (continued).	Apollinaire & Zac Maritim (ARCOS)
12:00 – 13:00	Introduction to Remote sensing and Digital Image Interpretation. Theoretical Concepts of Image Acquisition, Enhancement and Processing	Zac Maritim (ARCOS)
13:00 – 14:00	**Lunch break**	**Claudien**
14:00- 15:00	Practical Exercise: Image Enhancement and Image Processing Techniques (Geometric and Radiometric Correction, Mosaicking, Pan-sharpening, Subsetting and Layer-stacking)	Apollinaire (ARCOS)
15:00 – 16:00	Per-pixel Classification (Unsupervised and Supervised Classification) using ERDAS Imagine.	Zac Maritim (ARCOS)
16:00 – 16:15	**Coffee/tea break**	**Claudien**
16:15- 17:15	Extraction of Classified Raster images to vector polygons for incorporation to ArcGIS.	Apollinaire William (ARCOS)
17:15- 17:45	Individual Project Exercise: Exposition of Maps produced by analytical properties of GIS.	All
17:45- 18:00	Close day	

Day 3: Data Management and Publishing		
8:30 – 09:30	Practical Session: Database management Systems (DBMS). Using ArcGIS to create a Geodatabase for exploitation in the Spatial Component of the Portal	Zac Maritim (ARCOS)
09:30 – 10:30	Georeferencing of Attribute Data that has no coordinates (Using Place Names, admin boundaries, etc in ArcGIS)	Zac Maritim (ARCOS)
10:30 – 11:00	**Coffee/tea break**	**Claudien**
11:00 - 12:00	Spatial-linking of GIS Database to Portal	Zac Maritim (ARCOS)
12:00 – 13:00	**Lunch break**	**Claudien**
13:00 – 14:00	Biodiversity data quality and cleaning: Names and Taxonomy and GIS Data Cleaning and Quality Control + data Structuring for GIS and for the Portal (Data Discovery)	Zac Maritim & Faustin (ARCOS)
14:00 – 15:00	The GBIF data architecture and The GBIF Integrated Publishing Toolkit (IPT) Installation	Guest speaker - TanBIF
15:00 – 15:30	**Coffee/tea break**	**Claudien**
15:30 – 16:30	Practical demonstration: connection of the dataset produced by participants in the practical session to the GBIF Network using the GBIF Integrated Publishing Toolkit and to the ARCOS Data Portal	Zac Maritim (ARCOS)
16:30- 17:00	Individual Project Exercise: Final results and Follow up	All Working Groups
17:00- 17:15	Course Evaluation	All
17:15- 17:30	Closing Ceremony and Certificate Issuance	Sam Kanyamibwa (ARCOS)

Annex III. Group photo

ARCOS Network

UK Registered Office: Wellbrook Court, Birdlife International, Girton Road, Cambridge CB3 0NA, UK.
Regional Office, 1329, Nsambya, Kabalagala, P.O Box 9146, Kampala, Uganda,
Tel/Fax: +256-414- 530700, Email:info@arcosnetwork.org, Website: www.arcosnetwork.org