# Establishment of Ki-Geo and Kiribati Health Portal

A Health Statistics Data Repository Within A Geospatial Context for the Government of Republic of Kiribati


**9th April - 16th April 2015**
**Tarawa, Kiribati**


| **Sachindra Singh** | **Wolf Forstreuter** |
|---|---|
| Senior Geospatial Systems Architect | GIS/RS Team Leader |
| SPC Geoscience Division | SPC Geoscience Division |


The assignment entailed deployment of Ki-Geo and Kiribati Health Portal, along with provided relevant training to end-users and IT personnel within Kiribati Ministry of Health and Medical Services, Ministry of Lands and Ministry of Agriculture. In addition to the training workshop, a Kiribati GIS/RS User Meeting was convened. The data structure, inventory and backup is discussed and detailed in this report. An additional instruction paper will be delivered to Kiribati explaining the data link between SPC-GSD and Kiribati Lands Department.

# Table of Contents

# 1 Background

Kiribati Health Portal (KHP) is a secure online database system which enables Ministry of Health and Kiribati personnel to populate, collate and collaborate on outbreaks, household and utilities surveys, in addition to other related datasets, within a geospatial context.

KHP is designed to be integrated with a subset of PacGeo, the Pacific's Spatial Data Infrastructure (SDI) platform, called Ki-Geo, holding required structured geospatial datasets for Kiribati, critical to the operation of KPH.

# 2 GIS/RS User Meeting

The Kiribati GIS/RS User Meeting was held on Friday 10th April, and background information for Ki-Geo and the KHP was presented. Recent developments in GIS/RS within the region were presented and discussed.

| First Name | Family Name | Organisation | E Mail |
|---|---|---|---|
| Hiroki | Inone | MHMS | inopn.g@gmail.com |
| Ioane | Titaake | MFMRD | ioanetitaake@gmail.com |
| Itaaka | Tiaon | MHMS | itbwenawa@gmail.com |
| Kaiea | Ribanataake Awira | MFMRD | kaieaawira@gmail.com |
| Kirabwa | Redfern | LANDS LMD | kirabantetaake@gmail.com |
| Sachindra | Singh | SPC | sachindras@spc.int |
| Tamueru | Taverio | PUB | Ttamuerutav@gmail.com |
| Teanibuaka | Tabunga | MHMS | teanibuakatabunga@gmail.com |
| Tewaea | Keariki | LANDS | tkeariki@gmail.com |
| Wolf | Forstreuter | SPC | wforstreuter@yahoo.co.uk |

***Table 01:*** *Participants of GIS&RS User Meeting 10th April 2015*

# 3 Kiribati Health Portal

The health portal is an extensible enterprise-level online database system enabling Kiribati Ministry of Health (MOH) personnel to collate and collaborate on health-related datasets, including but not limited to:

- Outbreaks Data Collection
- Health Census
- Household Surveys
- Marine Water Sampling
- Public Water Sampling
- Well Water Sampling
- MOH Clinic and Personnel Management
- Documents and Reports Repository
- Integration with Geospatial Maps
- Administrative, Infrastructure and other Baseline Layers (eg: Villages, Buildings, Population)

KPH was deployed on the Local Area Network with MOH, therefore only accessible MOH personnel granted with appropriate user access. The system is designed around comprehensive user roles and fine-grained security, given that the database will store confidential information not oriented for general public consumption.

The platform has an extensible reporting system, which will enable KI-MOH personnel to query the system and generate reports required in an intuitive manner, both within Ki-Geo, and in non-spatial manner.

In addition to spatial maps, the portal will also be pre-populated with base-line read-only population and housing data from Kiribati national statistics datasets.

The system has simple update mechanisms, allowing IT personnel within the ministry to roll out updates remotely provided by SPC Geoscience Division.

# 4 Ki-Geo Data System

KHP implementation entails deployment, data population and provisioning relevant capacity building and support for an integrated enterprise-level Geospatial Data Services platform, that will enable the KI-MOH centrally collate, manage and expose it's geospatial and related health statistics data holdings.


***Figure 01:*** *The server installed at Ministry of Health*

SPC Geoscience Division installed the Ki-Geo spatial data platform on a server-class stand-alone machine and deployed it with collaboration with KI Ministry of Health and KI ICT personnels. The Ki-Geo platform will enable Ministry of Health to easily import and manage their geospatial data holding, and securely expose relevant datasets to other government ministries.

The platform enables user-friendly spatial data discovery via interactive and informative web mapping, along with exposing it via OGC-complaint services to desktop GIS clients on the government's local area network.

The platform has had batch data population according to the data structure and inventory completed for the Ministry of Lands earlier in 2014 by SPC Geoscience Division's GIS&RS Section.

Tight-integration with KHP enables the Ministry to spatially map out different and dynamic variables, such diarrhoea outbreaks information contrasted with well water attributes, or toilet types etc.

# 5 Data Structure and Backup

Data backup only works sustainably if there is a clear data structure with, as far as possible,

unique directory names. Having good backup facilities with safety copy in the cloud the typical situation hard disk crashed is not the main reason to loose data. Now old data overwrites new data is frequent occasion where data is lost even with best backup facilities.

## 5.1 Data Backup Situation at Lands, Agriculture and Health

During the workshop the participants reflected on the data backup situation at their working environment.

In Kiribati most data is stored at the ***Lands Department***. Data backup is performed on DVDs, external hard drives and on a server. The data backup is conducted every month. There is one person in charge: Tiaotin Enari. He participated in the workshop. There were already situations where the main server crashed and data could be recovered from external hard drives. However, this is a very time consuming procedure as long as data is not structured.

The second highest data volume is stored at the Kiribati ***Agriculture Department***. The department stores backup data on external hard drive and synchronises this with the Lands Department server. There is also a backup at SPC GIS&RS unit. The backup is not performed in regular intervals, it is carried out after a larger amount of data is created. Mostly Maio Tebania performs the backup but there is nobody with assigned responsibility. Data got lost four times when the computer corrupted. So far most data could be restored from Lands Department and SPC-GSD.

At the Kiribati ***Ministry of Health*** a server will be the backup media at a later stage. There is no regular data backup yet. The person responsible is a Japanese volunteer, Hiroki Inone. There was data loss due to computer crash without any recovery.

## 5.2 Backup Structure at SPC-GSD

At SPC-GSD every country which keeps a backup copy at SPC has an own external hard drive where data is stored in a structured form described below. Data is copied to this hard drive with Synchronise software which overwrites old files with new versions on the hard drive. With the current set-up the software does not delete files or directories on the external drive. If there a directory renamed the old and the new directory is kept at the backup media, which allows files stored twice. It is therefore essential to keep the exactly same file structure.

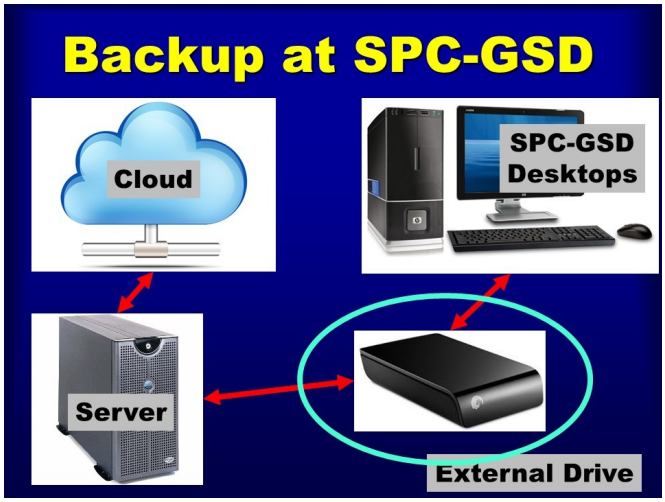



***Figure 02:*** *Data backup at SPC-GSD. The external hard drives keeping the master copies for the countries. From there servers and cloud backup get synchronised.*

From the external hard drive data is synchronised with a server installed at SPC-EDD. From there another backup copy will be placed in the cloud.

A copy of the external hard drive moves also between Kiribati and SPC-GSD to keep data synchronised. This is working since 2009.

## 5.3 Backup Structure in Kiribati and between Suva and Tarawa

Currently most spatial data in Kiribati is held by Lands Department and by Agriculture Department where Agriculture Department synchronises with Lands Department to have a safety backup on the Lands Department server. This synchronisation is performed by physically taken an external hard drive from Buota to Bairiki. During a GIS&RS User meeting in May 2014 it was discussed and agreed that also other Departments backup their spatial data on the Lands Department server.

The Ministry of health has now an own server but will also synchronise all none confidential data with the Lands Department.

Lands Department synchronises data with SPC-GSD. This synchronisation works via external drive which is transported with diplomatic courier between Suva and Tarawa. The synchronisation is important as in SPC-GSD and in Tarawa files are updated. For example Agriculture and SPC-GSD working on land cover mapping and create new layers. The synchronisation in both directions works manually so far. This will be more efficient through the regular data inventories. The data inventories are performed through an open source software and list all file names and file attributes. The software stores them in a small file which can be transferred as e-mail attachment. This file allows to identify: (i) files which are physically missing in Kiribati where a copy has to be sent to Tarawa, (ii) which files have been updated in either Kiribati or at SPC-GSD, (iii) files which are new in either Kiribati or SPC-GSD; mostly SPC-GSD has new files related to new image data, (iv) files which have the same file name and the same file size but are stored in different directories. These files can be easily identified, but then somebody with local knowledge has to make the decision which file to delete. It is extremely important to keep only one file as most data gets lost when an old version overwrites a new version. A file of similar content and same file name in a different directory can be mistakenly taken as latest version. If somebody works with this file it then becomes the physically the latest version and can overwrite the file which keeps the most correct data.

| Atoll_ID | Atoll Name |
|---|---|
| 000 | 000_national |
| ABA | Abaiang |
| ABE | Abemama |
| ARA | Aranuka |
| ARO | Arorae |
| BAN | Banaba |
| BER | Beru |
| BIE | Biernie |
| BUT | Butaritari |
| CAR | Caroline |
| CHR | Christmas |
| END | Enderbury |
| FAN | Fanning |
| FLI | Flint |
| KAN | Kanton |
| KUR | Kuria |
| MAI | Maiana |
| MAK | Makin |
| MAL | Malden |
| MAN | Manra |
| MAR | Marakei |
| MCK | McKean |
| NI2 | Nikumaroro |
| NIK | Nikunau |
| NON | Nonouti |
| ONO | Onotoa |
| ORO | Orona |
| RAW | Rawaki |
| STA | Starbuck |
| TA2 | Tabueran |
| TAB | Tabiteua |
| TAM | Tamana |
| TAR | Tarawa |
| TER | Teraina |
| VOS | Vostok |
| WAS | Washington |

***Table 02:** Atoll abbreviations*

## 5.4 Data Structure on External Hard Drive

As mentioned above, the external hard drive is the master structure and master data storage. From there it synchronises with the server in one way direction. New files or files which have been changed will update the server, however, the server data and structure do not update the external drive. Any change of structure where directories have been deleted has to be manually synchronised on the server. This also applies to files which have been deleted on purpose on the external drive. Therefore the directory structure is absolutely essential.

The proposed directory structure has three levels. To avoid any overwriting of data the following roles have to be followed:

1. Location before data type
2. Only three levels of directories no further sub directories.
3. Every directory contains an indicator of defined length reflecting the directory above, which makes the directory name unique.

### 5.4.1 Level 1 Directories, Atoll Name

The rule of location and then the data type is adapted from regional libraries where the books are stored in this order. The first level is therefore indicated by the island name. Table 02 has been created by Lands Department showing the island names and the atoll abbreviation. Normally 3 digits are sufficient to create a unique island ID. The exceptions are Nikumaroro which conflicts with Nikunau and Tabueran which conflicts with Tabiteua. For Tabueran TA2 is used as abbreviation and Nikumaroro got the abbreviation Ni2.

Data of national dimension such as EEZ boundary or maps showing island groups are stored in the corresponding sub-directory of a directory named 000_national.

The data is not stored in the first or second directory level but in the third level.

### 5.4.2 Level 2 Directories, Data Type

The second directory level shows the data types which were distinguished so far. This table will added over time. Currently the data type table is overlapping as topographic data (TOP) can be stored as vector data (VEC). There was not sufficient time yet to sort all data. This will be carried out over time. It has to be performed under the guidance of Lands Department as central institution looking after spatial data. It was mentioned in the last chapter that changing a directory name has major implications as every institution storing data has to follow (i) the directory renaming (ii) the new directory structure and (iii) the type of files stored in the directories. If these three rules are not followed the same file can be stored in different directories during the synchronisation process.

| Type_ID | Data Type |
|---|---|
| ADM | Administration information |
| AER | Aerial photographs |
| BTH | Bathymetry data |
| CAD | Cadastral data |
| CHD | Change detection data |
| DTM | Digital terrain models, digital surface models |
| GRD | Grids |
| REP | Reports |
| SAT | Satellite image data |
| STA | Statistic data |
| TOP | Topographic data |
| VEC | Vector data |
| WAT | Water survey data |

***Table 03:*** *The table shows the data type and the corresponding abbreviation*

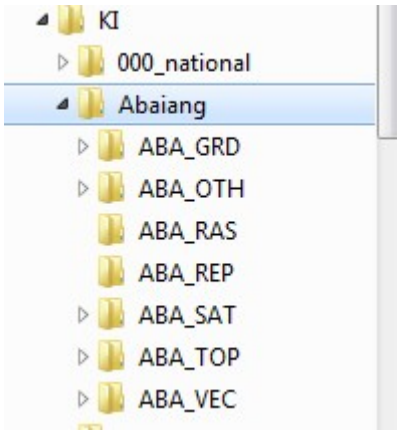



***Figure 03:*** *Second directory level structure for Abaiang*

As mentioned above, the second directory level also contains the abbreviation of the island name as part of the directory name. Through this the directory name becomes unique and it is fast to list files with normal operating instructions. The

directory names have the fixed structure of 3 digits indicating the island followed by an under score plus three digits indicating the data type.

Like the first directory level, the second level of directory does not contain data directly as data is stored in the third level of the directory structure.

### 5.4.3 Level 3 Directories, Sub Data Type

The sub data type is detailing the data type. If the data type indicates satellite image data (SAT) the sub data type explaining WV2 for image data from WorldView-2 satellite and COR for geometrically corrected data. On level three several abbreviations can indicate the type of data.

| Type ID | Data Type |
|---|---|
| ADM | Administration information |
| AER | Aerial photographs |
| BL | Base Line |
| BLD | Buildings |
| BP | Base Points |
| BTH | Bathymetry data |
| BW | Black & white aerial photograph |
| CAD | Cadastral data |
| COR | Geometrically corrected |
| DB | Database |
| DTM | Digital terrain models |
| GEO | GeoEye satellite image data |
| GRD | Grids |
| IKO | IKONOS satellite image data |
| INF | Infrastructure information |
| LS7 | Landsat 7 image data |
| LS8 | Landsat 8 image data |
| MED | Metadata |
| POI | Point data |
| POL | Polygons |
| QB1 | QuickBird satellite image data |
| RAS | Raster data |
| RD | Road layer |
| REP | Reports |
| RFL | Reef flat |
| SAT | Satellite image data |
| TOP | Topographic data |
| VEC | Vector data |
| VEG | Vegetation layers |
| WV2 | WorldView-2 image data |
| WV3 | WoldView-3 image data |

***Table 04:*** *Currently used abbreviations to describe the sub data type as directory name in the third directory level*

It is essential to avoid further sub directories as the data inventories are difficult with additional sub directories. Instead of sub directories the type of data can be described in detail in the third directory level. This provides a better overview and avoids that the same file is stored in different directories. Table 04 shows the current abbreviations which can be used parallel in the directory name.

The indication of data type is repeating as it is necessary to (i) repeat the name of the second directory level as this creates unique names and (ii) have a third level directory for data to avoid file storage in the second directory level. For example general vector data for Butaritari is stored in the directory:

**BUT_VEC_VEC**

BUT_VEC indicates the second directory name and the second abbreviation for vector data VEC indicates general vector data. In this way a unique name for the directory is created.

It has to be clearly stated again, if different vector data from a summary directory such as BUT_VEC_VEC is distributed to several more detailed named directories and the summary directory is then deleted because empty, the summary directory has to be manually deleted in the server and higher backup media as well. Otherwise the data is still kept in the summary directory and copied again with the new more detailed directory names. There is a possibility of a two way synchronisation, however, this requires very experienced operators. It is currently safer to keep one way synchronisation and manually deletion of empty directories.

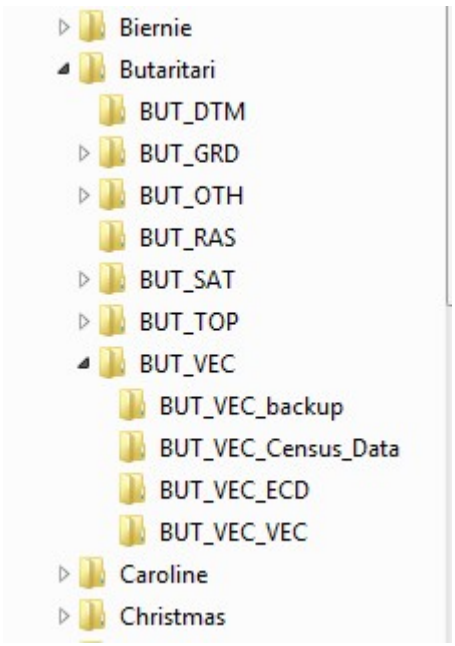



***Figure 04:*** *Third level of directories in which actually data is stored.*

# 6 Workshop

The workshop was conducted on 14th and 15th of April where Part A was held during the first and Part B during the second day. TSKL provided the facility.

## 6.1 Purpose Workshop Part A

The purpose of the workshop part A was that the participants understood the spatial data structure, data inventory and monitoring and data backup. Do be able to do this the participants were introduced several open source based tools and database handling.

## 6.2 Material Handed out

A memory stick was handed out from which the participants copied to their computers. The memory stick had following content:

1. Software (open source):
   - Synchronise
   - Ant Renamer
   - FileList
   - Agent RanSack
2. Exercises
3. Power point presentations

***Figure 05:*** *Training data backup, Tiaotin Enari (Lands) and Maio Tebania (Agriculture) with more experience in data handling assisted.*

## 6.3 Training Elements Workshop Part A

The participants were introduced to different possibilities of ***data loss*** and reflected about their own experience loosing data[1]. Methods

1 See chapter “Data Backup Situation in Lands, Agriculture and Health”

to avoid data loss were discussed, where the main element is a clear data structure with unique directory names and limited number of subdirectories.

File ***renaming software*** “Ant Renamer” was introduced and trained to create unique file names allowing to simplify the directory structure. An exercise trained to do string replacement of the file name and another exercise teached a string insertion. Both together created unique file names.

A further exercise trained the ***directory structure*** by actually restructuring a directory. The participants learned to follow the principal (i) location, (ii) data type and (iii) sub-data type. The data type abbreviations were discussed and a corresponding table was added. The same was conducted for the sub-data types. The workshop also had the purpose to create a table with unique abbreviations for all atolls, however, this was not necessary as this exists already at Lands Department[2].

The open source software product ***“Agent Ransack”*** allows to search for files and directories in an efficient way. It was trained to list files of similar names located in different directories and it was trained to locate files by searching for a sub-string of text as part of the file content. Other functions were demonstrated.

The afternoon of the day was used to train ***data inventory*** using the open source product ***“FileList”***. It was demonstrated to start the software from the Command Prompt, however, the actual training concentrated on creating an own batch file which allows to set all parameters in a more efficient way. The software lists the names of all files and directories available on a hard drive or other data media. With the names the software writes other file attributes into a text file. Following attributes are listed:

1. File name
2. File size
3. Last change
4. Last access
5. Creation date
6. Extension
7. Path to the file

As a next step it was trained to ***import the text file*** created by FileList software into an Access database and to utilise the database functions. Some of these functions are coded and only the application was demonstrated. However, the database will be handed out to Ministry of Health, Agriculture and to Lands Department. It is possible to:

- Run an automatic import of the text file created by the file list software
- List files of selected data type and/or of selected atoll
- Identify files of same name and same size stored in different directories
- Create a batch file which copies selected files from an external hard drive to another hard drive without manually moving in the backup media.

2 Current abbreviation tables for atoll names, data type and data sub-type see chapter “Data Structure”

## 6.4 Purpose Workshop Part B

Purpose of the workshop was to train MOH end-users in the data-entry, querying and general usage of K-Geo and KHP. The workshop also allowed participants to submit feedback on individual data entities, and data forms were fine-tuned during the workshop.

The participants were handed audio/visual materials and other documentation for the usage of the platforms.

## 6.5 Training Elements Workshop Part B

Ministry of Health's designated GIS and information management personnel were be trained on KPH and Ki-Geo data population and maintenance, including but not limited to:

1. Populating Survey Data
2. Deriving Maps
3. Style Maps and Layers
4. Building Metadata
5. Undertaking Spatial Queries
6. Defining Permissions and Security Management
7. Utilizing OGC WMS and WFS Services.
8. Sharing and Publishing Maps and Reports
9. Export and contextualizing data from KPH
10. Creating and modifying attributes within Ki-Geo.

***Figure 06:*** *Systems training at Tarawa Hospital. L-R: Sachindra Singh, Itaaka Tiaon, Hiroki Inoue*

## 6.6 List of Workshop Participants

| First Name | Family Name | Organisation | E_Mail | M/F |
|---|---|---|---|---|
| Maio | Tebania | AGRICULTURE KI | mtebania@msn.com | M |
| Tiaotin | Enari | LANDS | tiaotinenari@gmail.com | M |
| Kiing | Rotitaake | MHMS | krotitaake@gmail.com | F |
| Kiratevenga | Baina | MHMS | kmoerenga@gmail.com | F |
| Totiua | Puta | MHMS | totiuaputa@gmail.com | F |
| Rianako | Imota | MHMS | jorita211110@gmail.com | F |
| Maryanne | Utiera | HIU MHMS | marymsanne@gmail.com | F |
| Tebikau | Tibwe | EH MHMS | tnoran@gmail.com | M |
| Bungiia | Kaitaake | EH MHMS | bkirata87@gmail.com | F |
| Tenamau | Iotua | MHMS | titamaba@gmail.com | F |
| Itabera | Teekea | MHMS | tabeuia14@gmail.com | F |
| Maria | Tanro | MHMS | meerethomas@gmail.com | F |
| Kabwaki | Torere | MHMS | kabwakit@gmail.com | M |
| Ruti | Mikaio | MHMS | rutimikaio@qmail.com | F |
| Tirite | Irooti | MHMS | georgina.bwere@gmail.com | F |
| Teanibualea | Tabunga | HIS MHMS | teanibualeatabunga@gmail.com | M |

***Table 05:*** *List of workshop participants*

## 6.7 Follow-up Actions

KHP will be updated and delivered remotely to MHMS IT personnel by SPC Geoscience Division based on end-user feedback from MOH. Remote systems support, additional documentation and user assistance will also be provided via email.

Data storage structure has to be further discussed with Lands Department and instruction paper will be delivered. The file name database will be advance and updated in all three government institutions. A data synchronisation routine will be evolved.

## 6.8 Summary

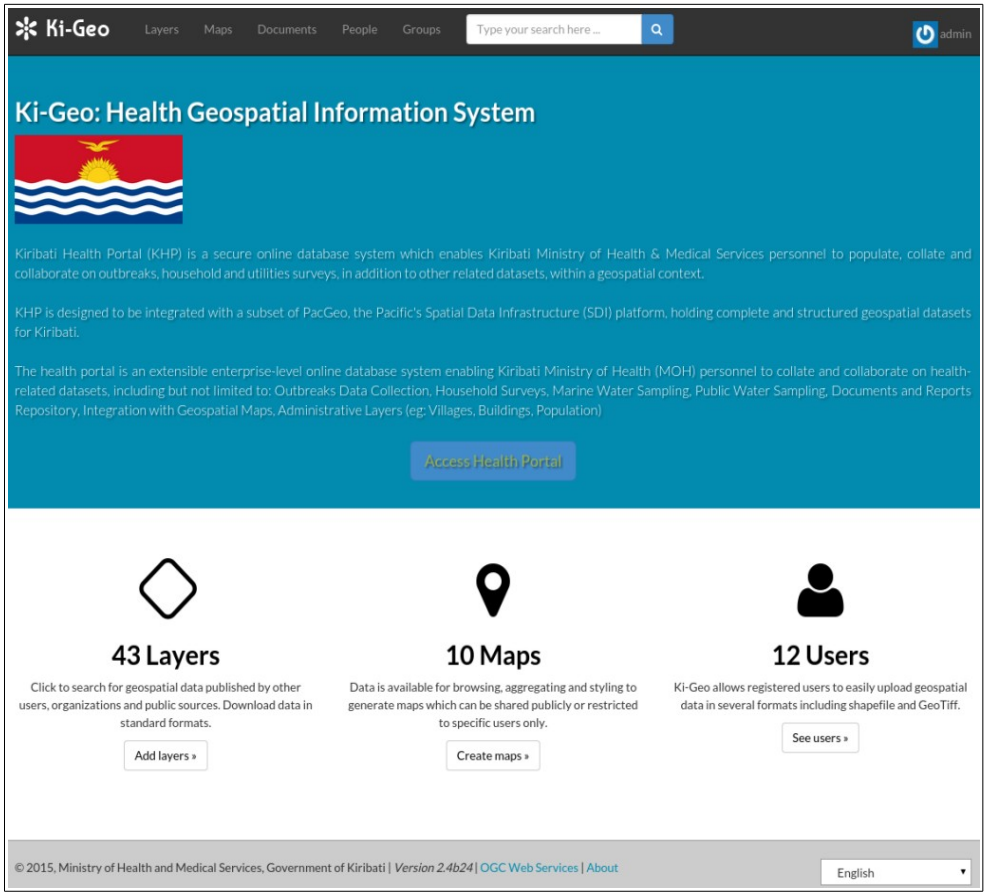
- Provisioning of appropriate server-class computational resources.
- Setup and configuration of Linux Operating System alongside Ki-Geo and related systems.
- Deployment of the KPH server within KI government's local area network.
- Creation of file and related web services on the local area network.
- Training MOH IT Section's designated personnel on maintenance, updating and disaster recovery mechanisms of the installed services.
- Training Ministry of KI-MOH GIS personnel on effectively utilising the KPH and PacGeo platform.
- Training Ministry of KI-MOH GIS personnel on updating the inventory and structure of data holdings.

# 7 References

KPH Prototype Implementation

http://ict.sopac.org/KiribatiHealth


PacGeo: open access geospatial data repository for the Pacific Region providing premier geophysical, geodetic, and marine spatial data sets.



http://www.pacgeo.org/beta