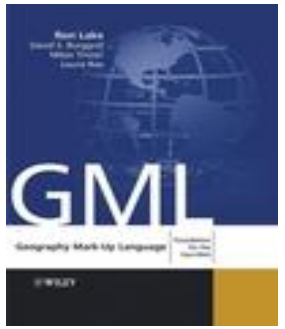
Integrated Information Systems - a Priority Evolving Platforms: Public - Private Partnerships

Steven Hagan
Vice President Engineering
Oracle Database Server Technologies
November, 2017

You Enhance Innovation & Sharing By Using **STANDARDS**

e.g. – The Spatial / Semantics/ Statistics Data Domains

- **ISO**
  - **TC 211; TC 204, 19115**
- **Open Geospatial Consortium**
  - **Simple Features; GML; Web Services**
- **De-facto Standards**
  - **SHP, MGE, DXF, KML**
- **Professional Standards**
  - **ISPRS, FIG, WMO,DDI, SDMX**
- **Java, .NET, Flash**
- **W3C: RDF,OWL, SPARQL, GeoSPARQL**
- **TAGGED METADATA – agree on tags**



SDMX

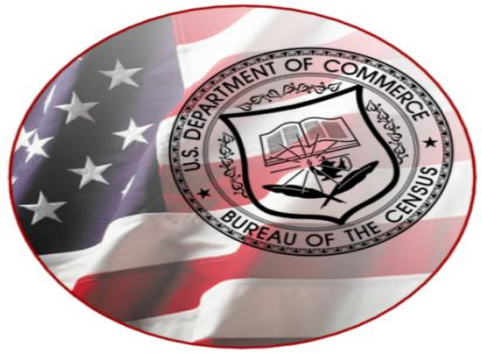
PLATFORM: HARDWARE / SOFTWARE TRENDS:

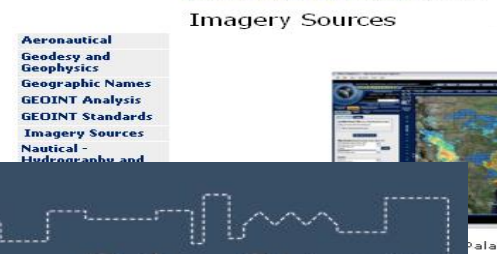
Hardware - **EVOLUTIONARY** – Moore's law still holding

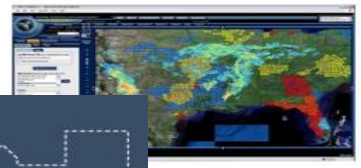
- New possibilities at Research Level – not yet proven – DNA, Quantum, Holography, Graphene …

- Software – **DISRUPTIVE** – **Parallelism** => clusters of **10,000+** computers, **CLOUD, ML, AI**
  - Sharding, Federated Databases
- Software: AVAILABLE NOW - Supporting all Data types in Databases
  - Databases/persistent stores: ***POLYGLOT PERSISTENCE*** now can handle ***ALL*** types of data
  - Software – ***GRAPH STORAGE, SEMANTICS, ONTOLOGIES, STATISTICS***
    - – Add all types of data, build *NEW* relationships
  - Enables ***MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE*** (ML, AI)
  - Stream data arriving; Filter the data; ML: Keep what matches your requirements; aggregate it, make it accessible for **ALL SEVENTEEN (17)** goals.
  - SECURITY – PRIVACY – Encryption improvements

Acquiring/Keeping Data for ***17*** Sustainable Development Goals: Need A Platform for **ALL** Variety, Velocity, Volume of Data







- VIDEO: UAVs, DRONES, SURVEILLANCE
- IMAGERY/Raster: (Satellites, Medical)
- Sensors (IOT), LIDAR, 3D, RFID, Wearables
- Social Media, Web Scraping, Mobile Phones
- New data products for: Land and Water mgmt, Agriculture, Environment Transportation, Terrain and City Models, SDIs for planning, maintenance, Emergency response, Defense, Intelligence, Consumers , Healthcare
- Genomics (DNA Sequencing)
- Semantics , Ontologies →
- Machine Learning, AI, Statistics
- **Location is a Powerful Organizing Principle**
- **MULTIPLE VERSIONS OF THE ABOVE**

PRIORITY ONE = ??

HOW FAST DO YOU WANT TO MEET SDG GOALS ?

PARTNERING WITH PRIVATE INDUSTRY!

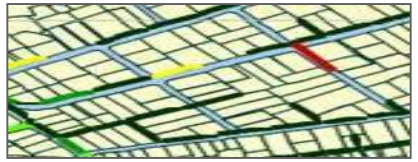
Success examples of Public Private Partnerships
You tell us what you want – we implement it in Products

- U.S. Census Taught Oracle Spatial how to do Planar Topology
  - Now shipping for years & improved
- Ireland Ordnance Survey – Needed Linked Open Data –
  - Now shipping & Improved
- Ordnance Survey – England, Ireland – needed Workspace Manager
  - multiple scenarios for what-if analyses or multiple editions of data for publication
- Many – needed Parallelism for Spatial Operators – use 1000s of Cores
- Kegg Japan – Explained need for and how to do RDF - Semantics

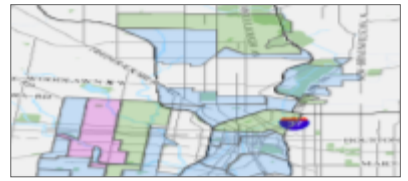
Managing All Spatial, Graph, Statistic Data – in One Store

Location and Statistics analysis with Secure, scalable storage for enterprise data

"Points"



"Lines"



"Polygons"

SQL access

Web Services (OGC)

SPARQL End Point

40+ Graph Analysis Functions (PGX)

Oracle Spatial and Graph

ORACLE DATABASE 12c

Spatial Vector Acceleration

Geocoding

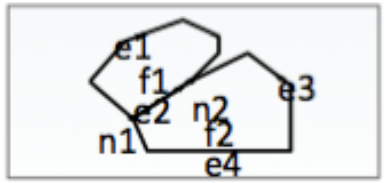
Routing

Inferencing

Property Graphs

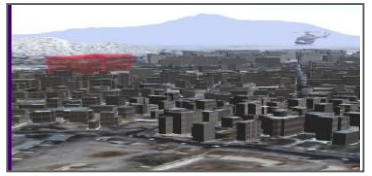
Network Graphs

RDF Semantic Graphs



Topologies

Rasters

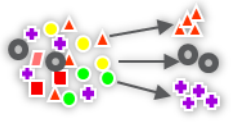


3D, point clouds (LiDAR)

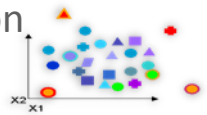
Oracle Statistics / Analytics Machine Learning Algorithms

- **Classification**
  - Logistic Regression
  - Decision Tree
  - Random Forest
  - Neural Network
  - Support Vector Machine
  - Naïve Bayes
  - Explicit Semantic Analysis
  - Gaussian Mixture Models

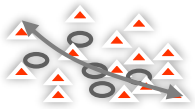
- **Clustering**
  - Hierarchical K-Means
  - Hierarchical O-Cluster
  - Expectation Maximization



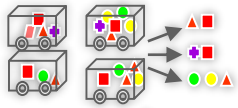
- **Anomaly Detection**
  - One-Class Support Vector Machine



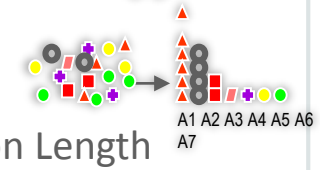
- **Regression**
  - Generalized Linear Model
  - Support Vector Machine
  - Random Forest
  - Linear Model
  - Stepwise Linear regression
  - LASSO



- **Association Rules**
  - A priori



- **Attribute Importance**
  - Minimum Description Length
  - Principal Component Analysis
  - Unsupervised Pairwise KL Divergence



- **SQL Predictive Queries**

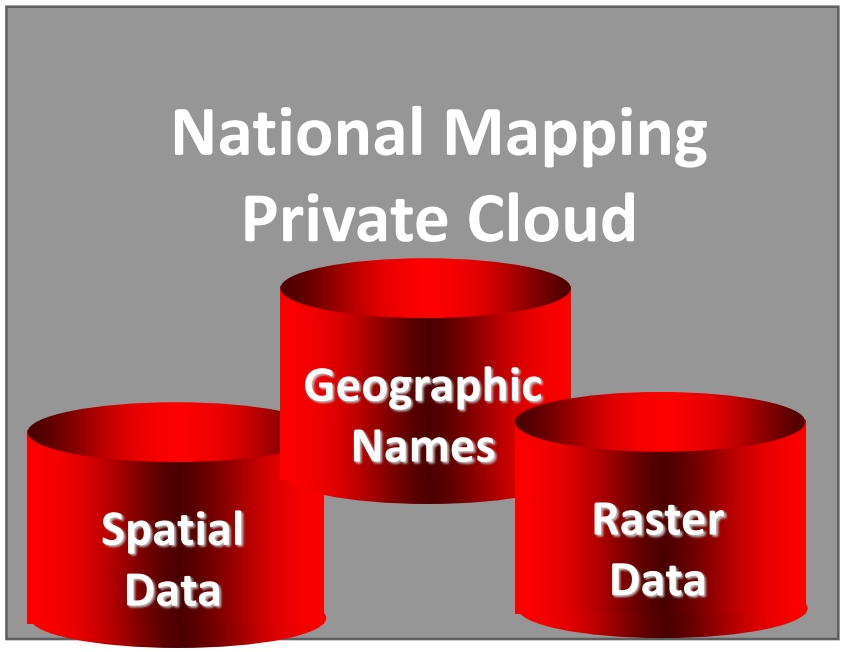
- **Statistical Functions**

- **Algorithm Text Support**
  - Algorithms support text type
  - Tokenization and theme extraction
  - Document similarity

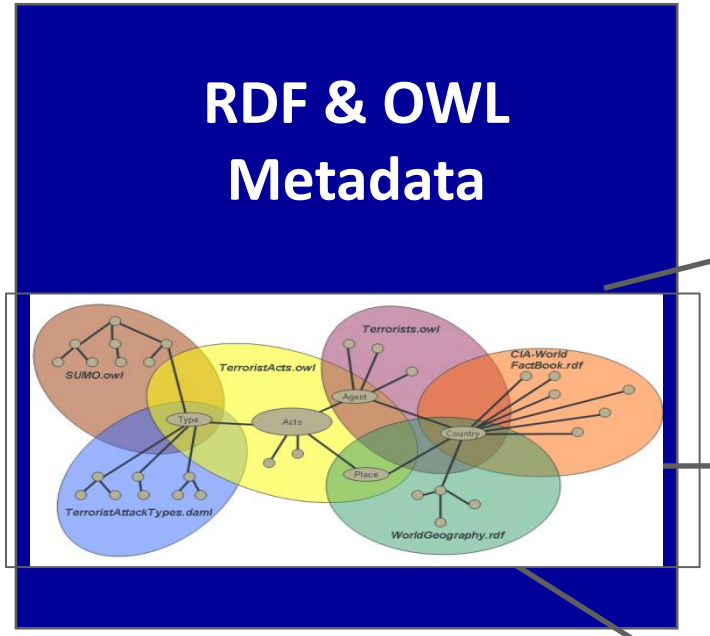
- **Feature Extraction**
  - Principal Component Analysis
  - Non-negative Matrix Factorization
  - Singular Value Decomposition
- **Time Series**
  - Single Exponential Smoothing
  - Double Exponential Smoothing
- **Open Source ML Algorithms**
  - CRAN R Algorithm Packages through Embedded R Execution
  - Spark MLlib algorithm integration

Meet Sustainable Goals: Repurposing Data: Ontology-driven Enable Shared, Actionable Knowledge

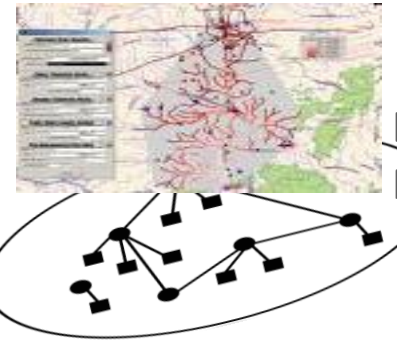
Application Ontologies



- Simple Features
- GeoRaster
- Topology
- Networks
- Gazetteers
- …



- Data Integration
- National Map schemas
- Geographic names
- Temporal
- Naïve Geography
- …



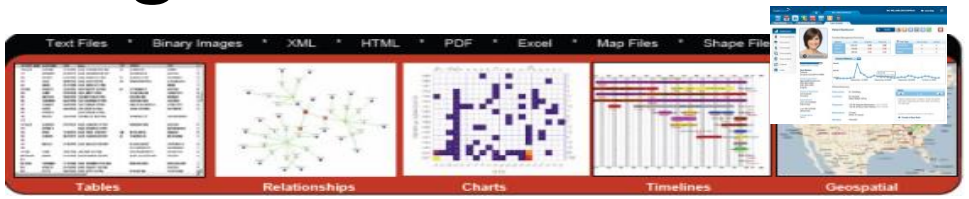
Environmental Monitoring

Disaster Recovery

Healthcare Biotech

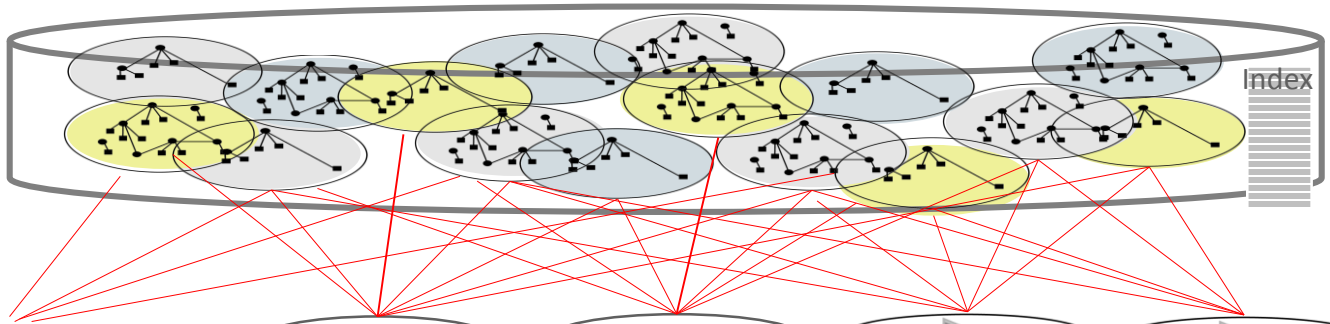
Harmonizing the Electronic Health Care Ecosystem – Goal 3 Using Semantics, Ontologies

Enterprise-wide, Patient-centric, longitudinal Record System



Domain Ontologies

(business metadata + Ontologies)

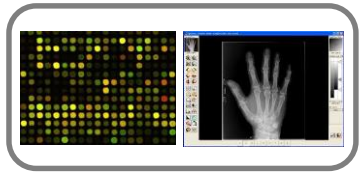


Data Servers

Lab/clinical Care | Research | Content Mgmt | Billings/Claims | Reporting/BI

Data Sources / Data Types

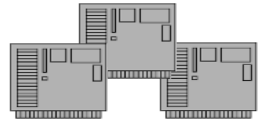
Social Media



Medical Devices

Lab Information Systems

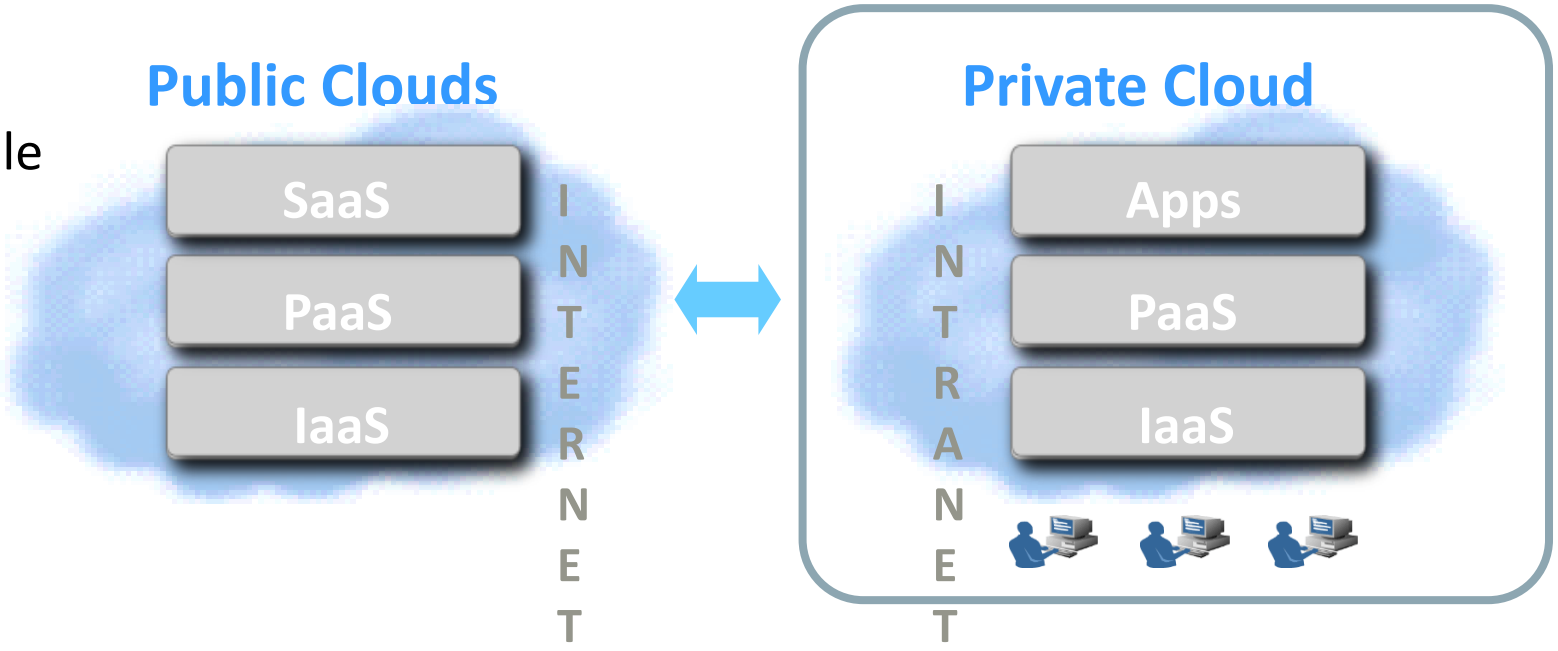
Subscription Services



Legacy Patient Records

Public Clouds, Private Clouds & WEB SERVICES

- Used by multiple tenants on a shared basis
- Hosted and managed by cloud service provider



- Exclusively used by a single organization
- Controlled and managed by in-house IT

Trade-offs

Lower *upfront* costs ⟷ Lower *total* costs

Outsourced management ⟷ Greater control over **SECURITY, COMPLIANCE, QOS**

OpEx ⟷ CapEx & OpEx

ELASTICITY is key value of Clouds

Oracle Technology Supplies both Public and Private clouds

YOU MAY NEED A CLOUD IN EACH COUNTRY ---DEPENDS ON THEIR LAWS

To Meet 2030 Goals: Do NOT Build Your Solutions From Scratch

Long Term Cost of Ownership rises with custom construction & Open Source

Time to Build

Optimizations

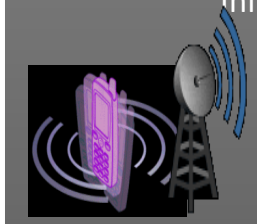
Maintenance

UN-GGIM: “train the individuals is at least five years”

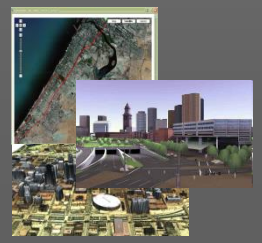
Sustainable Goals: All Data Types /Statistics/ ML / AI Bases: Success Enhanced with **MULTI-MODEL DATABASE PLATFORM**

Big & Fast Data

Volunteered Geographic Statistical Information



Sensors Streaming Data



Geo-referenced Video, 3D, LiDAR Satellites

Simplify Statistics IT

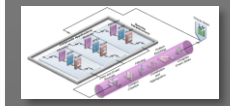
Support for Open Standards

Spatial Database, Application Server, BI, tools

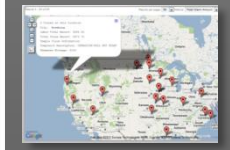
Support by Leading Partner solutions

Multi-Model Engineered Systems

Deep Analytics



Real-time Complex Event Processing



Dense Visualization

Spatial Analysis Graph Analytics

On Premise, On Cloud, Shared Services

Shared GeoSpatial Services
Location Aware Everything

Fully Parallel and SECURE