Towards maximizing geospatial data usage in north eastern India using open-source scalable user-centric applications

Puyam S Singh [1], Dibyajyoti Chutia [1], Nilay Nishant [1], Avinash Chouhan[1],Victor Saikhom [1], Shiv Prasad Aggarwal[1], Rahul Kumar[1]

[1] North Eastern Space Applications Centre,
Government of India, Department of Space – ss.puyam@nesac.gov.in

Keywords: Open Source, Maximize data Usage, Interoperability, Governance Applications, Capacity Building

Abstract

The north eastern region of India is characterized by limited physical connectivity and challenging terrain, that requires innovative approaches for effective developmental planning and monitoring using geospatial data and user-centric applications. This paper focuses on leveraging space-based data and applications to address these challenges and enhance decision-making processes in governance. A unified framework for interoperable spatial data sharing has been developed, ensuring that data is easily discoverable, searchable, and consumable through metadata standards and OGC interoperable data services. A dedicated spatial data repository has been established, offering real-time access to over 1400 geospatial datasets and facilitating better decision-making and geo-analytics among diverse user groups. The repository operates on an open-source platform with integrated map viewers and geo-analytics tools, allowing for interactive exploration and visualization of geospatial data. Additionally, 35 user-centric geospatial applications have been operationalized, all developed using open-source geospatial tools and standards. Capacity building programs have been conducted to increase awareness and maximize the utilization of geospatial data services and applications across the region. The platform has boosted utilization of geospatial data services and applications in the region for overall planning and developmental activities for smart governance and societal benefits.

1.Introduction

The North Eastern Region of India, known for its picturesque landscapes and rich cultural heritage, faces significant developmental challenges due to its limited physical connectivity and challenging terrain. These geographical constraints necessitate innovative approaches for effective developmental planning and monitoring. In this context, the use of geospatial data and user-centric applications emerges as a vital solution. This paper focuses on the strategic leverage of space-based data and applications to address these challenges, aiming to enhance the decision-making processes in governance within the region. To facilitate this, a unified framework known as the North Eastern Spatial Data Repository (NeSDR) has been developed. The NeSDR framework is designed to ensure that spatial data is easily discoverable, searchable, and consumable. This is achieved through the implementation of robust metadata standards and Open Geospatial Consortium (OGC) interoperable data services. The establishment of a dedicated spatial data repository stands at the core of this framework, providing real-time access to over 1400 geospatial datasets. This repository plays a crucial role in improving decision-making and geo-analytics capabilities among diverse user groups within the region. The repository operates on an open-source platform, which is equipped with integrated map viewers and geo-analytics tools. These tools facilitate the interactive exploration and visualization of geospatial data, making the data more accessible and usable for various stakeholders. In addition to the repository, a suite of 35 user-centric geospatial applications has been developed and operationalized. These applications are all built using open-source geospatial tools and standards, ensuring broad accessibility and flexibility.

Recognizing the importance of capacity building, numerous programs have been conducted to raise awareness and maximize the utilization of geospatial data services and applications. These programs aim to empower local governments, organizations, and communities to leverage these tools effectively for planning and developmental activities.

The implementation of the NeSDR platform and its associated applications has significantly boosted the utilization of geospatial data services in the North Eastern Region. This has had an impact on the overall planning and developmental activities, enhancing smart governance and delivering tangible societal benefits. The NeSDR database and its applications have significantly benefited numerous departments and academic institutions in the region. To further raise awareness and highlight the utility of these applications, various training programs and workshops have been organized. Through this paper, we explore the development, deployment, and impact of the NeSDR framework, showcasing its potential as a model for other regions facing similar challenges.

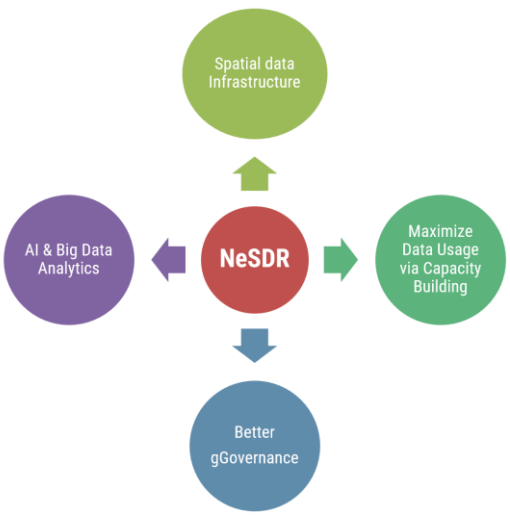

Figure 1. NeSDR and Analytics for better gGovernance

2. Data and Methodology

2.1 Interoperable data service architecture

The spatial data along with ISRO’s National Natural Resources Management System metadata standards is published following OGC data services for facilitating interoperable access by different groups of users. A network of decentralised spatial data network is established across 8 states of North East India for seamless data publishing and access. The centralised geospatial portal is populated with 1392+ geospatial products pertaining to land and water resources, administrative base information, terrain, action plan inputs, infrastructure, utilities, weather and climate, disaster support inputs etc. of NE states. It provides the datasets as per OGC standards for visualization via web services. The interface has been designed to be more responsive in all kinds of platforms and devices. It’s a platform for seamless data accessibility across various users for data visualization, interactive analysis, search and discovery of spatial based user’s interest. Further all the users can run Geo-Analytics with processing capability to generate useful user-define data products. More than 800 openly available raster datasets is leveraged for this.

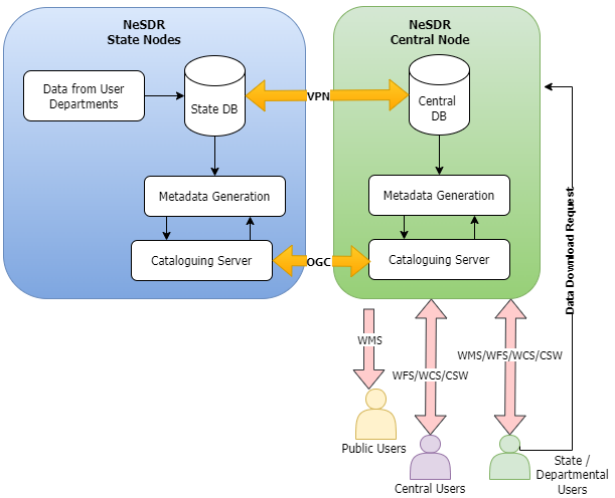

Figure 2. Spatial Data network with OGC based different accessing mechanisms

The broad architecture also enables the major goals such as (a) better g-Governance by providing robust platform for space based spatial data discovery, sharing & hosting (b) ensures data availability and carry out gaps analysis & reduce duplication of efforts & cost (c) accelerating governance via needs-oriented Spatial Decision Support System (SDSS) applications (d) adoption of emerging technologies through use of AI & big data analytics for geospatial data insights and (e) rapid development and deployment of governance applications with scalable geospatial frameworks at centralized node

2.2 NeSDR GeoAnalytics

The ability to consume available raster imagery in different resolution and apply geo-processing functionalities leads to generation of user-defined data products. NeSDR Analytics comprises model catalogues for geospatial data processing, sperate catalogue server for structurally cataloguing of geospatial raster datasets and web map server and tile map server for real time display of results. The web processing component is built using the Python programming language and the Django framework. This processing framework includes a comprehensive library with a wide range of functions, from simple mathematical operations to advanced geostatistical analysis, machine learning, and image processing capabilities. The analytics platform features tools for computing statistics on the results, as well as options for downloading the computed data.

The Figure 3 represents a three-layered architecture for a geospatial data processing system. At the top, the Client Layer includes user-facing components such as the Map Viewer for displaying maps, a Raster Calculator for performing operations on raster data, and APIs for connecting external services or applications. This layer allows users to interact with the system. The Processing Layer, in the middle, handles the actual computation and data manipulation. It consists of a WMS (Web Map Service) Server for serving map images over the web, a Compiler Engine to process user queries or commands, and the GDAL Framework for geospatial data processing. At the bottom, the Data Layer provides the storage and data infrastructure. It includes Distributed Storage for storing large datasets across different locations, a Peer-to-Peer Network for decentralized data sharing, and COG (Cloud Optimized GeoTIFF) (J. Yu et al., 2021) as a format optimized for cloud-based geospatial data storage and retrieval. Together, these layers support data storage, processing, and interaction in a scalable and efficient geospatial system.

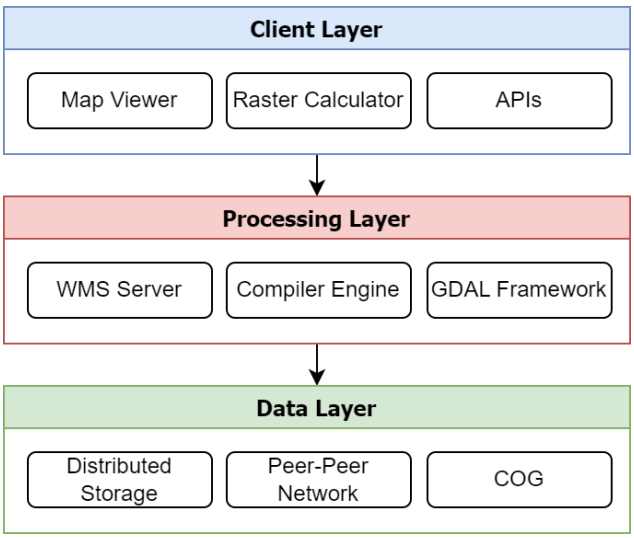

Figure 3. User on-demand computation for custom data product generation

2.3 Geospatial Technology for User oriented applications for Governance

The system can ingest multi-source data, processes and applications for building rapid user oriented applications for governance by adopting open source stacks and interoperable OGC compliance data access and delivery. The architecture above represents a comprehensive geospatial data management and processing system aimed at supporting governance applications and decision-making processes through a series of interconnected layers and modules. At the base of the architecture is the NeSDR Spatial Data Repository (C. Barnett et al., 2003), which serves as the foundation of the system. This repository collects, stores, and manages large volumes of geospatial data from various sources, including satellite

imagery, drones, and ground-based sensors. It also manages the associated metadata and ensures that all data services are compliant with the Open Geospatial Consortium (OGC) standards, enabling seamless integration and exchange of data across different platforms. The data stored in NeSDR is fed into the Processing and Analysis Layer, where it undergoes various forms of analysis. This layer is crucial for transforming raw geospatial data into actionable insights. It facilitates geospatial analysis, monitoring of activities, planning, and decision support, making it a central point for processing complex geospatial operations. This layer also establishes a feedback loop with the higher governance application layers, ensuring that insights and outcomes from the analysis are used to refine and optimize the system over time. Feedback from real-world applications helps in continuously improving the accuracy and efficiency of the geospatial services. The micro-services Layer comprises of a range of specialized geospatial tools and operations. These micro-services (Bebortta, S et al., 2020) are customized to perform specific tasks that are essential for analyzing and processing geospatial data. For instance, functions such as routing, tracking, proximity analysis, and buffer operations are handled in this layer, ensuring that different types of spatial analysis can be performed efficiently. Change and trend analysis services are also part of this layer, which enable the detection of changes over time in spatial datasets. These operations support both real-time and historical analysis, giving stakeholders insights into past trends and future projections.

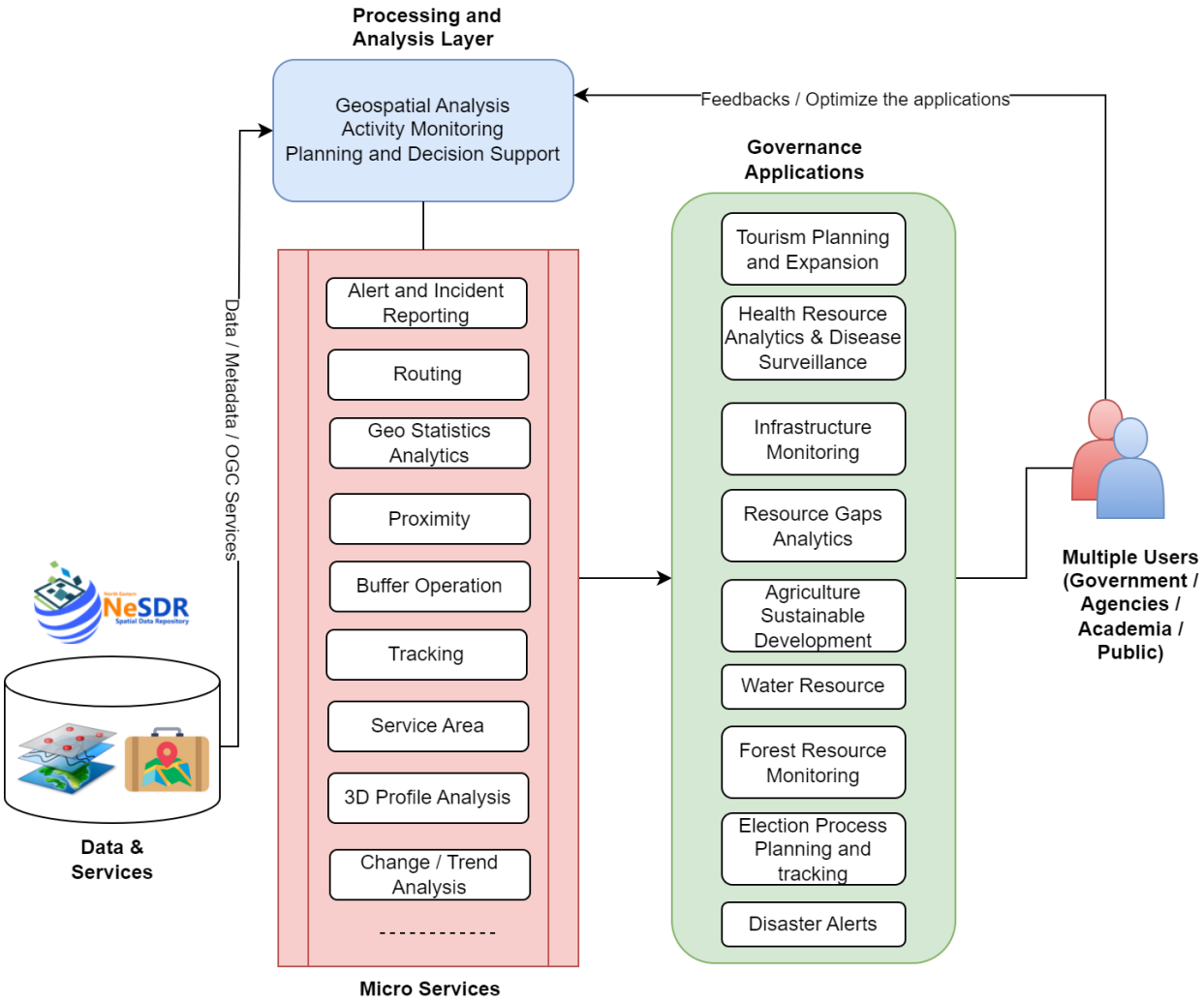

Figure 4. Micro-service based architecture for building rapid and configurable user applications

All these micro-services and analyses can be fed into the Governance Applications, which represent the practical, user-facing aspect of the system. Various sectors, including tourism planning, health resource management, infrastructure monitoring, agriculture, water and forest resources, and disaster management, are supported by the geospatial analysis provided by the system. These governance applications rely on the processed data and analytical insights to make informed decisions, optimize resource management, and improve overall planning. The system is therefore designed to serve a diverse range of users, including government agencies, academia, and the general public. Each of these user groups can access and interact with the geospatial data and services through the governance applications. The multi-user design ensures that the system has a broad impact, supporting both public services and research initiatives.

3. Results

3.1 GeoAnalytics UI for custom data product generation

A web-based UI interface is provided along with inbuilt raster

calculator allowing users to perform custom band math operations with ease. Queries to the NeSDR analytics platform are constructed using functional composition and evaluation. Users build queries by chaining together operations that includes over multiple functions. These functions vary in complexity, ranging from simple mathematical operations to advanced geostatistical analyses, machine learning algorithms, and image processing tasks.

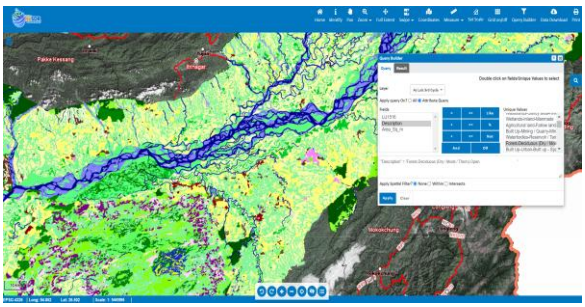

Figure 5. In-Browser interactive Visualization and Analysis

The NeSDR analytics presents a workflow for geospatial data analysis using raster and vector operations, applicable to various domains like vegetation, weather, water resources, and satellite imagery. It begins with selecting a specific domain for analysis, such as vegetation or weather, allowing users to focus on the environmental parameters relevant to their needs. After selecting the domain, users can choose a specific dataset and apply relevant processes over a defined date range. This allows the user to analyze temporal changes in the data, such as monitoring vegetation growth over different seasons or evaluating weather patterns over a given period. The interface provides options for dataset selection, data processing operations, and time-specific filtering.

The central part of the diagram shows key operations like Change and Trend Analysis and the Raster Calculator. The Change and Trend Analysis tool helps in identifying spatial and temporal changes within the dataset, offering insights into patterns and shifts over time, like vegetation health or water levels. The Raster Calculator enables users to perform advanced computations, applying mathematical formulas to the raster data for custom analysis, such as calculating averages, creating indices, or comparing data from different time periods. The final outputs are generated in the form of new raster maps, generated as a result of the operations. These raster products visually represent the processed data and can highlight important trends, such as vegetation density, weather anomalies, or water resource availability. The outputs are colour-coded maps that are useful for environmental monitoring, decision-making and strategic planning, providing a clearer understanding of the region’s dynamic conditions. Thus NeSDR Analytics architecture streamlines the process of accessing, processing, and analyzing multi-source geospatial data to generate valuable insights, crucial for resource monitoring and environmental management.

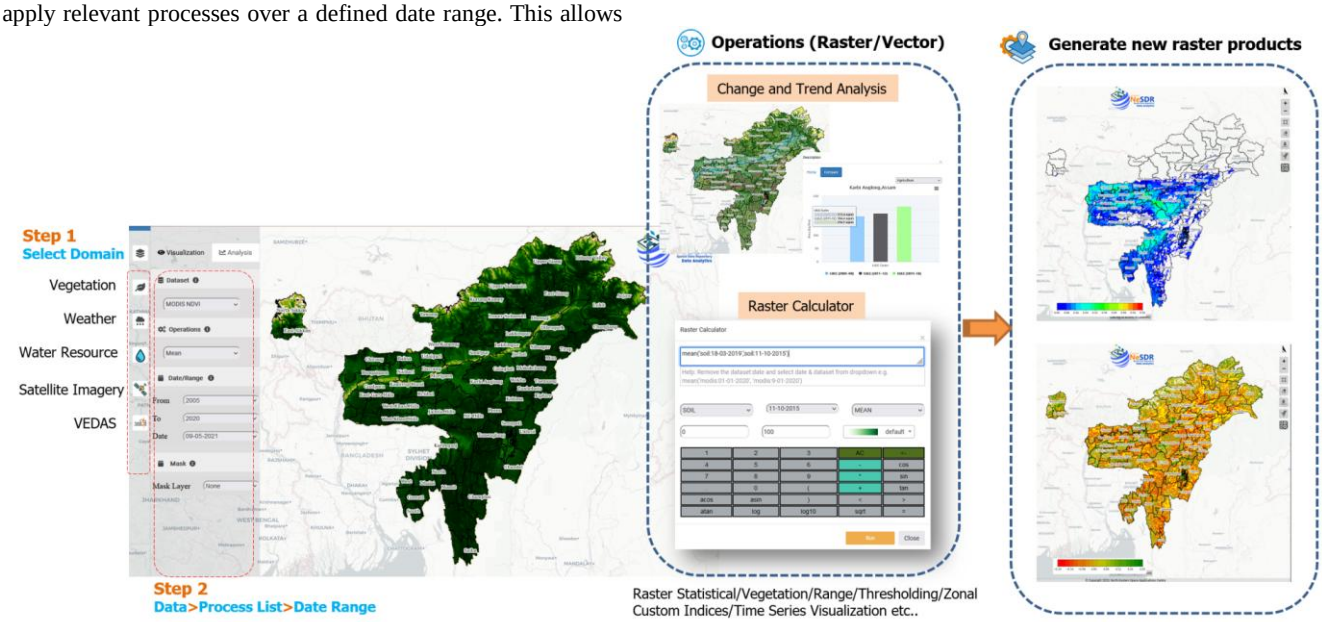

Figure 6. UI interface NeSDR Analytics

3.2 Efforts towards good g-Governance

Several geospatial based g-Governance applications have been designed, hosted and put to operational using interoperable data services of NeSDR. These applications are utilized by various government departments for purposes such as i) health resource management and planning. They aid in monitoring, including disease hotspot analysis, ii) geospatial application to enhance the electoral process by tracking incidents and alerts to ensure timely event detection and smooth elections (Figure 8), iii) space-based application to monitor the physical progress of projects by combining satellite, drone, and app-based spatio-temporal data (Figure 9), iv) support sustainable agriculture, improve tourism planning, and expand tourism in the region etc. As recognition to these impactful applications built with open source, some of these applications were awarded with National Award for e-Governance 2022 (https://nesac.gov.in/events/2022/11/national-award-for-e-governance-2021-22-for-north-eastern-spatial-data-repository/) and eNorth East awards 2023 (https://nesac.gov.in/events/2023/11/the-enortheast-award-2023-for-nec-geotagging-geo-monitoring-project/).

Empowering Electoral process in NER	• Mapping of Electoral Assets • Real time tracking of Alerts and Incidents
Ease of Projects Monitoring	• Physical Progress Monitoring via Mobile, Satellite, UAVs & IoTs
Better Health Surveillance	• Health Resource Optimization • Incidents reporting on diseases and hotspots
Early Warnings on Disasters	• Smart Axom (Assam) for disseminating alerts and early warnings of lightning & flood
Increase Agriculture Productivity	• Drone based Agro-advisory • Rice Information Systems
Forestry resource management	• Forest Resource Monitoring and Alerts • Demarcation of forest pattas
Expanding Tourism in NER	• Tourism planning : 1153+ tourist locations • Expansion of Tourism: Tourism potential index
Infrastructure Management	• Gap Area Assessment of health and education
GeoAI for better Governance	• AI based road quality monitoring • Precision Agriculture | Change detection

Figure 7. Geospatial g-Governance developed using Open-source data services and applications

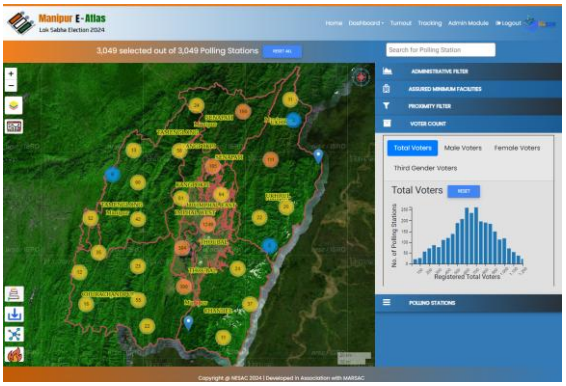

Figure 8. Election planning and monitoring dashboard, https://apps.nesdr.gov.in/election/

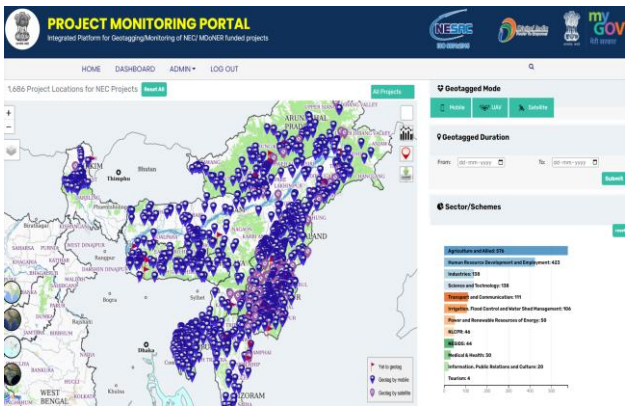

Figure 9. Project monitoring portal, https://nec.nesdr.gov.in/

3.3 Outreach and capacity building for utilization

To promote maximum geospatial data utilization, reduce data duplicity and promote data sharing across diverse groups of users in the region, several wide scale capacity building programmes have been conducted and sensitized. Specific capacity building programmes were organised for remote sensing centres, state run departments and disaster management groups etc. Further several workshops have been organized via online/offline modes and as a result approximate 1500 personnel have been trained and benefited so far.

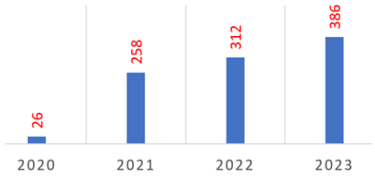

Figure 10. Increase use of geospatial data services by users

4. Conclusion

The continuous availability of remote sensing-derived datasets, combined with user-centric open-source and scalable geospatial applications, will continue to facilitate ongoing planning and monitoring activities in the region. Platforms like NeSDR will remain pivotal for efficiently searching and discovering these datasets. With the recent opening of the new geospatial policy to public and private sector enterprises, a substantial increase in data generation and consumption across various sectors is anticipated. In this regard, the datasets and tools offered by NeSDR will further enhance planning efforts in different areas which can accelerate developmental progress in the region. Increased sensitization and awareness regarding interoperable data sharing will benefit a wider audience. The emergence of advanced open-source tools and standards for geospatial data dissemination, visualization, and analytics will continue to contribute significantly to the enhance governance, benefits academia community, and build more socially relevant applications. The platform ensures the availability of up-to-date and diverse geospatial big data in various resolutions, scales, and formats. It also focuses on adding more user-oriented and demand-based geo-analytics to cater to larger user departments. By adopting new and emerging tools, the platform aims to deliver services more efficiently. Additionally, the platform supports developing eGovernance solutions to target a broad range of user groups. The success of NeSDR has led to replication of this open source and scalable technology to other states, with necessary customizations to meet their specific needs.

5. References

J. Yu *et al.*, "Realtime Remote Sensing Image Publishing System Based on COG Technology," *2021 28th International Conference on Geoinformatics*, Nanchang, China, 2021, pp. 1-5, doi: 10.1109/IEEECONF54055.2021.9687665.

C. Barnett et al., "Improving GRASP: a geospatial repository for analysis and safety planning," IEEE Systems and Information Engineering Design Symposium, 2003, Charlottesville, VA, USA, 2003, pp. 113-118, doi: 10.1109/SIEDS.2003.158013

Bebortta, S.; Das, S.K.; Kandpal, M.; Barik, R.K.; Dubey, H. Geospatial Serverless Computing: Architectures, Tools and Future Directions. *ISPRS Int. J. Geo-Inf.* 2020, *9*, 311. https://doi.org/10.3390/ijgi9050311