HARNESSING SPATIAL DATA INFRASTRUCTURES: A SCALABLE, BOTTOM-UP, DISTRIBUTED APPROACH

Dr. Tubolayefa Warekuromor

INTRODUCTION

- Utilises the case of the Nigerian SDI attempt
- Explores the use of free and open source data to create an SDI demonstrator
- Deploys a framework within which the protocol put forward by the demonstrator can flourish

MOTIVATIONS

- There are still notable problems with the access, use and maintenance of spatial data.

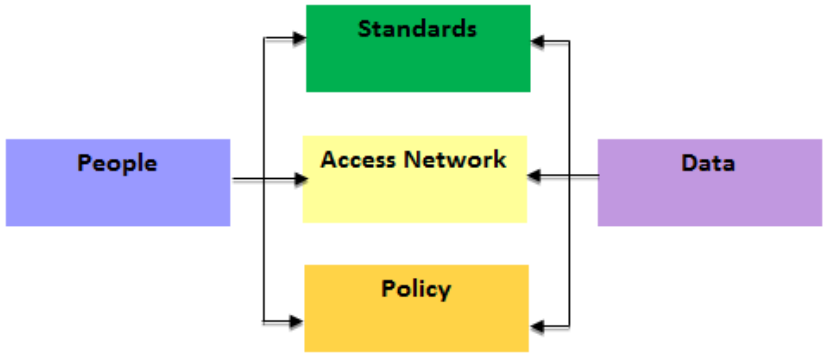
  - Availability
  - Poor sensitisation and awareness
  - Access
    - Administrative restrictions and the bureaucratic procedures
    - Security, privacy and access control
    - Unwillingness to share data due to
    - Lack of collaboration
    - Absence of a unified source of accessing domain-specific data
  - Accuracy and quality
    - The presence of data sets with missing components, diverse scales, formats, standard and resolution
  - Usability
    - Data does not satisfy the user requirements.
    - Incomplete, inconsistent and inaccurate data
    - Metadata has not been properly documented or is unavailable

Data Standardization

- Data standardization entails the procedures and protocols employed to define the data components, specifications, fields and values to ensure conformance to stipulated standards as well as the interoperable use of the data
- Open Geospatial Consortium (OGC) and International Standard Organisation (ISO) standards have been advocated to help foster interoperability
- They develop open source spatial standards to support the development of geospatial data and geospatial systems on both proprietary and non-propriety interfaces.
- SDI employs these interoperable standards with the goal of enabling the seamless access to geospatial data

INTRODUCTION

- SDIs are yet to sufficiently address the problems of seamless spatial data access, collaboration and sharing.
- There are still issues with fully implementing the access networks at all levels.
- There are also cases where regulations do not effectively translate into practice
- A continuously evolving concept that can contribute significantly to economic development, environmental management and social stability in both developing and developed countries if effectively adopted



SDI

- Anticipated to integrate data from different sources, in the right format, and enable the interoperable access to that data via a clearinghouse.

- Earlier implementation of SDI typically follows the top-down approach, which involves the definition of the policy, technical standards and data standards before implementing the clearing house

- The implementation, data standardization and data preparation is usually initiated and managed by the government or its designated agency

OVERACHING QUESTION

Using the Nigerian case, we assessed the insufficiency of the NGDI to provide comprehensive spatial data access in the context of environmental management.

- How can a scalable and sustainable SDI be developed which overcomes failings of the NGDI project?

WHY OPEN SOURCE?

- Free and open source geospatial technologies offer cost effective ways to deploy scalable, secure and interoperable SDIs but they are yet to be fully exploited.

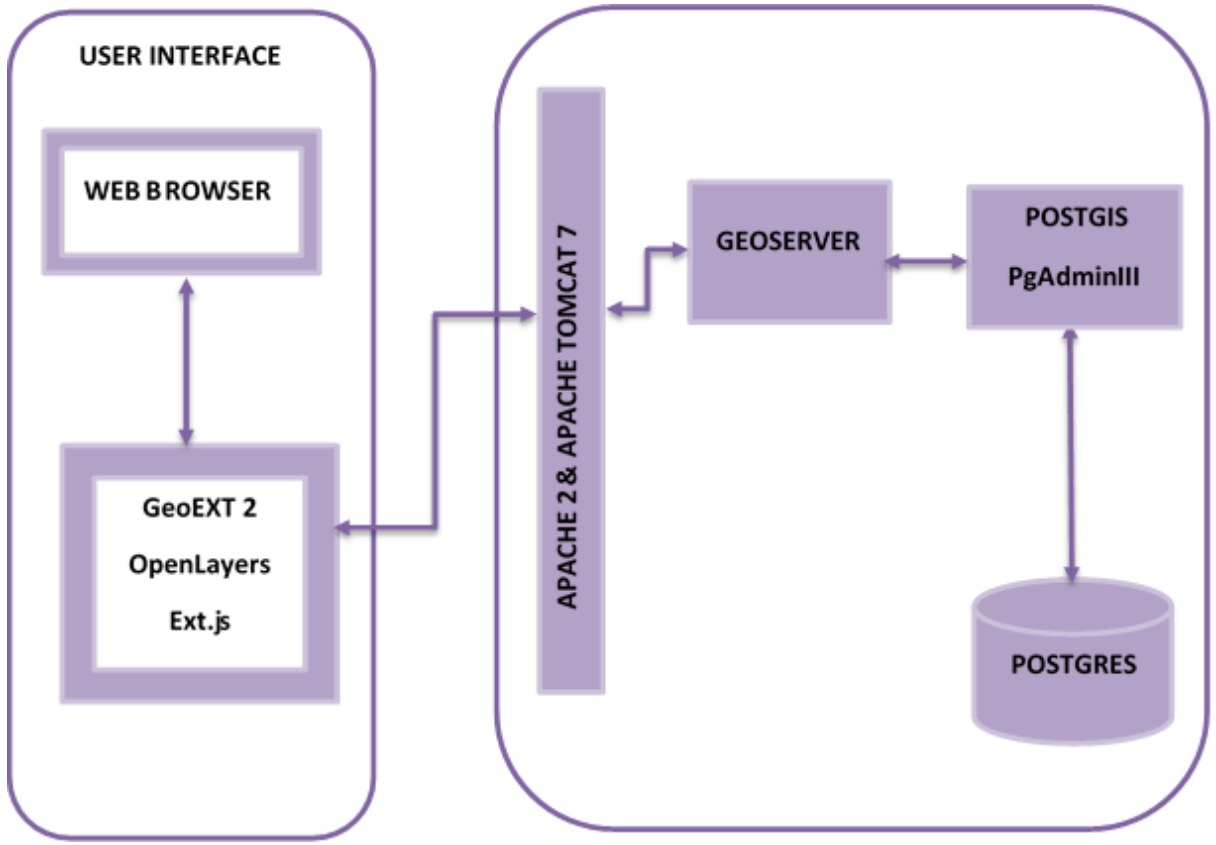
DEMONSTRATING SDI TO STAKEHOLDERS

- Prototype demonstrated a lightweight and flexible SDI data access protocol that can be accessed using mobile devices as well as computers.
- Deployed as a prototype demonstrator to unbundle the concept of SDIs to stakeholders, as well as communicate its benefits and feasibility in practice.
- Prototype was deployed as a distributed, scalable system using web services and open source software; Geoserver, GeoExt, Ext.js, OpenLayers, Postgres, PostgreSQL, AWS, Apache2, Apache Tomcat7, QGIS, PostGIS, PgAdminIII
- Prototype was evaluated by stakeholders using the Prototype Performance and User evaluation (PPU) deployed as part of the research.
- The prototype was attested as an effective demonstration of accessing spatial data from an SDI.

REQUIRED COMPONENTS

Requirements

- UML tool
- Map client
- Geospatial data server
- Data repository
- Java application server



System was deployed in Amazon web service (AWS) Ubuntu 14.04 virtual machine.

ASSEMBLY OF COMPONENTS

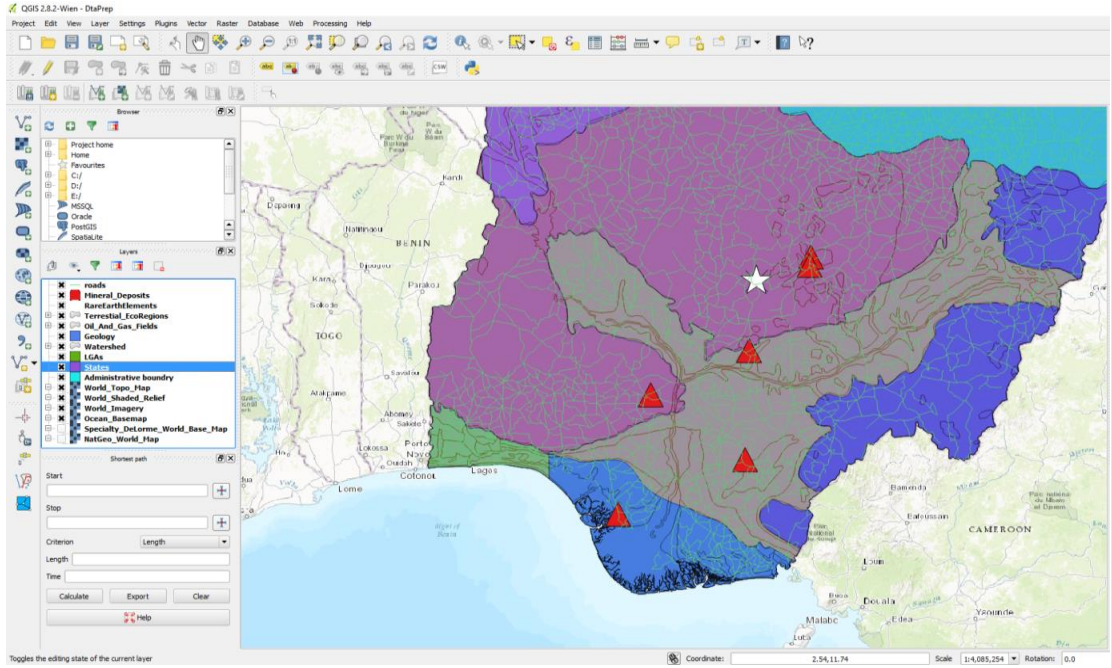
- ArgoUML and Eclipse EE the design of the prototype structure and the implementation of the back end.
  - The classes and activity diagrams were created in ArgoUML were imported into Eclipse EE for the implementation of the Java perspectives for the web application development.

- QGIS desktop client was used to prepare, create and update spatial data.

ASSEMBLY OF COMPONENTS

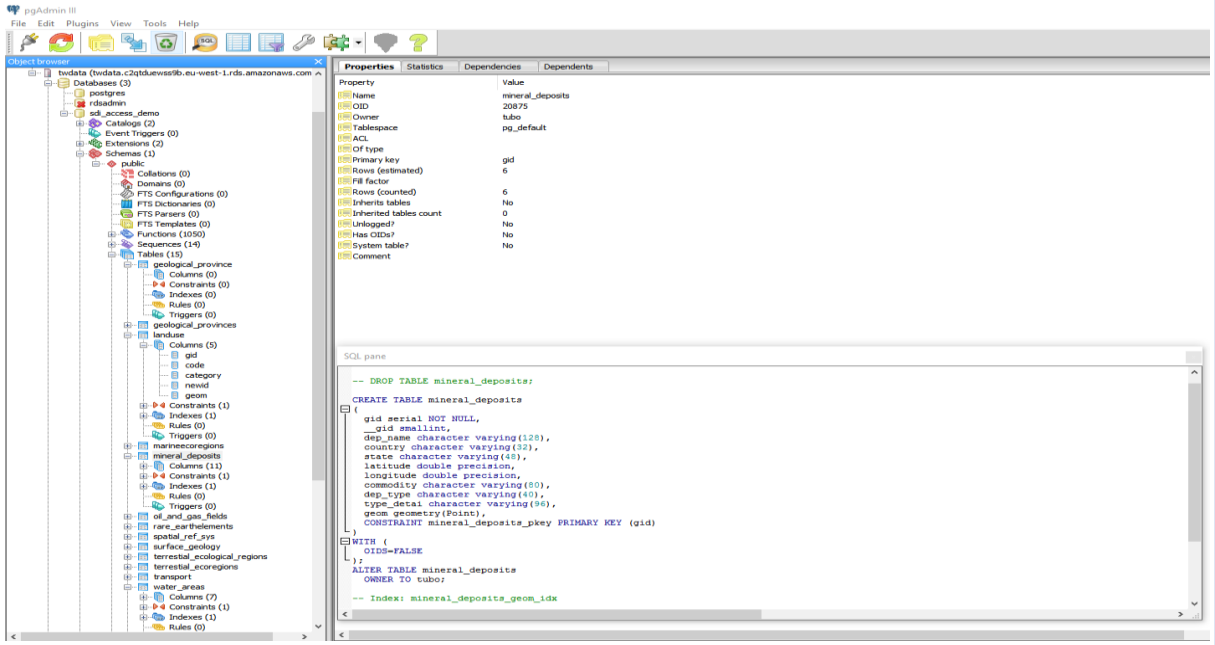
- Classes and activity diagrams were created in ArgoUML and Eclipse EE

- Data was prepared and cleaned up in QGIS to mitigate known issues with free and open source datasets
  - Completeness , validity, consistency, accuracy, and reusability .

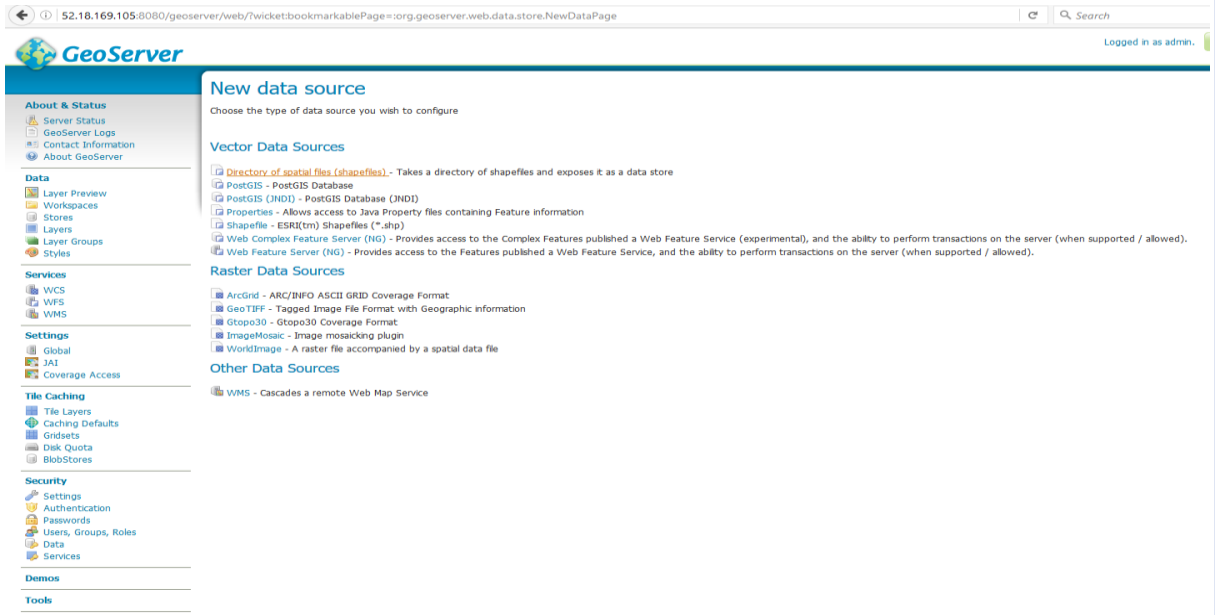


- Updated datasets were then imported into the PostgreSQL server created in the AWS RDF instance.

- The data is imported into PostgreSQL with their corresponding spatial layers and tables using the PostGIS extension.



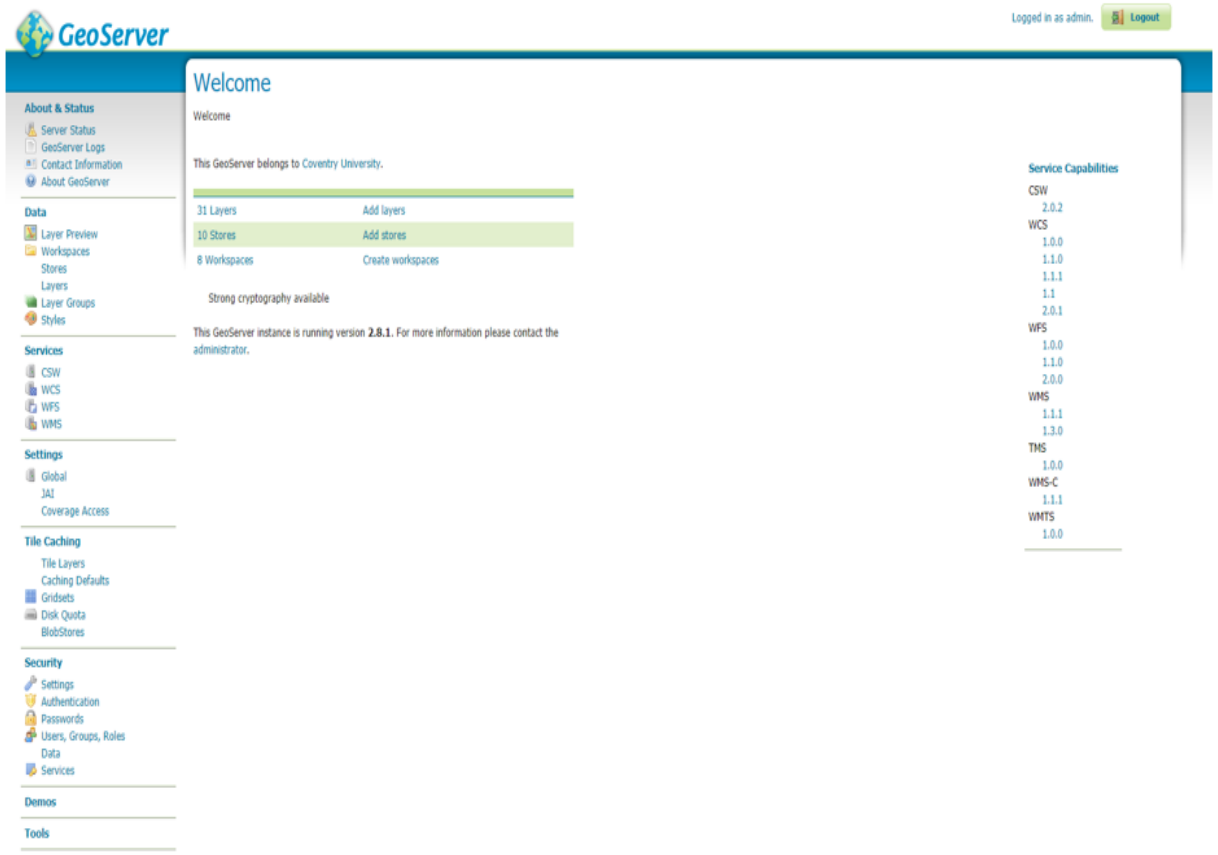
- The data is maintained in the PostgreSQL instance within the AWS relational data base service (RDS) and harvested into the Geoserver web service.



- Allows the allocation of user roles.
- Allows the seamless flow and combination of data from multiple sources

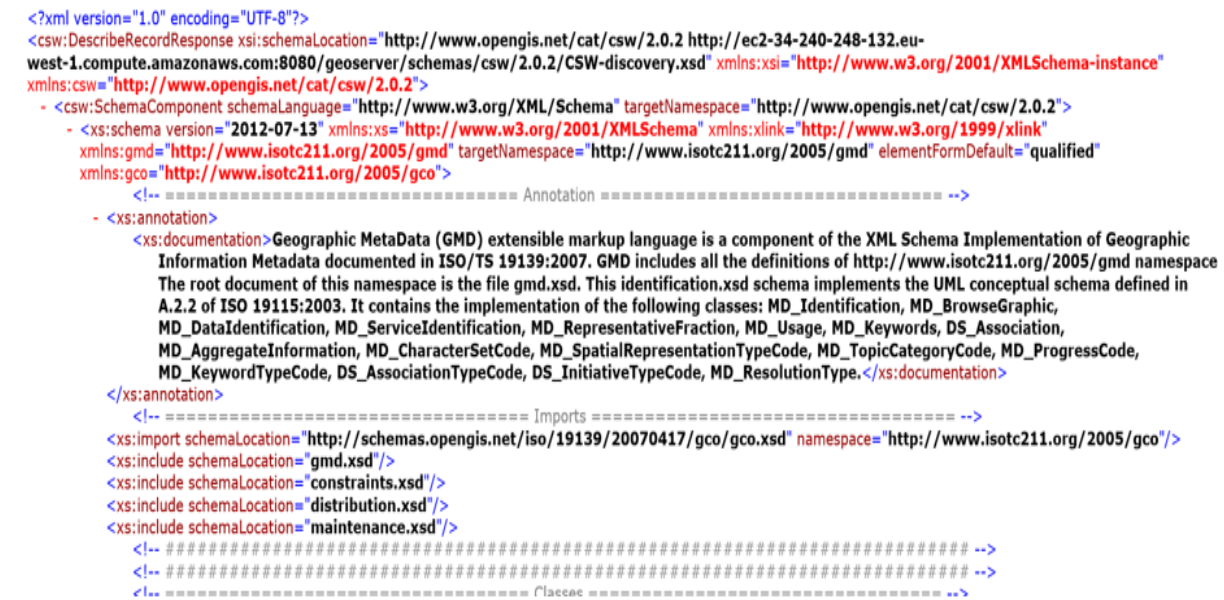
- Adapted to implement ISO19115 metadata standard by configuring the Catalog Service for the Web (CSW) OGC service into the Geoserver data directory using the ISO Metadata Profile Mapping file MD_Metadata

- A Geonetwork is however recommended for more extensive studies

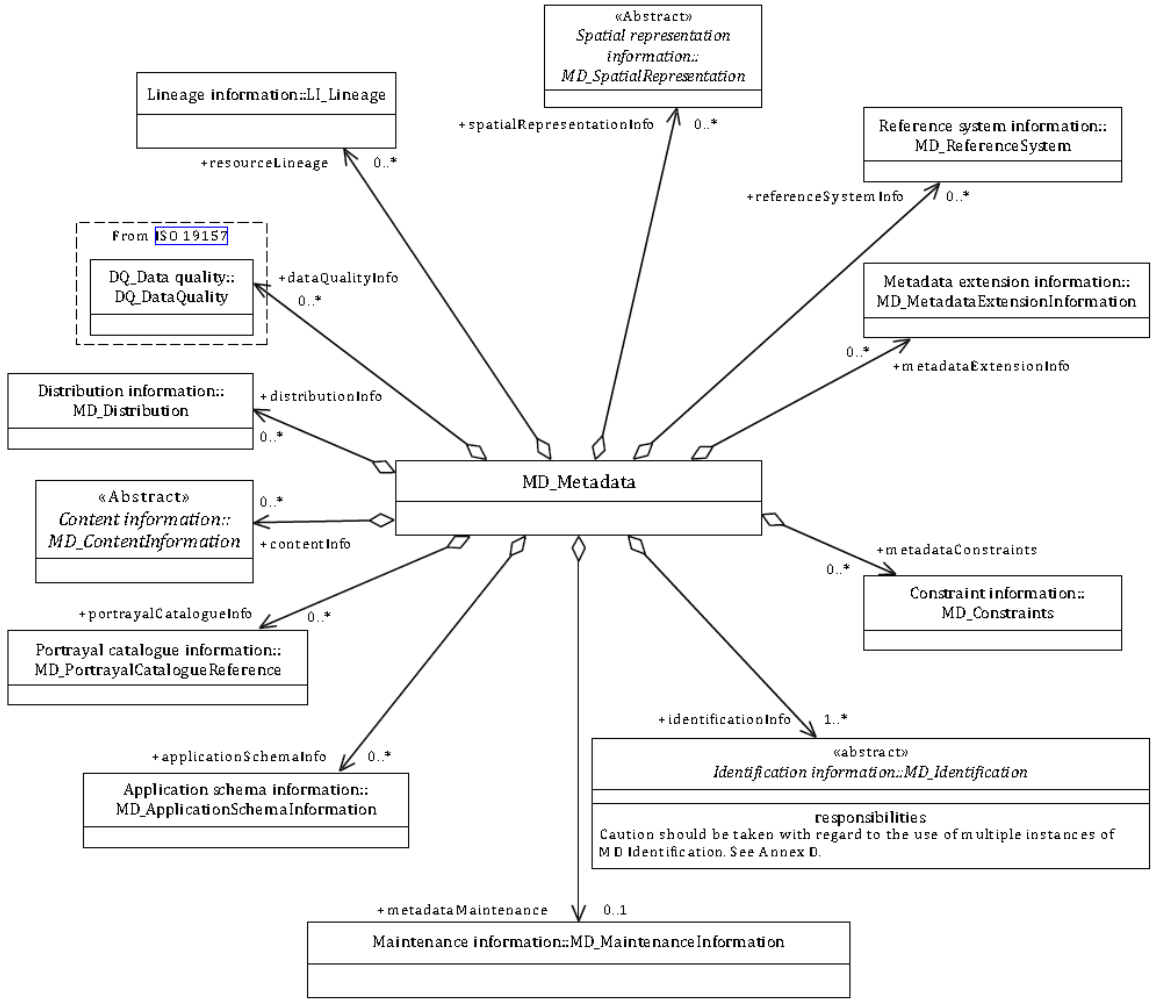


- The Geoserver implements the ISO 19115 Geographic Information - Metadata standard, prepared by ISO Technical Committee 211 (ISO TC211)

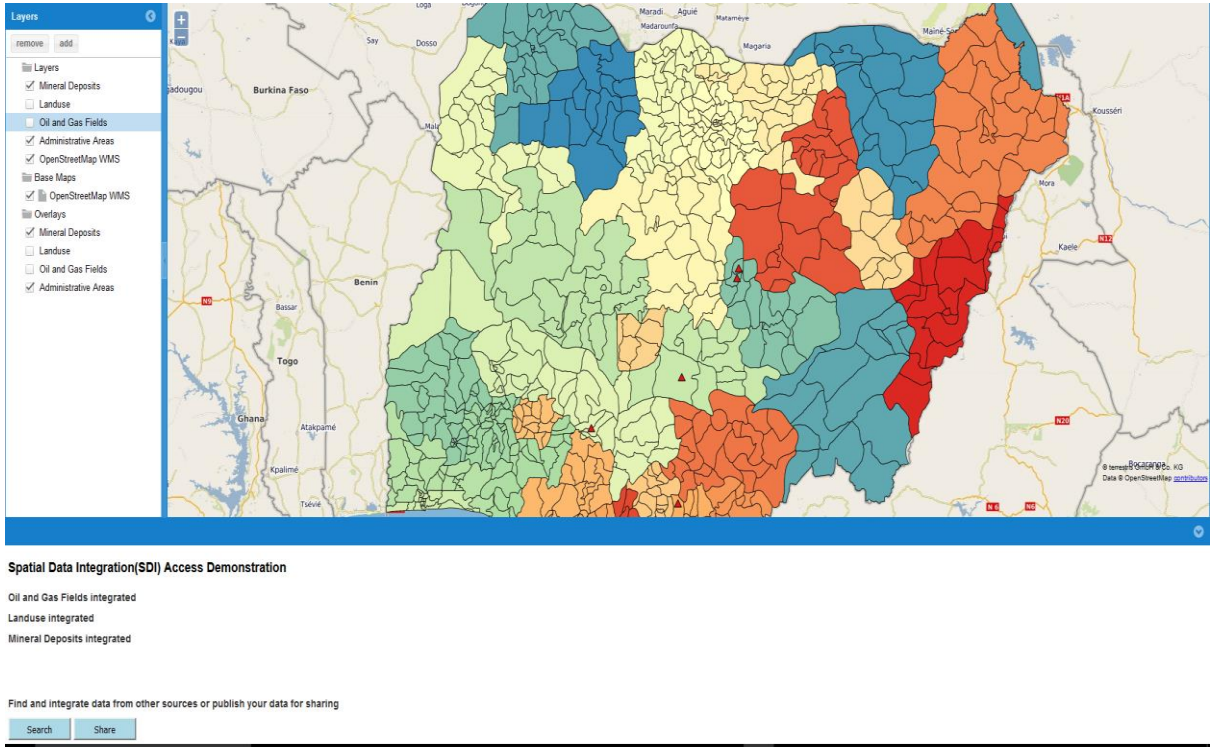
- Elements are encoded using XML (Extensible Mark-up Language) Meta Language and they are structured to adhere to the relevant schemas and structures.



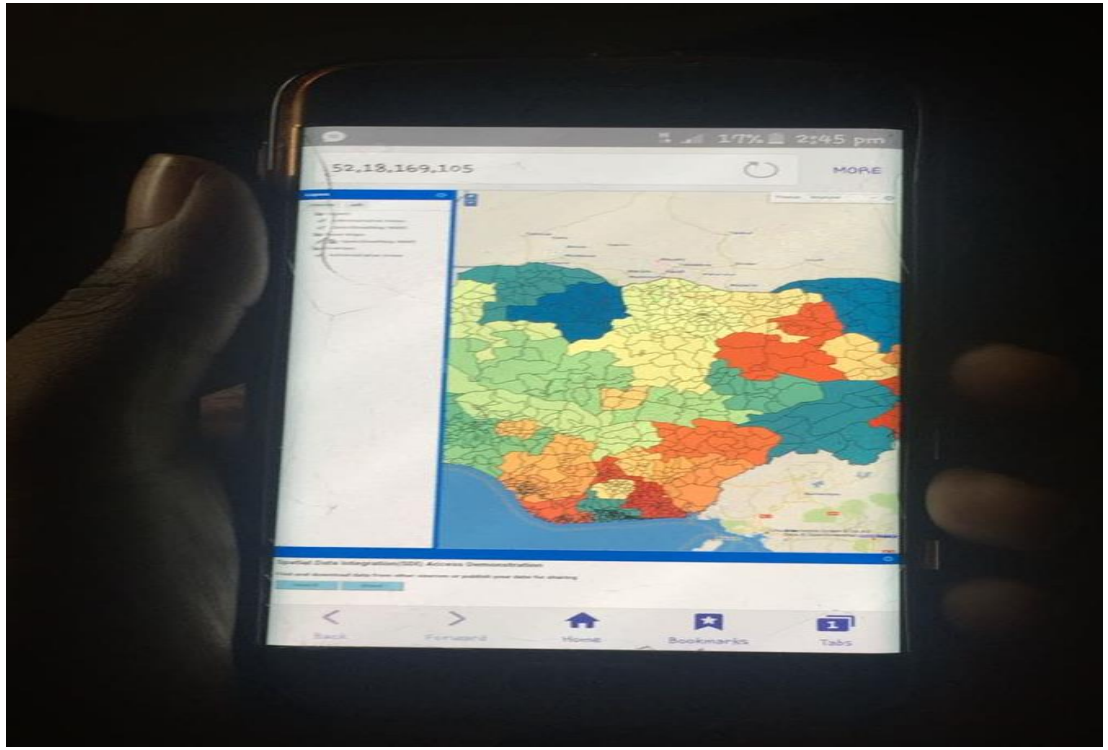
ISO19115 SCHEMA SHOWING THE MD_METADATA CLASSES



USER INTERFACE



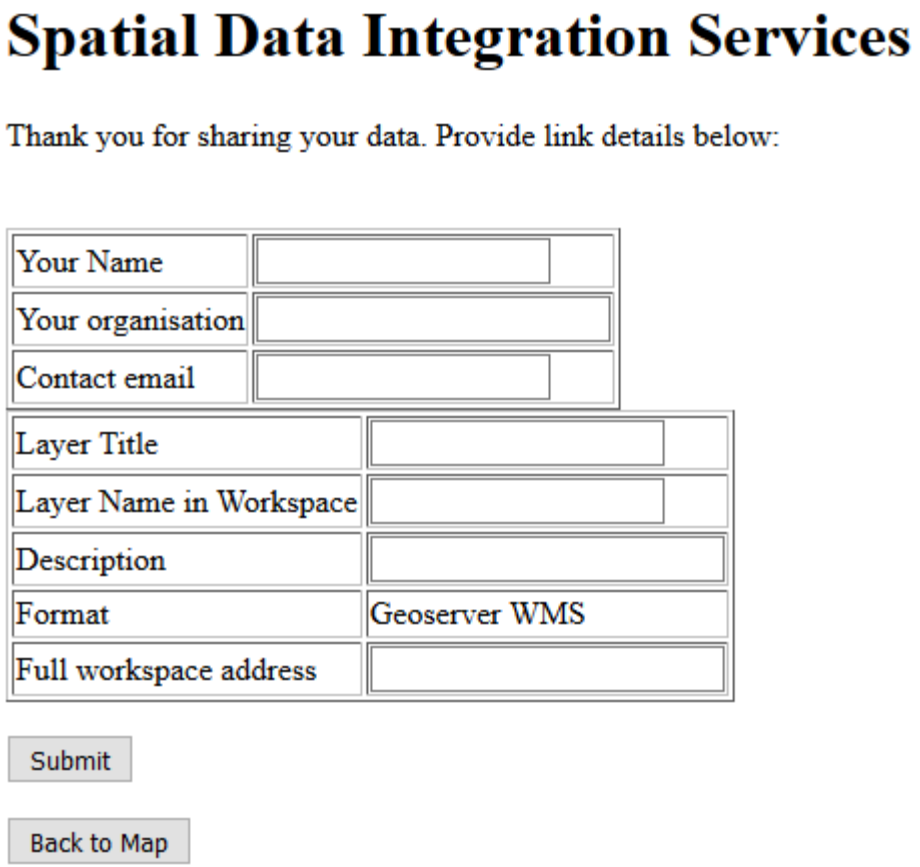
Web application accessed via a desktop



Web application accessed via a mobile phone

- GeoEXT, Ext.js and OpenLayers was used create a map window to access data stored in the Geoserver and postgres server

- Allows for spatial data integration from other sources.
- Demonstration was limited to Geoserver instances but is not limited to the Geoservers only.



Spatial Data Integration Services

Thank you for sharing your data. Provide link details below:

Your Name	
Your organisation	
Contact email	
Layer Title	
Layer Name in Workspace	
Description	
Format	Geoserver WMS
Full workspace address	

Submit

Back to Map

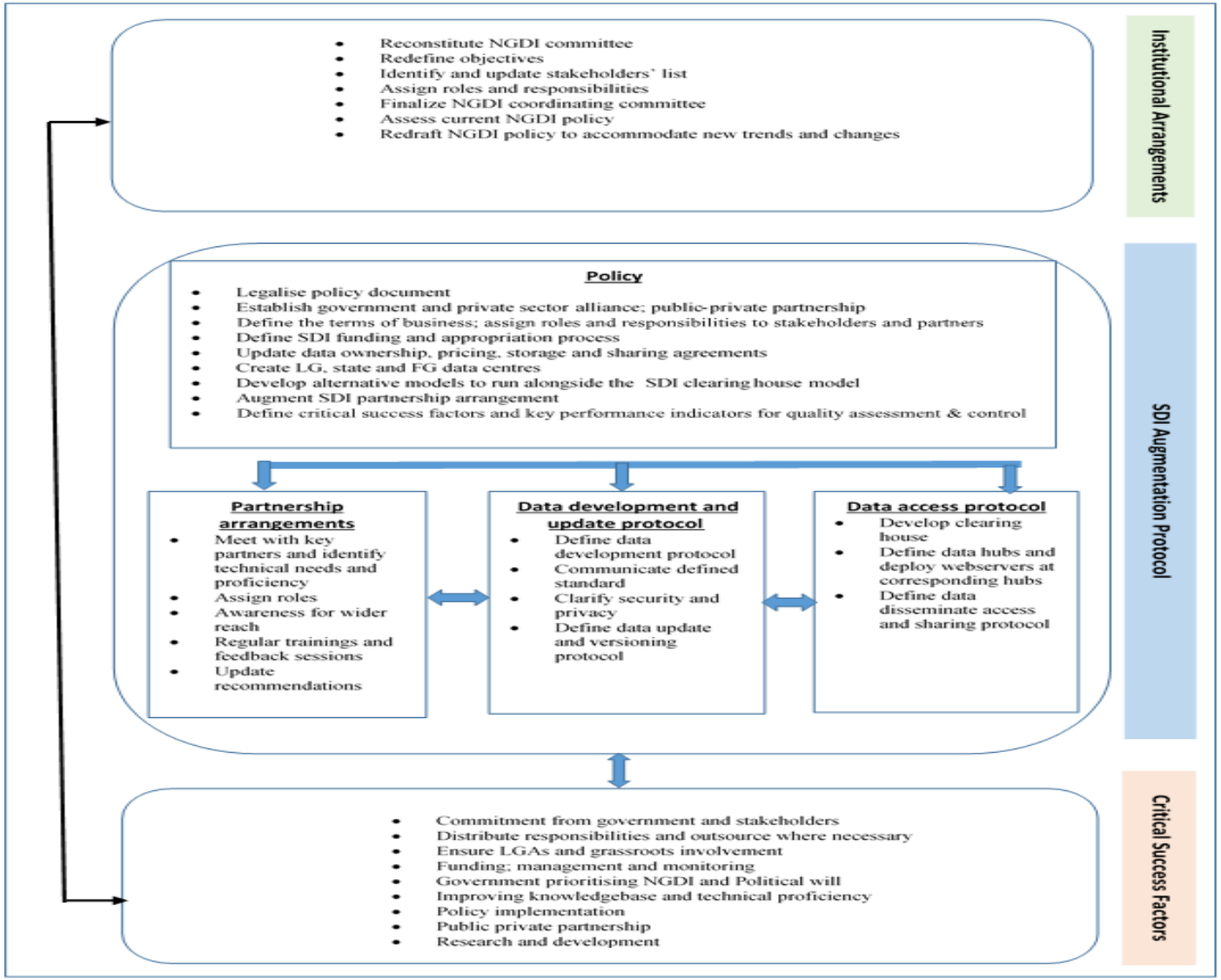
IMPLEMENTATION IN PRACTICE

- Prototype was assessed by industry experts following a demonstration and some evaluation tasks.
- Prototype was asserted a sufficient demonstration of SDI access protocol.
- However, a technological protocol in itself cannot address the problem.
- A framework for to incorporate technological protocol and current efforts was important.

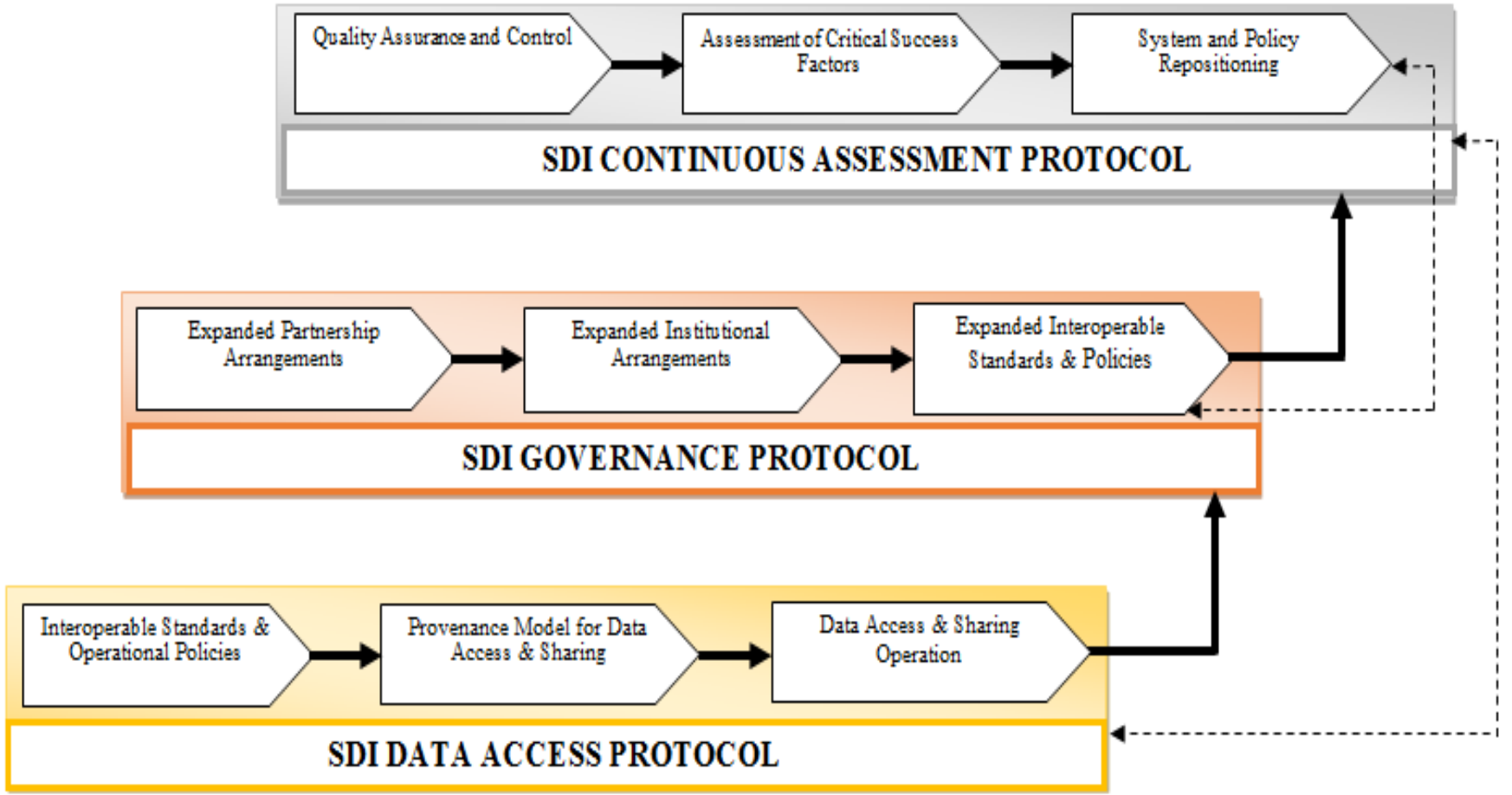
DEVELOPING AN AUGUMENTATION FRAMEWORK

- We included a provenance model following recommendations for a versioning system to improve the current metadata documentation.
- It was initially developed following the traditional top-down approach with three sections; Institutional Arrangements, SDI Augmentation Protocol and Critical Success Factors
- On further analysis and the incorporation of a key finding of the factors critical to NSDI implementation in Nigeria, (integrating grassroots structures from bottom to top), the framework was redeployed following the bottom-up approach.

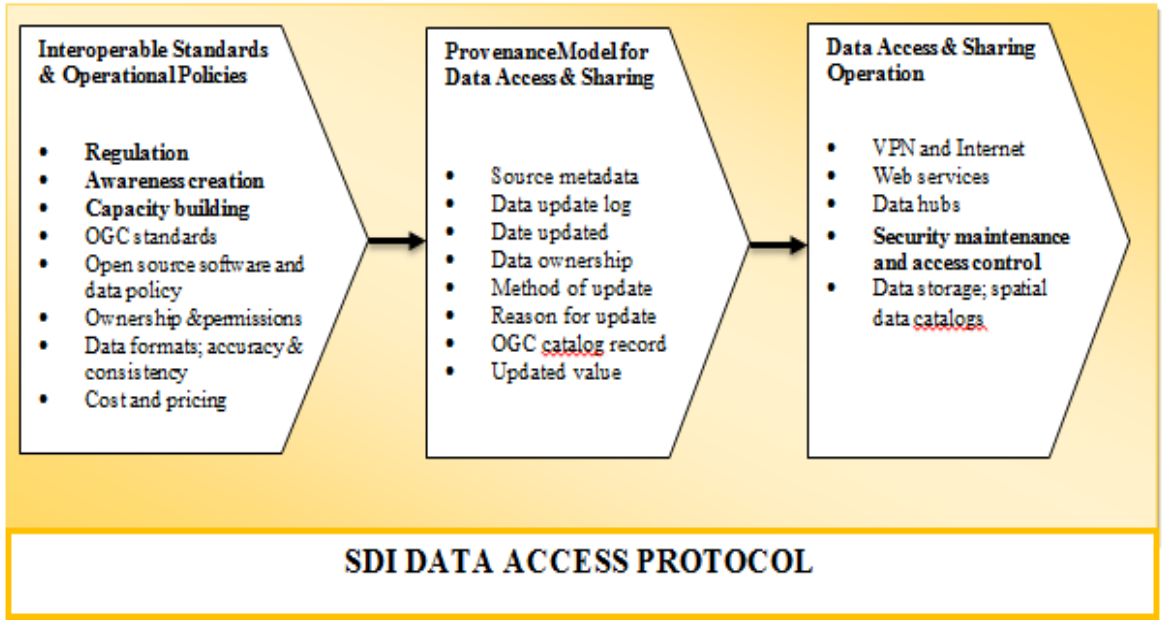
INITIAL ATTEMPT



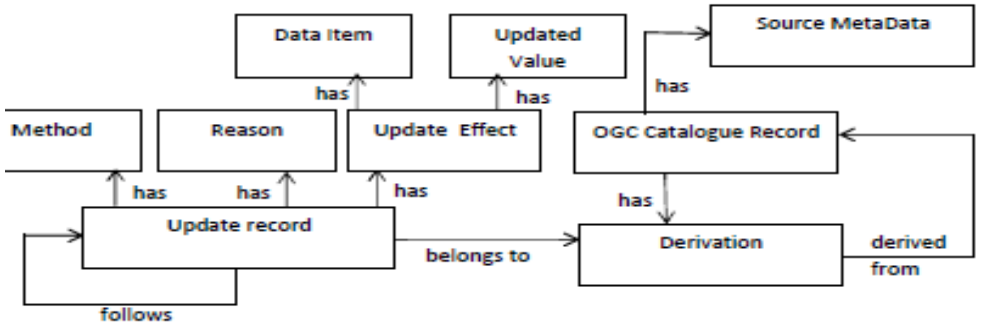
SDI AUGMENTATION FRAMEWORK



LEVEL ONE



Components of the SDI Data Access Protocol



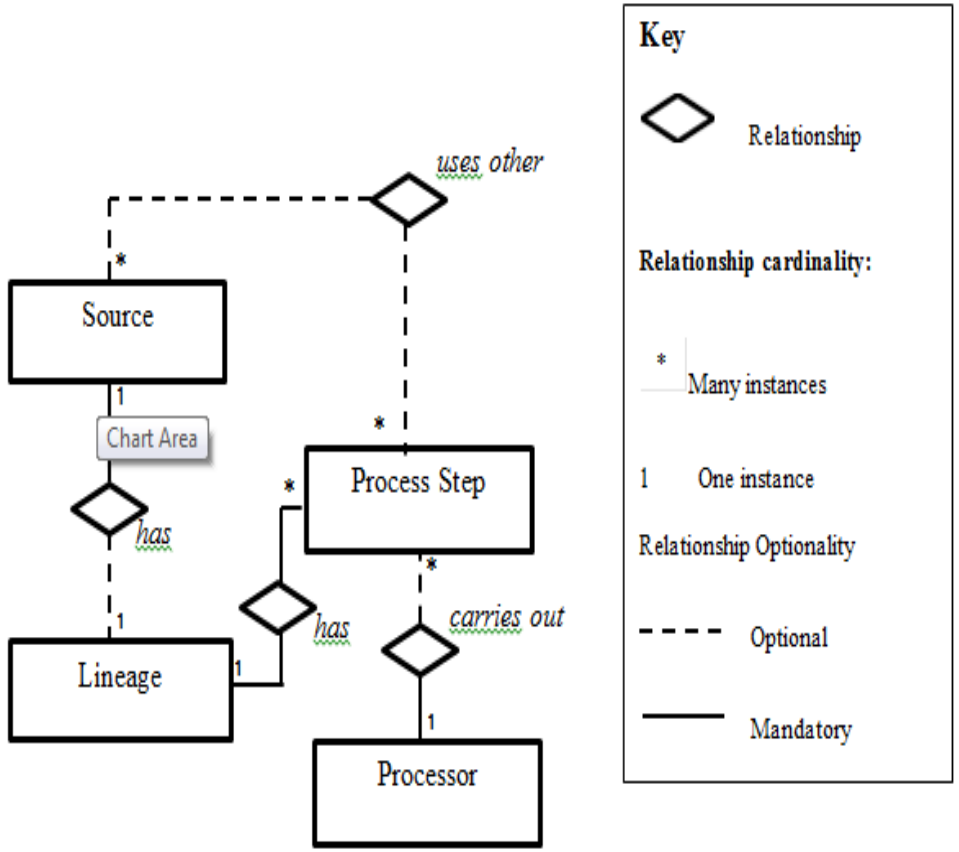
Provenance Model

- The supports data quality while ensuring the systems adheres strictly to interoperable standards and policies.
- Provenance model is flexible and can be adopted to varying levels of granularity to suit the user requirements.

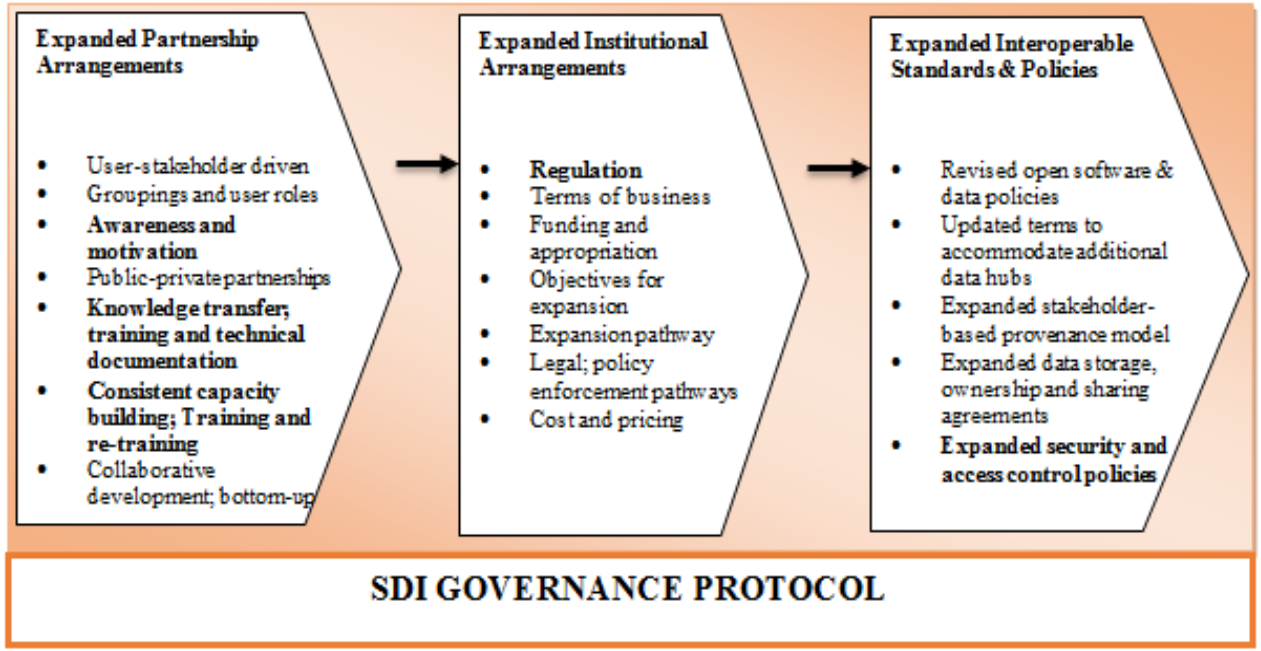
Concept	ISO19115	OPM	W3C Prov	SDI Access Protocol
Actor that manipulates the data	Processor	Agent	Agent	Processor
Data	Source	Artifact	Entity	Source
Manipulation of the data	Process Step	Process	Activity	Update record

- Considered the OPM, WC3-Prov and ISO19115 standards.

- Each source may have a Lineage record which will have at least one process step.
- Each process step belongs to a Lineage record and may use a number of sources.
- Process Step record will hold information about the process itself .



LEVEL TWO

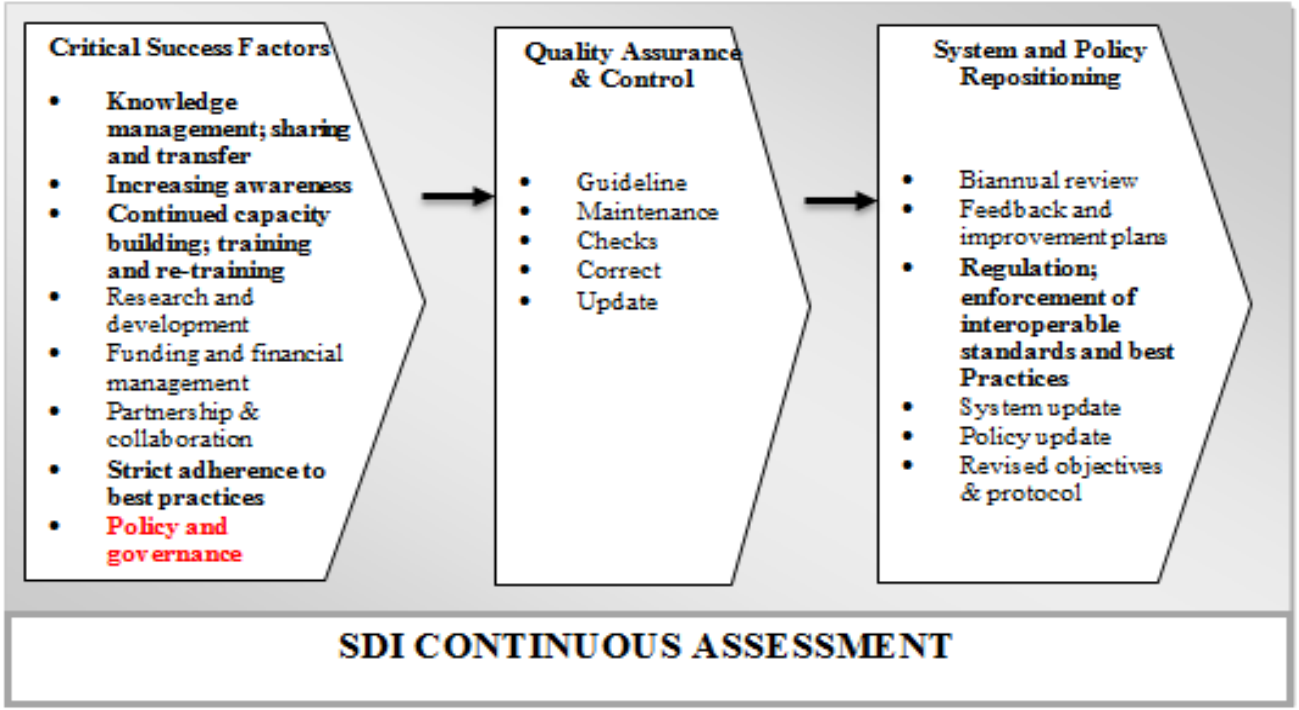


Components of the SDI Governance Protocol

Network Architecture of the Data Access Protocol Expansion

- At this level the SDI expands into multiple hubs that is harvested following agreed policies to create a 'clearinghouse'
- The internet is assumed as the underlying connection but VPNs can be established for applications requiring increased security

LEVEL THREE



- Overcomes a fundamental problem with implemented architectures which tend to fail due to poor monitoring, the alignment of architecture to changing needs of the users and the repositioning of the architecture to fit the up-to-date methods or technologies.

NEXT STEPS???

THANK YOU!

Credits:

- **This presentation contains aspects of the PhD research conducted by Dr. Tubolayefa Warekuromor.**
- **Research was supervised by Prof. Anne James and Dr. Babatunde Anifowose.**