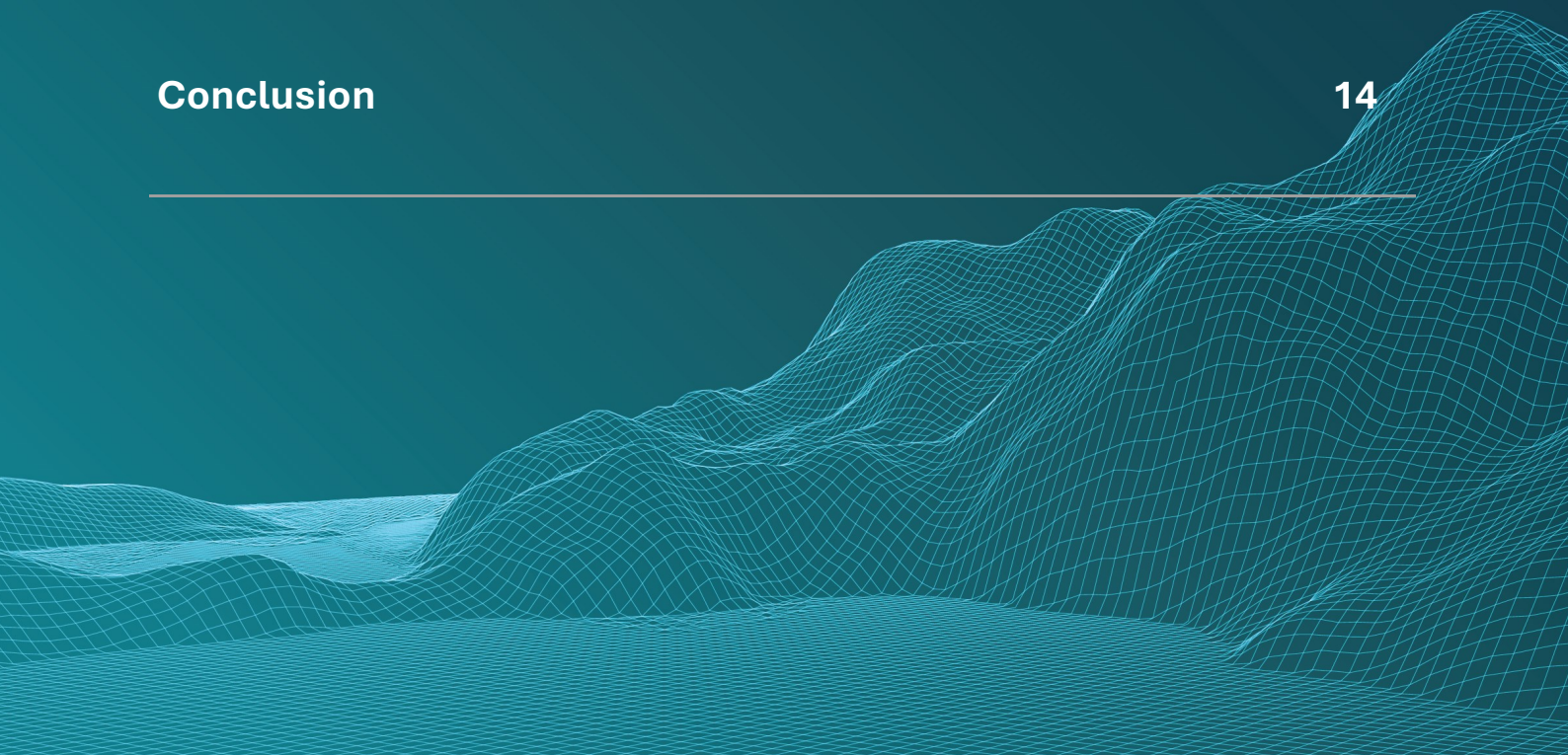


Integrated Geospatial Data-Sharing Interface: Building Digital Public Infrastructure for Geospatial Data



Table of Contents

Introduction	1
Data exchange platforms as Digital Public Infrastructure	3
GDI : The data exchange platform for geospatial data	5
Some technical details of GDI	6
Advantages of GDI	9
Some use cases	11
Conclusion	14



INTRODUCTION

The focus of the National Geospatial Policy 2022, is to make geospatial technology and data, the agents of transformation for achieving the Sustainable Development Goals (SDGs). The aim is to establish a coherent national framework within the country, leveraging it to transition towards a digital economy and improve citizen services. This includes ensuring the accessibility of valuable geospatial data, funded by the public sector, to both businesses and the general populace. In addition to that it should also foster a robust geospatial industry, encouraging active participation from private enterprises.

Furthermore, efforts are directed towards establishing and enhancing an integrated interface for all digital data with spatial components, facilitating seamless access, sharing, and reuse. A pivotal aspect involves the development of a Geospatial Knowledge Infrastructure (GKI), supported by an Integrated Data and Information Framework, to streamline information dissemination and utilization within the geospatial domain.

As technology plays an increasing role, the effective exchange of geospatial data is becoming critical for the sustainable development and

success across sectors like agriculture, healthcare, urban planning and disaster management, etc. The ability to collect, analyse, and share geospatial data has the power to drive innovations and empower the stakeholders like application developers, researchers and policymakers.

Geospatial data exchange enables seamless flow of data and insights among stakeholders in the geospatial ecosystem. It involves the sharing of diverse datasets, such as geo-tagged locations, satellite based imagery, aerial photographs, topographic maps, land use and land cover data, soil characteristics, climate information, infrastructure details, and environmental indicators. This exchange allows for informed decision-making, improved resource management, and improved planning and response across the entire ecosystem. One of the key drivers behind the need for geospatial data exchange is the belief in FAIR principles (**F**indable, **A**ccessible, **I**nteroperable and **R**eusable). This entails the need to break the data silos, make the data discoverable; seamless sharing in standardised manner; interoperable datasets to avoid painstaking georeferencing, and conversion across coordinate systems, making usage terms and licensing easily discoverable, and create essential interfaces for identity verification.



Geospatial data exchange plays a pivotal role in various domains, facilitating informed decision-making and strategic planning. In urban settings, it enables the analysis of Land Use Land Cover (LULC) patterns, which is crucial for urban planning and development. Additionally, geospatial data aids in monitoring urban deforestation and supports infrastructure planning, ensuring sustainable growth. In the agricultural sector, geospatial data may facilitate soil monitoring, enabling farmers to make informed decisions regarding crop management. Moreover, it aids in farm delineation and yield estimation, optimizing agricultural practices and enhancing productivity. In disaster management, geospatial data exchange is indispensable for predicting and mitigating natural disasters, enabling authorities to allocate resources effectively and implement preventive measures. Geospatial data exchange also aids in conservation efforts, water resource management

and facilitating proactive measures to mitigate environmental degradation. Geospatial data exchange can also enhance healthcare by improved disease tracking, resource allocation, and patient care delivery. Overall, geospatial data exchange serves as a cornerstone for informed decision-making across various sectors, fostering sustainable development and resilience.

The benefits of geospatial data exchange extend beyond individual stakeholders. By facilitating the exchange of data, the entire geospatial community can leverage collective intelligence and drive systemic improvements. It enables the identification of industry-wide trends, the benchmarking of performance, and the dissemination of best practices. This collective effort contributes to the growth and competitiveness of the geospatial sector as a whole.



DATA EXCHANGE PLATFORMS AS DIGITAL PUBLIC INFRASTRUCTURE

Last year, the Hon'ble Prime Minister of India, Shri Narendra Modi, spoke about Digital Public Infrastructure (DPI) during his speech to the U.S. Joint Session of Congress, stressing about its importance for our nation and India's willingness to help the Global South build their respective DPI platforms by leveraging India's experience and technology.

DPI is a set of digital platforms and IT systems that serve public good, forming the foundation for the nation's digital economy. As per a report by NASSCOM, by 2030, economic value addition from DPIs has the potential to increase to 2.9-4.2 percent of GDP, from 0.9 percent in 2022.

Three aspects of the digital economy are envisaged as part of the foundation - the triad of personal identity, financial payments and data exchange. The first two are quite mature and well-established in India through Aadhar and UPI respectively. The third aspect, data exchange, is very diverse and unique within each sector. The data exchanges are the third piece of the India stack, which focuses on establishing a new data governance model in India. National Geospatial Policy 2022 focuses on making geospatial technology and data as agents of transformation and achieving the SDGs.

Exchanging data to enable the creation of applications or services is complicated for a wide variety of reasons. In fact, the availability of data is the single hardest issue when an application developer wishes to create an application to deliver a public service. The ability for application developers and particularly start-ups to experiment and innovate with new data-driven services must be fostered. Hence the need for a data exchange platform, which must provide the following capabilities:

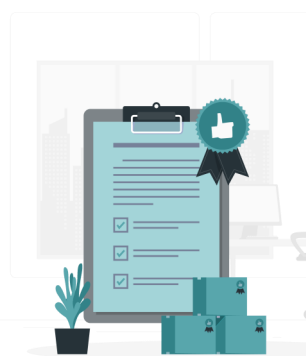
- **Discoverability of data:**

Most innovators and organizations waste a lot of resources in finding and acquiring appropriate data. The ability to identify pertinent data with a searchable catalogue is an essential aspect of a Data Exchange. The catalogue must be programmatically searchable and contain standardized meta-data descriptions.



- **Standardisation of API's and data models:**

A major drain of resources comes from dealing with heterogeneity. Different data providers represent data in different formats, and this imposes costs and complexity on application development. Access to data through standardisation of API's and data models is an important part of Data Exchange.



- **Controlling data access:**

Data providers must have the ability to control who can access a data asset, in part or whole, and restrict access until an agreement is in place or payment is complete.



- **Policy-based and consent-based data sharing:**

A policy-driven architecture is one where data is shared with a data consumer in a manner consistent with a specified data sharing policy. For personal data, a data exchange will preserve the privacy of such data by ensuring that any personal data is shared only if explicit consent is provided by the concerned individual.



In addition to these functional requirements, there are several non-functional requirements that a data exchange designed for DPI must satisfy.



- **Decentralisation & Federation:**

There must be no requirement to centralise control or storage of data. Each data provider must retain full control and possession of their own data. There may be hundreds of data exchanges, based on sectoral or application requirements. A catalogue of catalogues will provide a federated single-system view of the set of data exchanges, with the ability to search across them.

- **Open-source:**

The exchange should be designed to be an open-source software system. In general, any DPI must be based on open-source, or it will be under the control of a proprietary vendor. This will also enable all data exchanges within the country to be based on a common open source code base that can be sub-setted and customized for specific deployments or sectors.



These requirements are very different from those of an **Open Data Portal** as a Data Exchange can control who can get access based on defined policies. The requirement for decentralisation and federation is essential for DPI and makes the design of a data exchange quite different from data sharing platforms common in enterprise settings such as **Data Warehouses** and **Data Lakes**.

GDI : THE DATA EXCHANGE PLATFORM FOR GEOSPATIAL DATA

The Department of Science and Technology (DST), and the Indian Institute of Science, Bengaluru (IISc) have collaborated over the past year to create the nation's first Integrated Geospatial Data-sharing Interface (GDI). GDI is a platform that embodies all the requirements of the DPI mentioned earlier and is intended to address data sharing in the Geospatial sector. GDI is based upon the India Urban Data Exchange (IUDX), a pioneer in enabling the use of data for public good and created in IISc under the auspices of the Ministry of Housing and Urban Affairs (MoHUA). IUDX has enabled a variety of platforms, tools, and applications that have benefitted urban citizens.

One of the biggest barriers to creating innovative applications and services to support the geospatial

sector in India is the lack of access to high-quality and up-to-date geospatial data. Start-ups and other companies involved in such efforts can spend many weeks or even months searching for data and petitioning for access. GDI democratizes and standardizes access to geospatial data, helping innovators focus on solution creation for various sectors, including agriculture, urban planning, disaster management, environmental monitoring, infrastructure development, and healthcare. The cloud-based software platform facilitates the secure, standards-based exchange of data between users of geospatial data (e.g. application developers) and geospatial data providers (e.g. government agencies, private companies, NGO's, universities, etc).

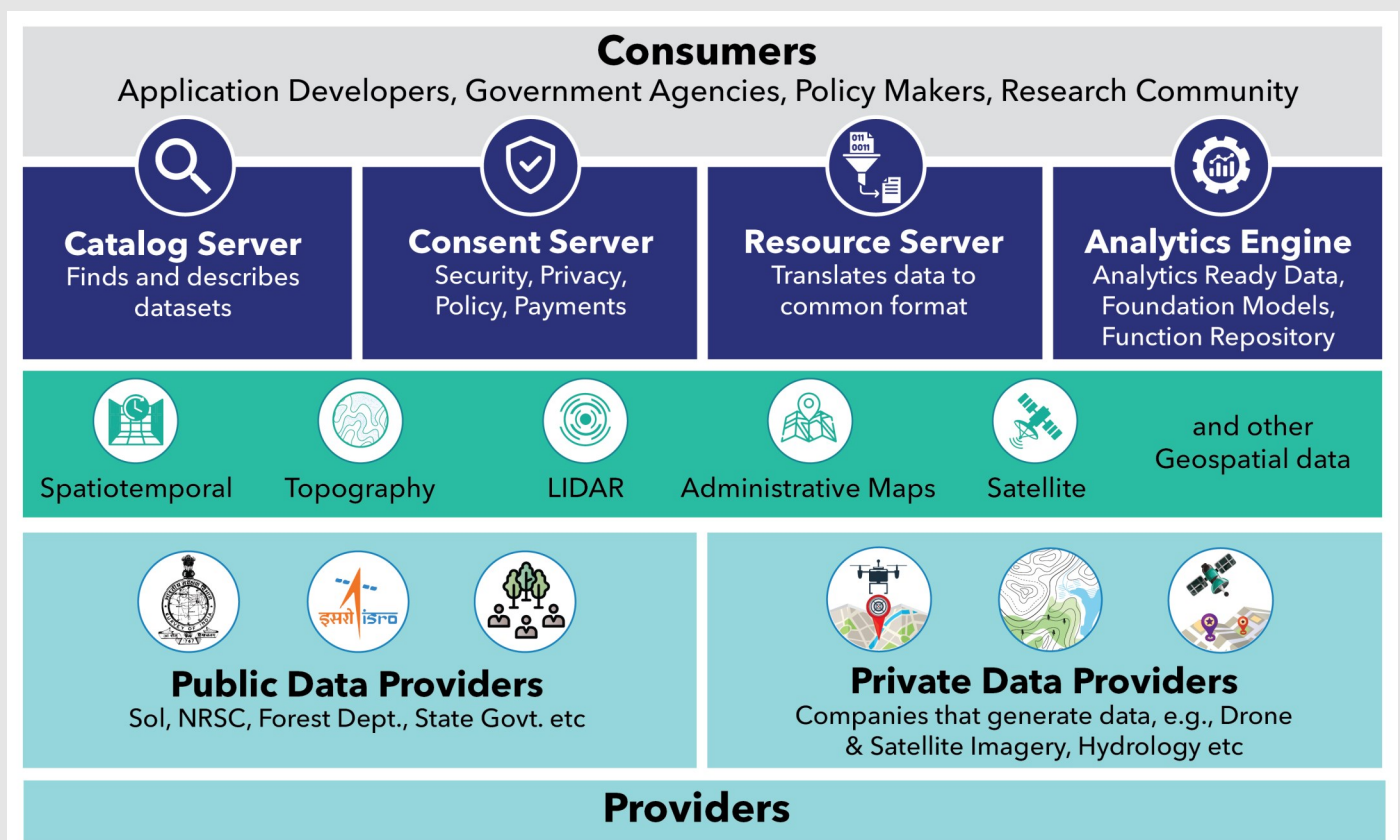


Figure - 1

GDI has the potential to revolutionise the way we approach the geospatial sector. By embracing data sharing, stakeholders can harness the power of information to drive innovation, improve sustainability, and address the challenges posed by a growing population.

SOME TECHNICAL DETAILS OF GDI

GDI does not store any data itself, but provides a standards-based catalogue of available geospatial data, a comprehensive set of authorization and consent management capabilities, standardized API's and data models to make data available, and a suite of analytics to help the creation of use cases and solutions that can benefit the ecosystem. Like IUDX and ADeX, GDI is a highly scalable, interoperable, open-source and secure platform, based on state-of-the-art software methods and practices. The platform provides full control to the data owners as to what data to expose and to whom. Built-in accounting mechanisms enable connection with payment gateways which will form the foundations for a data marketplace. The whole platform is developer friendly, via definitions of open APIs and data schema templates (formats for interpreting data), so that a whole new application ecosystem gets created.

Utilizing advanced geospatial data analysis technologies and harnessing their benefits is crucial for achieving the Sustainable Development Goals (SDGs). A vital component in this effort is ensuring that application developers have easy access to geospatial data. The National Geospatial Policy aims to provide efficient access to fundamental geospatial data held by various stakeholders through electronic data querying and processing services.

Integrated Geospatial Data-sharing Interface (GDI) described in this document consists of services facilitating the exchange of geospatial data between providers and consumers via one or multiple resource servers, contingent upon explicit authorization acquired from the data provider. The interfaces are shown in Figure 1 and a high level architecture diagram is shown in Figure 2.

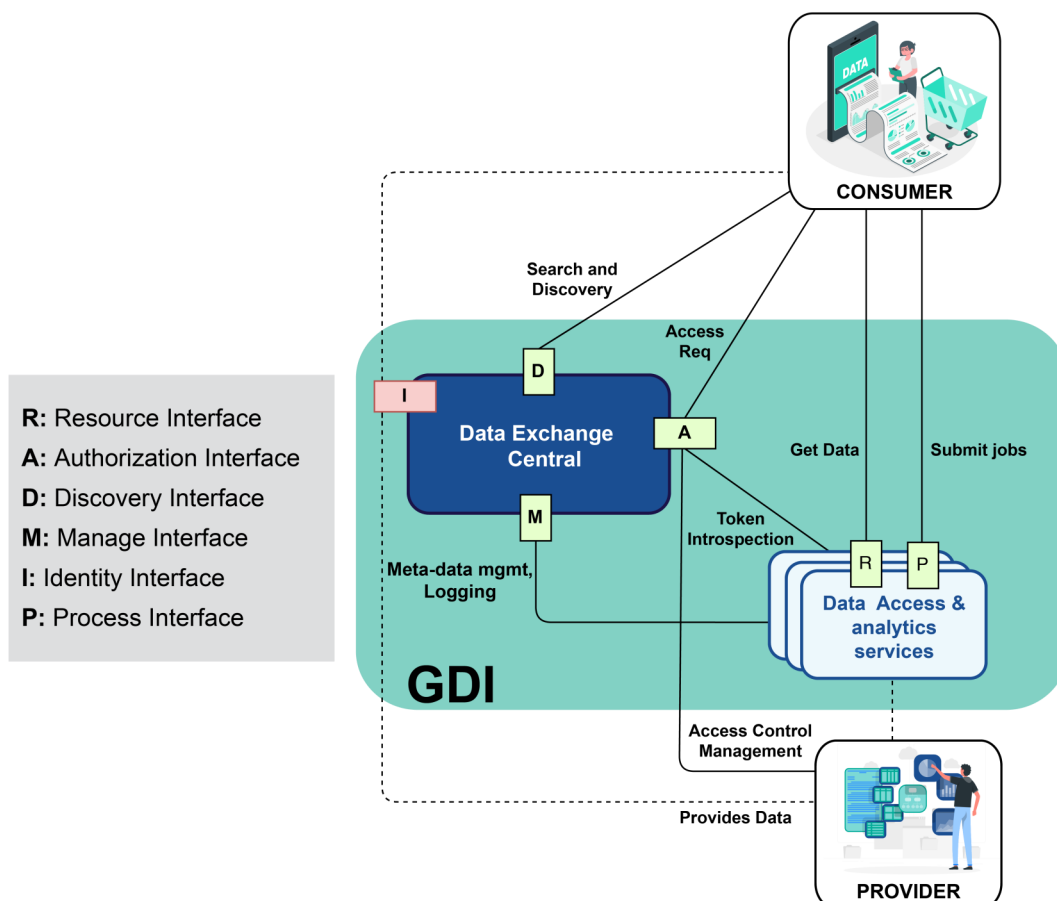


Figure 2: GDI Interfaces

The GDI platform comprises the following interfaces:

- Management interface
- Discover interface
- Authorization interface
- Resource interface
- Identity interface
- Process Interface

Resources, administered by a data provider, reside on one or more resource servers and are accessible to entities via meta-information descriptions (such as data descriptors, formats, data providers, etc.) through a catalogue, also known as a data explorer, utilizing the Discovery Interface (D). The discovery interface supports searches based on geo-search, temporal search, attribute search, text search etc. The catalogue entries are represented in JSON-LD format which is both human-readable and machine-understandable.

The provider registers using the Identity Interface (I) and manages the metadata of its resources and their associated access control policies through the Manage Interface (M) of GDI. The metadata of each resource should aid data consumers (application developers) in facilitating resource consumption to create valuable applications. Data consumers can register with GDI to receive notifications about any changes to the metadata of resources of interest to the consumer.

Data consumer seeks authorization to consume resources via the Authorization Interface (A) by acquiring an authorization token. Any request by a consumer application to access a provider's

resource undergoes scrutiny against existing access control policies. If no decision can be made, GDI may facilitate an access request notification to be generated. Communication between the provider and the consumer occurs outside the scope of this specification and can be established using various messaging technologies such as SMS, OTP, EMAIL, etc. Data licensing terms and conditions are also beyond the scope of this specification; however, references to licenses can be provided in the resource metadata.

Upon obtaining consent and an authorization token, the entity utilizes the Resource Interface (R) to access the resource. The resource interface provides data access using standard published OGC REST APIs such as OGC Feature APIs, OGC Coverage APIS, OGC Tiles APIs etc. GDI resource access service component will be available as a deployable open-source component that the data providers can deploy to provide access using standardized interfaces to the GDI data consumers. GDI supports a distributed architecture for data access which, apart from avoiding data bottlenecks, also frees up the providers from routing their data through a centralized data exchange.

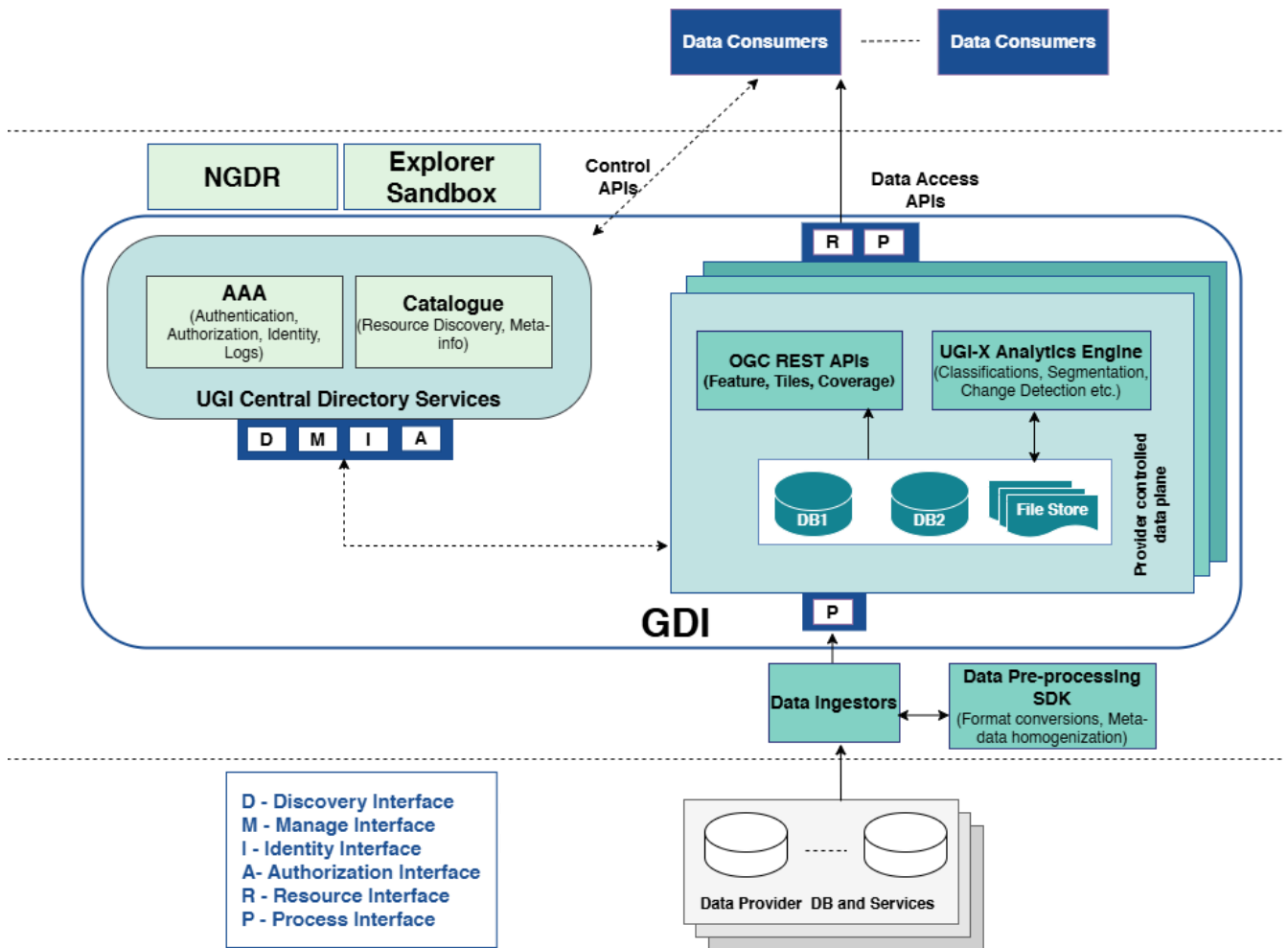


Figure 3: GDI High Level Architecture

Further, GDI makes available a deployable software component to implement a process execution framework (P), collocated with GDI resource server, that enables local computation on the geo-spatial data. It consists of a set of pre-defined (or partially defined) optimized analytic building blocks to run on collocated data and make query specific results available through APIs.

GDI system is comprised of the below core components that are mentioned below:

1. **Data Explore Service:** This service is used for the discovery and management of meta-data of the data available with the data exchange. This service shall be used by any provider, consumer to onboard or discover datasets in GDI.

2. **API Gateway Service:** This service is used for serving the data associated with a resource in encrypted or unencrypted format with an A after verifying the authorisation token. This service shall be used by any Data providers and Data Consumers who wish to use GDI.
3. **Authorization and Identity Service:** This service is used for user profile management on GDI and requesting access tokens which can be presented at the API gateway for accessing resources. This service shall be used by any provider and consumer who wishes to use GDI.

The deployment of GDI is technology neutral, reliable, consumer centric, auto scalable, cloud ready, vendor agnostic with security and privacy being the primary design consideration.

ADVANTAGES OF GDI

1. Improved Data Accessibility

- **Centralized Repository:** A unified platform allows government agencies to access a comprehensive repository of geospatial data, including satellite imagery, terrain maps, and demographic information.
- **User-Friendly Interface:** A well-designed platform can simplify data retrieval and visualization for non-technical users, enabling broader usage across agencies and domains.



2. Enhanced Decision-Making

- **Informed Policy Formulation:** Agencies can leverage geospatial data to make informed decisions related to urban planning, disaster management, and resource allocation.
- **Data-Driven Insights:** Access to accurate geospatial data facilitates evidence-based insights that can inform strategies and policies.

3. Collaboration Across Departments

- **Inter-Agency Coordination:** The platform can foster collaboration among various government agencies, enabling them to share data, insights, and best practices.
- **Integrated Solutions:** Combining data from different sectors (e.g., agriculture, transportation, health) allows for a holistic approach to problem-solving.



4. Real-Time Monitoring and Response

- **Disaster Management:** Agencies can use real-time geospatial data for monitoring natural disasters (like floods, earthquakes) and implementing effective response strategies.
- **Infrastructure Monitoring:** Continuous tracking of infrastructure projects can help in timely maintenance and issue resolution.

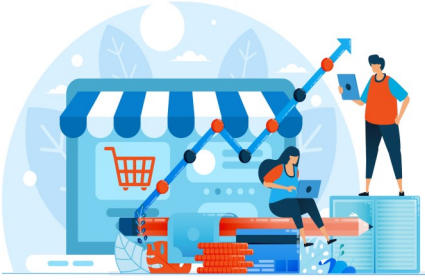
5. Public Services and Citizen Engagement

- **Transparency and Accountability:** By making geospatial data accessible to the public, agencies can enhance transparency in governance and build trust with citizens.
- **Citizen Services:** A geospatial platform can support various public services, such as land registration, urban planning, and transportation services, improving service delivery.



6. Rapid Application Development

- **Product Development:** The start-up ecosystem and private companies benefit from access to datasets from diverse sources. Datasets which are interoperable with rich metadata and data descriptors help in smaller product development cycles.
- **Data Marketplace:** The platform can aid organisations possessing valuable geospatial data to monetise their data thus creating a data marketplace.



7. Capacity Building and Training

- **Skill Development:** A centralized platform can facilitate training programs for government employees on data usage and analysis, enhancing their technical capabilities.
- **Knowledge Sharing:** Regular workshops and seminars can promote knowledge sharing on geospatial technologies and their applications.



8. Regulatory Compliance and Policy Adherence

- **Compliance Monitoring:** Agencies can use geospatial data to ensure adherence to environmental regulations, land use policies, and urban development guidelines.
- **Policy Evaluation:** The platform can assist in evaluating the effectiveness of existing policies by analyzing spatial data over time.



9. Sustainable Development Goals (SDGs)

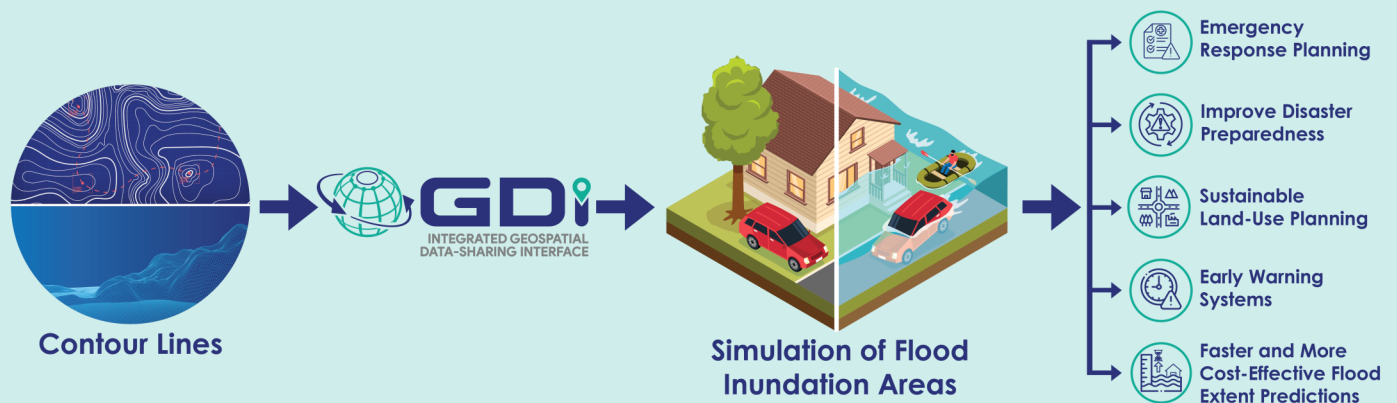
- **Tracking Progress:** Geospatial data can help monitor progress towards achieving SDGs, such as sustainable cities, climate action, and responsible consumption.
- **Resource Allocation:** Targeted interventions based on spatial data can lead to better resource allocation and management.



SOME USE CASES

Flood Inundation Model

Flood inundation, the submergence of land due to rising water levels from extreme weather events like heavy rainfall or river overflow, is a critical concern for disaster preparedness, emergency response, and effective damage assessment. Accurate mapping of flood-prone areas enables authorities to develop early warning systems, plan evacuation routes, and allocate resources efficiently. Contour data available on GDI helps in the identification of flood zones using foundational models. These outputs are invaluable for disaster preparedness, with static maps highlighting high-risk areas for potential protection or evacuation, and animated sequences offering insights into flood progression under different conditions. Together, these tools support early warning systems, optimize resource allocation during floods, and aid in post-flood damage assessment.



Mapping Natural drains

Natural drains serve as vital channels for collecting and transporting surface water, significantly impacting ecosystems and providing critical insights for water resource management, flood risk assessment, and environmental protection. Contour data available in GDI can be used to map natural drains. This mapped network of natural drains has diverse applications, including identifying areas at risk of water pollution, supporting biodiversity conservation by preserving critical watercourses, and assessing flood risk since changes to natural drainage can worsen flooding. Additionally, understanding natural drainage patterns aids in locating potential groundwater recharge zones, makes it a powerful tool for sustainable water resource planning.



Routing Applications

Routing is a fundamental aspect of modern mapping and navigation, enabling users to find the shortest or most efficient path between two points on a road network. Routing can be implemented to calculate distance, duration, and the shortest path between points of interest (POIs). Practical applications of this solution include enhanced navigation for tourists, optimized local transportation and delivery routes, and efficient emergency response routing. This case study highlights how integrating routing with local POI data can improve user experiences and offer valuable insights for urban planning and city management.



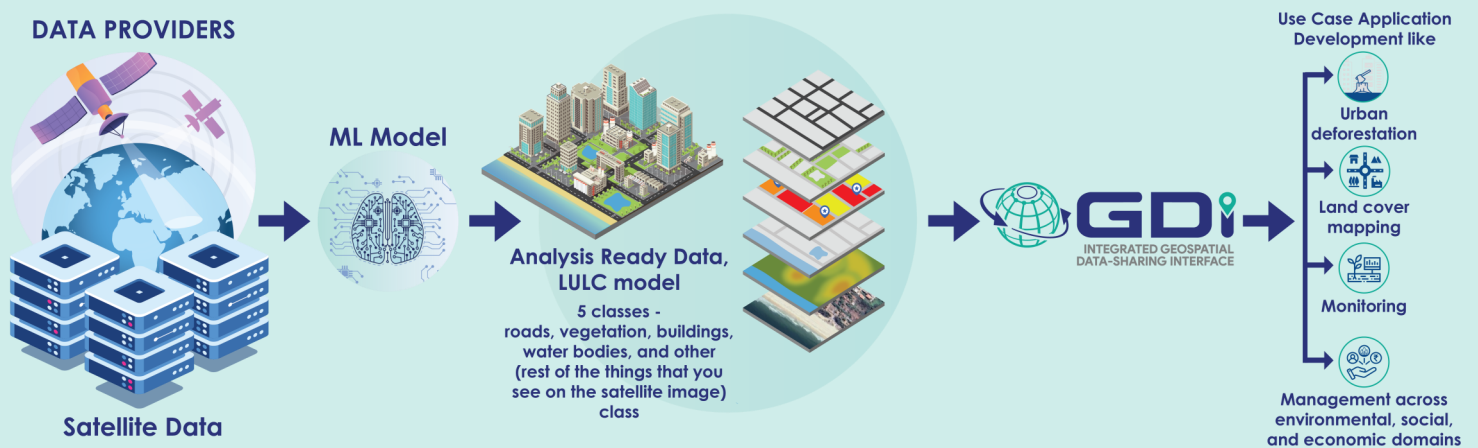
Modelling Agricultural Land Suitability

Agricultural land suitability can be modelled focusing on soil properties to optimize agricultural productivity. The approach assesses factors like soil depth, texture, pH, drainage, and slope, assigning scores to indicate each area's agricultural potential. The final land suitability map can then be visualized, offering a clear view of areas with the greatest potential for productive farming. This mapping approach not only helps maximize agricultural land use but also supports sustainable practices by minimizing the need for chemical inputs in areas where crops naturally thrive. The suitability map provides farmers with data-driven insights that promote both economic gains and environmental goals, showcasing how technology and soil data can guide efficient, sustainable agricultural practices.



Land Use Land Cover

Land Use and Land Cover (LULC) segmentation classifies Earth's surface into categories like urban areas, forests, agricultural fields, and water bodies, using satellite or aerial images. This segmentation finds applications in Change Mapping, Natural Resource Monitoring, Disaster Management, Green Space Allocation and Tree Canopy Management, Crop Monitoring, Environmental Impact Assessment, Smart City Analytics and IoT Planning, Utilities and Resource Allocation, and Road Network Analysis and Traffic Management.



CONCLUSION

A data exchange platform for geospatial data in India can transform how government agencies operate, enabling them to become more responsive, efficient, and collaborative. By leveraging geospatial data, these agencies can address complex challenges, improve public services, and contribute to sustainable development goals.

