GIS and Spatial Data Systems Guidelines

The purpose and scope of GIS and spatial data guidelines are to establish a common framework that enhances the quality, consistency, and efficient use of GIS Technology and data in Government of Rwanda Institutions. This will lead to more effective decision-making, better planning, and improved collaboration and data sharing among stakeholders across various sectors and disciplines. The guidelines cover GIS Data Collection, storage integration, data sharing, quality and data analysis.

- Scope and Applicability
- Data Storage and Management
- Data Integration
- Data Sharing
- Data Quality
- Data Analysis and Visualization
    - Data Analysis
    - Data Visualization

Scope and Applicability

GIS and spatial data systems play a role in capturing, analyzing, and visualizing geographic information, enabling informed decision-making across various industries. In Rwanda, Government institutions are using GIS in many ways such as National land use planning, Infrastructure development, Agriculture, Environmental conservation, Disaster management and Urban planning.
The purpose and scope of GIS and spatial data guidelines are to establish a common framework that enhances the quality, consistency, and efficient use of GIS Technology and data in Government of Rwanda Institutions. This will lead to more effective decision-making, better planning, and improved collaboration and data sharing among stakeholders across various sectors and disciplines. The guidelines cover GIS Data Collection, storage integration, data sharing, quality and data analysis.

Data Storage and Management

Data storage guidelines address how spatial data should be organized, stored, and maintained in database systems or files. Consistent storage standards ensure data integrity, accessibility, and ease of retrieval for users and applications. Key guidelines for data storage and management are listed below:

- Use a geospatial database management system (DBMS) like PostgreSQL with the PostGIS extension, which offers spatial data support and advanced geospatial functions
- Organize spatial data in a structured manner, such as by geographic regions and data types
- Implement a version control system (such as Git) to track changes made to spatial data and associated metadata. Encourage collaboration among team members when making changes to the data and establish a review process before finalizing new versions
- Establish clear data retention policies based on legal requirements, organizational needs and data significance. Determine the specific retention periods for different types of spatial data based on their importance and relevance
- Develop clear data governance policies that define data ownership, access controls, update procedures, and data lifecycle management
- Establish a backup system to safeguard against accidental data loss or corruption

Data Integration

Data integration processes for GIS spatial data include data conversion, data merging and data format standardization. These are essential for creating a consistent and unified spatial dataset. Key guidelines include:

- Adopt an open data format like Shapefile (SHP), GeoJSON, or Keyhole Markup Language (KML). Standardizing formats facilitates data exchange between different systems and minimizes data conversion issues
- Data conversion may be required to transform one spatial data format to another to achieve compatibility. Common spatial data formats include Shapefile, GeoJSON, KML, GeoTIFF, and others. Use data transformation and integration tools, like GDAL (Geospatial Data Abstraction Library), to convert data between different formats and systems efficiently
- Data formats should be standardised to ensure consistency in how geographical information is represented and stored across the integrated dataset. Examples includes establishing uniform naming conventions, units of measurement, and attribute values
- Follow recognized spatial data standards like those set by Open Geospatial Consortium (OGC)
- Adopt metadata standards such as ISO 19115 to document data content, quality and other relevant information
- Coordinate reference systems define how spatial data is positioned on the Earth's surface. Ensure that all data layers are correctly aligned with a consistent CRS to avoid spatial discrepancies

Data Sharing

Data sharing guidelines aim to support data sharing among institutions through adoption of common standards and tools. Recommendations include:

- Establish data sharing protocols and agreements when collaborating with other organizations. Clearly define data formats and quality requirements to ensure compatibility
- Spatial Data Transfer Standard (SDTS) is a standard for transferring spatial data between different computer systems. It defines a structure and format for various types of spatial data, such as vector, raster, and attribute data. By using SDTS, government institutions can ensure that their spatial data can be exchanged with other agencies seamlessly.
- Web Map Service (WMS) is a standard protocol developed by the Open Geospatial Consortium (OGC) for serving georeferenced map images over the Internet. It allows government institutions to share spatial data as map images that can be visualized in various GIS software and web mapping applications.
- Web Feature Service (WFS) also an OGC standard, WFS enables the sharing of geographic features and their attributes over the internet. It allows users to access and retrieve vector-based geographic data directly rather than just images, facilitating data analysis and manipulation
- Spatial data can be large, so using efficient compression techniques and optimized transfer protocols can improve data exchange performance
- Adhere to widely accepted standards and protocols such as those defined by the OGC to help ensure compatibility between systems and data formats

Data Quality

The guidelines outline procedures for data validation, error checking, and data accuracy assessments. Data of high quality is essential for making reliable decisions and producing accurate results. Below are guidelines on data quality.

- Implement data validation processes to identify and correct errors, inconsistencies, and outliers in the spatial data. This involves cross-checking data against predefined rules, range checks, and logical validations to maintain data integrity
- Thoroughly document metadata, providing essential information about the spatial data, including its source, collection methodology, accuracy, and update frequency
- Follow established spatial data standards and protocols for data collection, storage, and dissemination. These standards ensure compatibility and interoperability across different systems and applications, enabling data sharing and integration

Data Analysis and Visualization

These guidelines aim to help institutions to effectively communicate information through GIS data analysis and visualizations. Key recommendations include:

Data Analysis

i. Geoprocessing:

- Verify the spatial relationships to avoid overlaps, gaps, or invalid geometries.
- Ensure all spatial datasets have the same coordinate system or project them to a common one for accurate analysis.
- Convert data between different formats (e.g., vector to raster) based on analysis requirements.

ii. Spatial Analysis:

- Utilize buffering and proximity analysis to study the relationships between features and their surroundings.
- Combine spatial datasets and aggregate data to analyze relationships and summarize information.
- Use appropriate spatial interpolation methods to estimate values at unsampled locations.
- Apply network analysis to model movement and connectivity within spatial networks.

iii. Statistical Modelling:

- Split the dataset into training and testing subsets for model training and evaluation.
- Clearly define the dependent variable and select relevant independent variables for the model.
- Choose appropriate statistical models based on the nature of the data. (e.g., linear regression, logistic regression).
- Assess model performance using appropriate metrics (e.g., accuracy, RMSE, R-squared) and cross-validation techniques.
- Interpret the results of the statistical model in the context of the research objectives and data patterns.
- Validate the model on unseen data to ensure its generalizability and avoid overfitting.

Data Visualization

i. Map Design Principles:

- Simplicity - Maps should be visually clean, focusing on the most important information while removing unnecessary elements.
- Legibility - Ensure that the map's text, labels, and symbols are easy to read and understand, even when zoomed in or out.
- Balance - Maintain a visual balance on the map by evenly distributing elements, avoiding overcrowding in any particular area.
- Contrast - Use contrasting colors and symbols to make the map visually appealing and enhance the distinction between different features.
- Hierarchy - Arrange map elements in a clear hierarchy, with more important or prominent features receiving greater visual emphasis.
- Consistency - Maintain a consistent design style, color scheme, and symbology throughout the map to create a coherent visual experience.

ii. Use of Symbology:

- Color - Use color wisely to convey meaning, such as warm colors for higher values and cool colors for lower values.
- Size - Use different symbol sizes to represent varying magnitudes or quantities, ensuring that the size variation is easily distinguishable.
- Shape—Differentiate between various features using distinct symbol shapes, especially when dealing with multiple datasets on the same map.
- Iconography—Utilize appropriate icons to represent specific features or categories, making the map more visually informative.
- Labels—Provide clear and concise labels for important features or regions to enhance map readability
- Transparency—Use transparency for overlapping elements to prevent visual clutter and improve comprehension.