Geospatial Data and Processing in Apache Projects

George Percivall
CTO, Chief Engineer – Open Geospatial Consortium
percivall@apache.org

ApacheCon NA 2019

OGC®

Comprehensive global community-driven forward-looking expertise in location

Using location, we connect people, communities, technology and decision making to create a sustainable future for us, our kids and future generations

- *By specializing in making location more Findable, Accessible, Interoperable and Reusable*
- *Via a proven collaborative and agile process combining standards, innovation and partnerships*

Communities- Tech & Market Domains

Partnerships & Alliances

Process for Standards & Innovation

Open Source and Open Standards

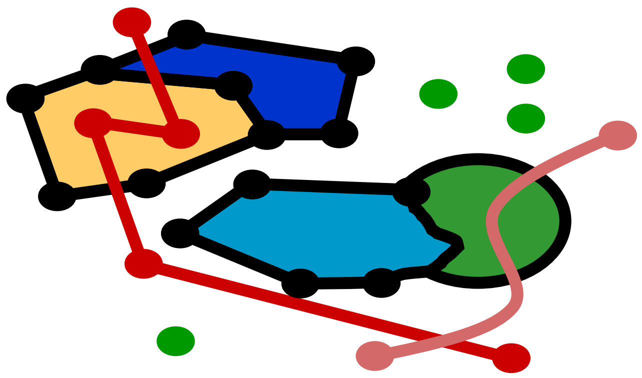
Open consensus standards as building blocks in Open Source

- –Increase code quality
- –Reduce development effort
- –Stable and proven APIs and encodings for interoperability
- –Interchangeability of software components

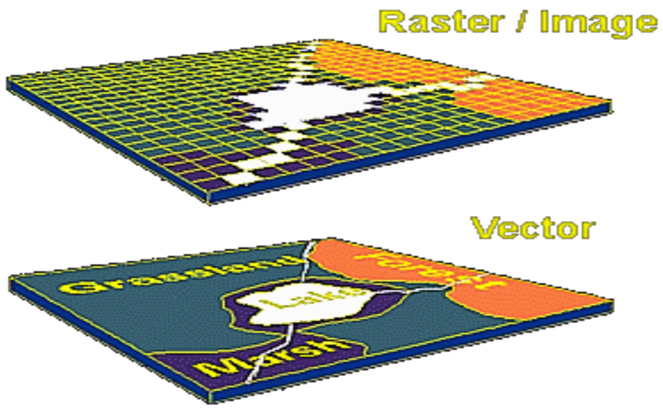
Spatial Things

"Anything with spatial extent; size, shape, position, e.g. people, places, objects"

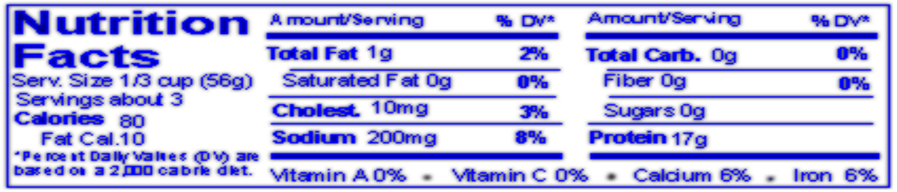
Feature Data



Coverage Data



Metadata



Maps

OGC®

Best Practices for Spatial Data

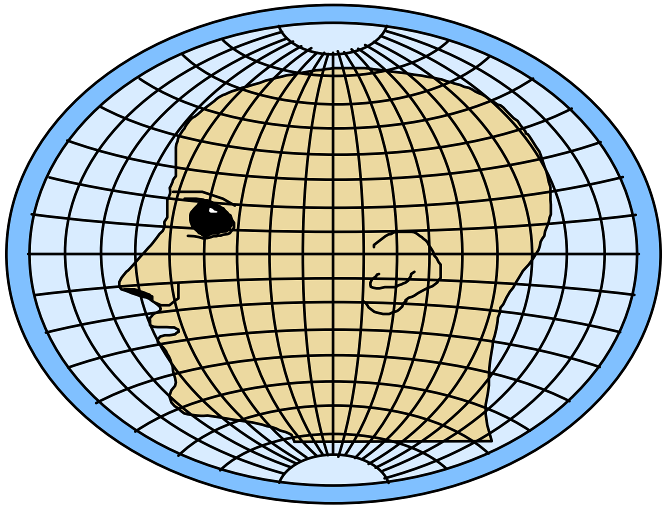
- Choose coordinate reference systems to suit your user's applications
- Provide geometries in a usable way: right level of accuracy, precision, size
- Use spatial data encodings that match your target audience
- Make data available through APIs
- Link resources together to create the Web of data
- Spatial analytics - discover meaningful patterns in spatial data

In the beginning…

All was simple, everyone agreed: the world was flat.

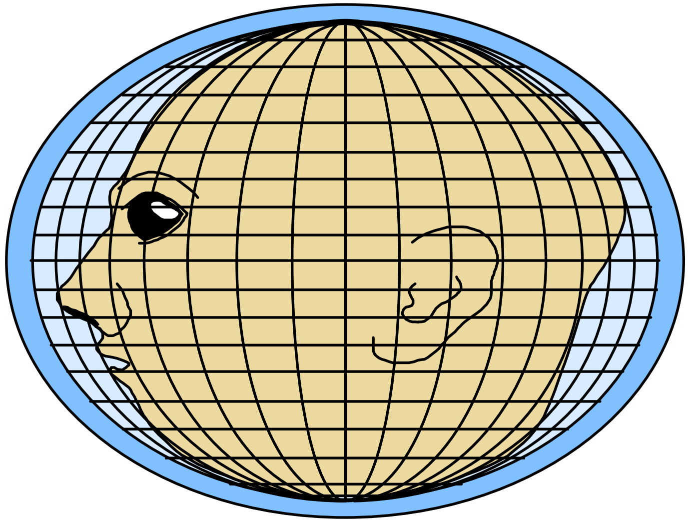
OGC®

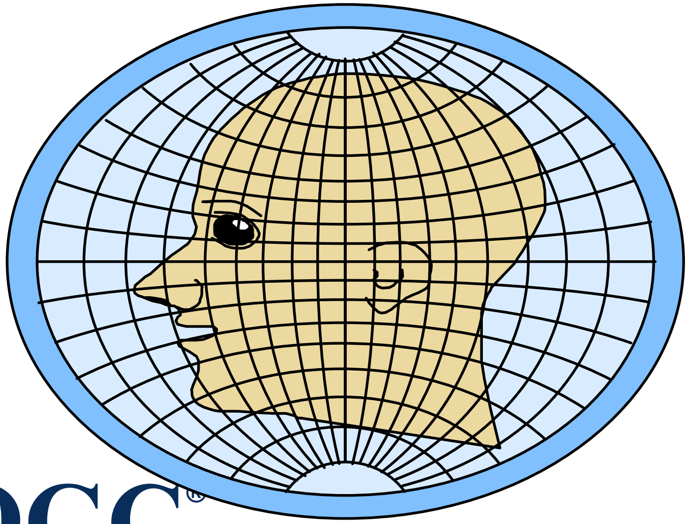
‘continent’ in different map projections



Globular projection

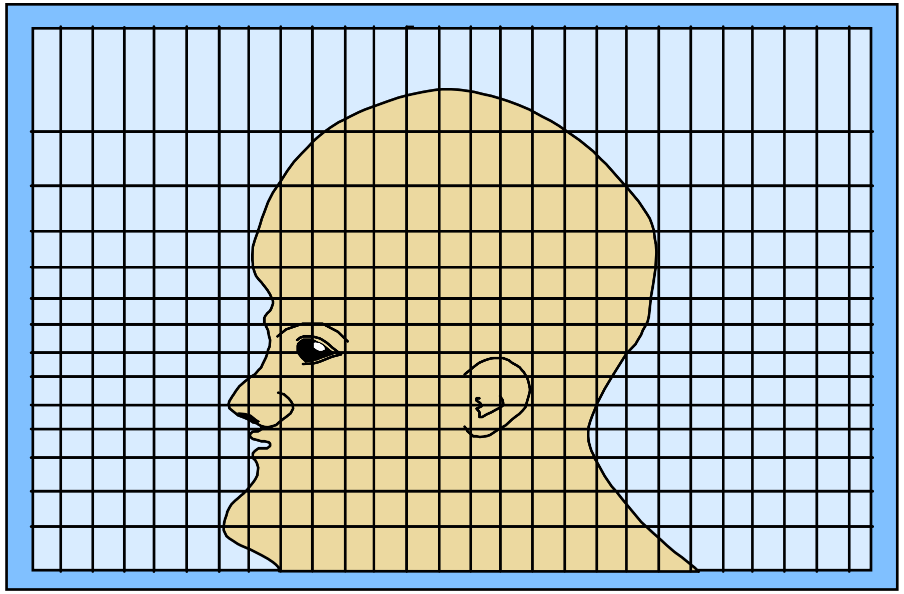
Orthographic projection



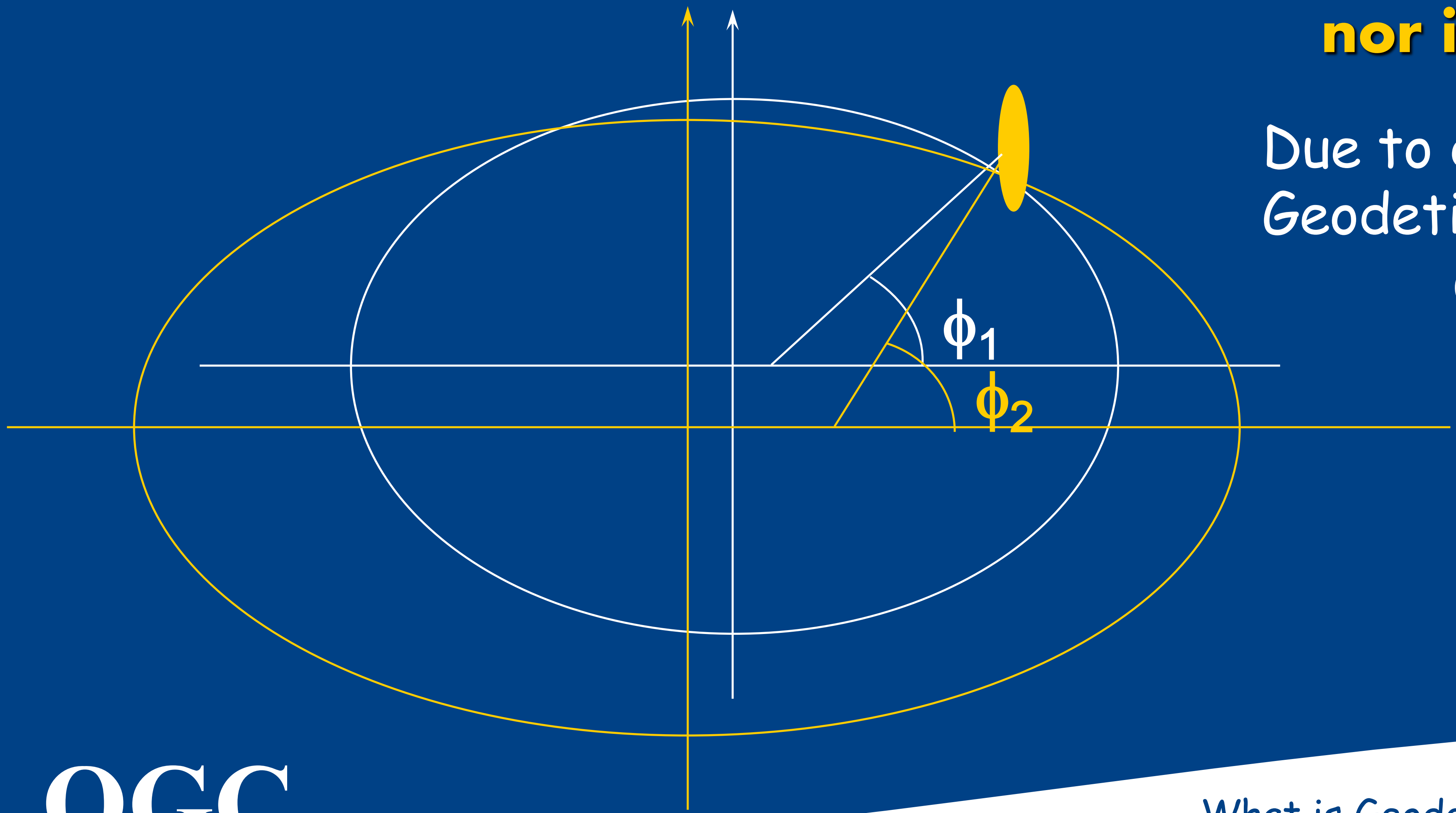


Stereographic projection

Mercator projection



Surprise: Latitude is not unique



nor is Longitude

Due to different Geodetic Datums:

$$\phi_1 \neq \phi_2$$

What errors can you expect?

- Wrong geodetic datum:
  - several hundreds of metres
- Incorrect ellipsoid:
  - horizontally: several tens of metres
  - height: not effected, or tens to several hundred metres
- Wrong map projection:
  - entirely the wrong projection:
    hundreds, even thousands of kilometres (at least easy to spot!)
  - partly wrong (i.e. one or more parameters are wrong):
    several metres to many hundreds of kilometres
- No geodetic metadata → coordinates cannot be interpreted
  - Datum; ellipsoid; prime meridian; map projection

OGC®

CRS WKT: Dynamic CRS, ellipsoidal 3D coordinate system

```
GEOGCRS["WGS 84 (G1762)",
  DYNAMIC[FRAMEEPOCH[2005.0]],
  TRF["World Geodetic System 1984 (G1762)",
    ELLIPSOID["WGS84",6378137,298.257223563,LENGTHUNIT["metre",1.0]]],

  CS[ellipsoidal,3],
    AXIS["(lat)",north,ANGLEUNIT["degree",0.0174532925199433]],
    AXIS["(lon)",east,ANGLEUNIT["degree",0.0174532925199433]],
    AXIS["ellipsoidal height (h)",up,LENGTHUNIT["metre",1.0]]

  ID["EPSG",4269],
  REMARK["1986 realisation"]
```

OGC®

Apache Spatial Information System

http://sis.apache.org

Reference system by EPSG code

```
CRS.forCode("EPSG::26747");            // NAD27 / California zone VII
```

```
WARNING: Code "EPSG:26747" is deprecated and superseded by 26799.
Reason is: Error in dependent projection record.
```

Apache SIS warns user if there is known issue with provided parameters

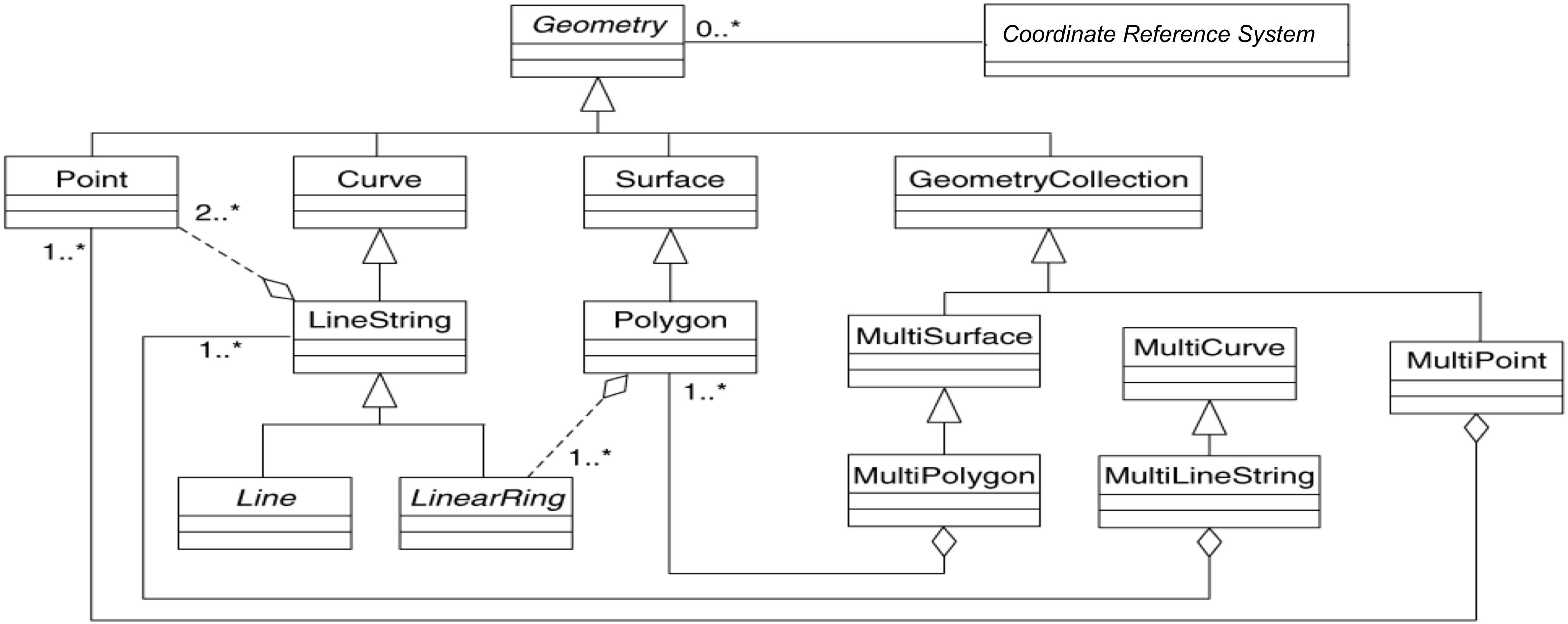
Coordinate operations

```
import org.opengis.referencing.operation.CoordinateOperation;
// Class declaration omitted for brevety

CoordinateReferenceSystem sourceCRS = // may be parsed from WKT, EPSG code, etc.
CoordinateReferenceSystem targetCRS = // idem
CoordinateOperation op = CRS.findOperation(sourceCRS, targetCRS, region);

// Verify domain of valididty and accuracy
System.out.println("Valid in " + CRS.getGeographicBoundingBox(op));
System.out.println("Accuracy " + CRS.getLinearAccuracy(op) + " m");
```

Simple Feature: Geometries

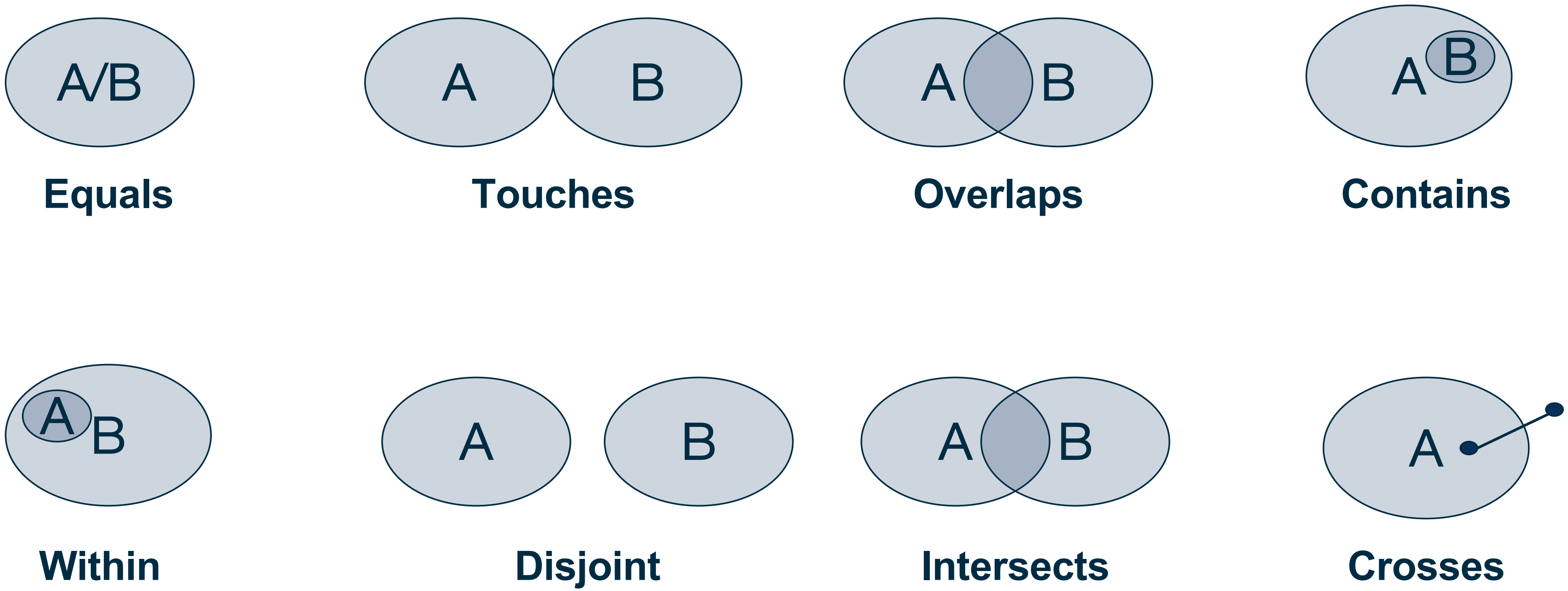


OGC simple features geometries restricted to 0, 1 and 2-dimensional geometric objects that exist in 2-dimensional coordinate space (R2)

OGC®

http://www.opengeospatial.org/standards/sfa

Topological Relations between Spatial Objects



Defined in OGC Simple Features Access Standard

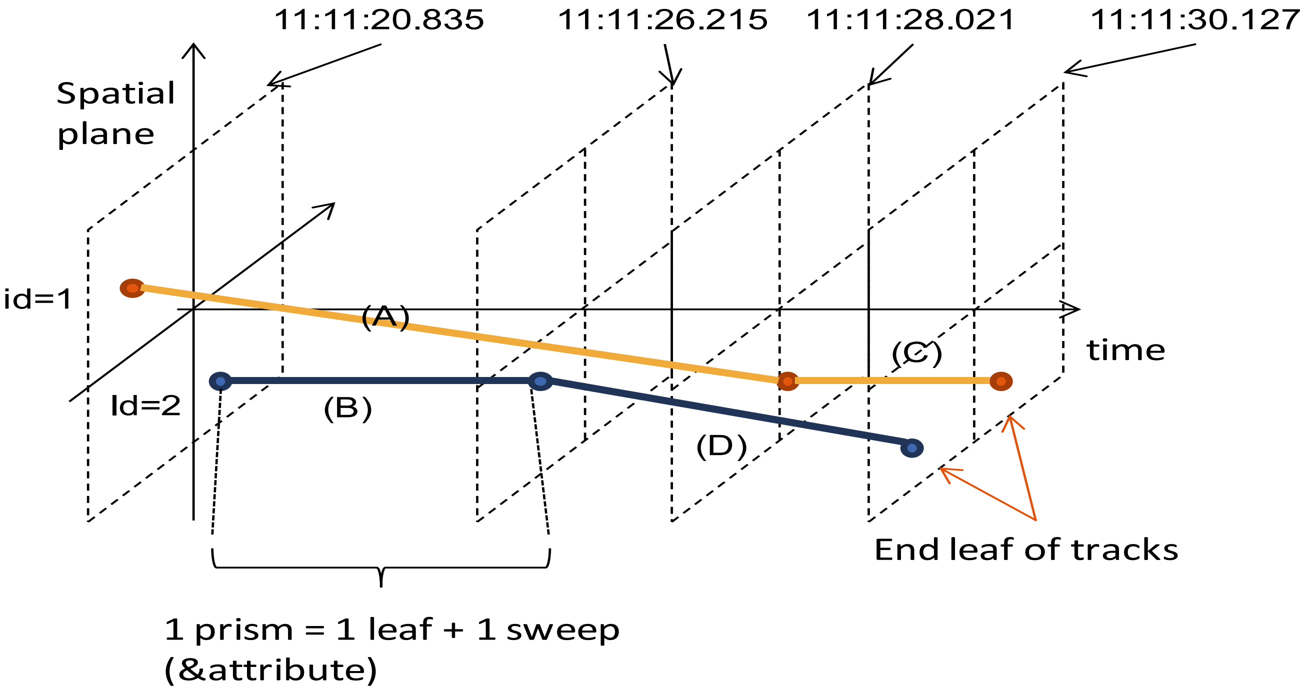
http://www.opengeospatial.org/standards/sfa

Also defined in GeoSPARQL

http://www.opengeospatial.org/standards/geosparql

OGC®

Spatial Temporal Geometry



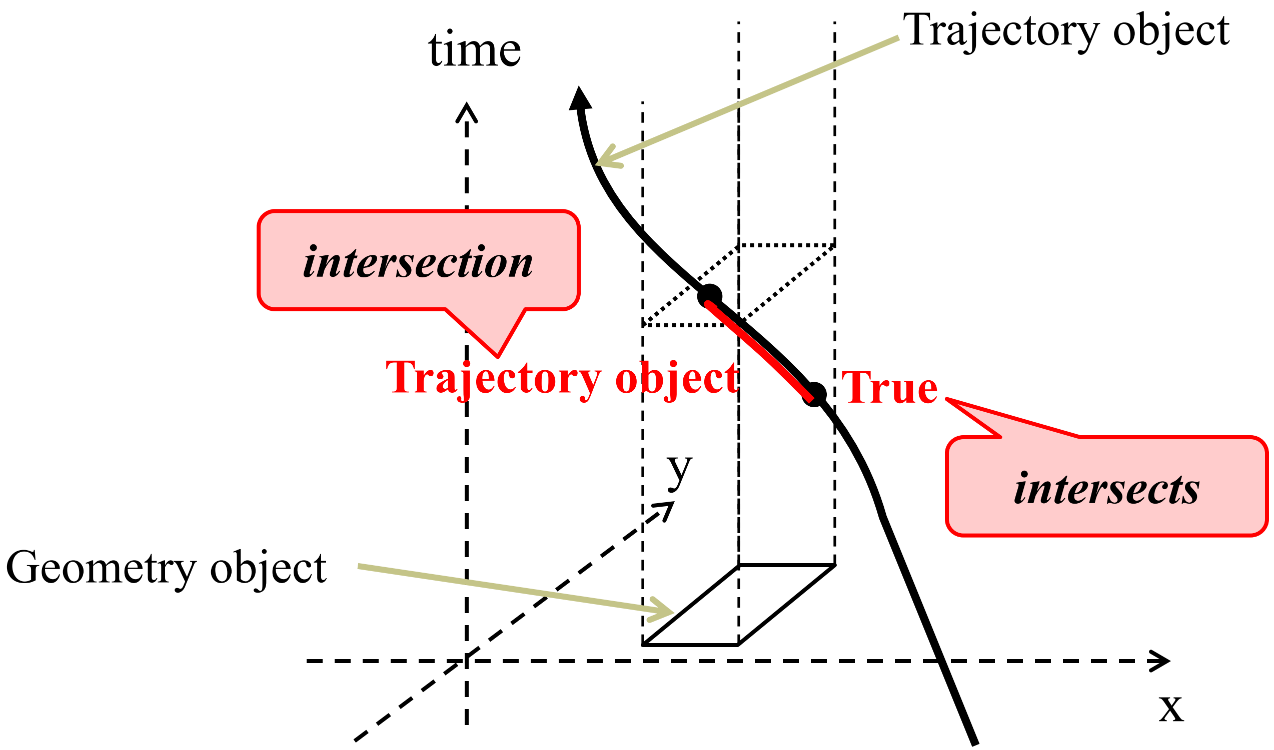
OGC Moving Features Standard

OGC®

Moving Feature Access: Trajectory and a Geometry

Operations between one trajectory object and one or more geometry objects

An example is “intersection” between a geometry object and a trajectory of a moving feature like a car, a person, a vessel, an aircraft, and a hurricane.

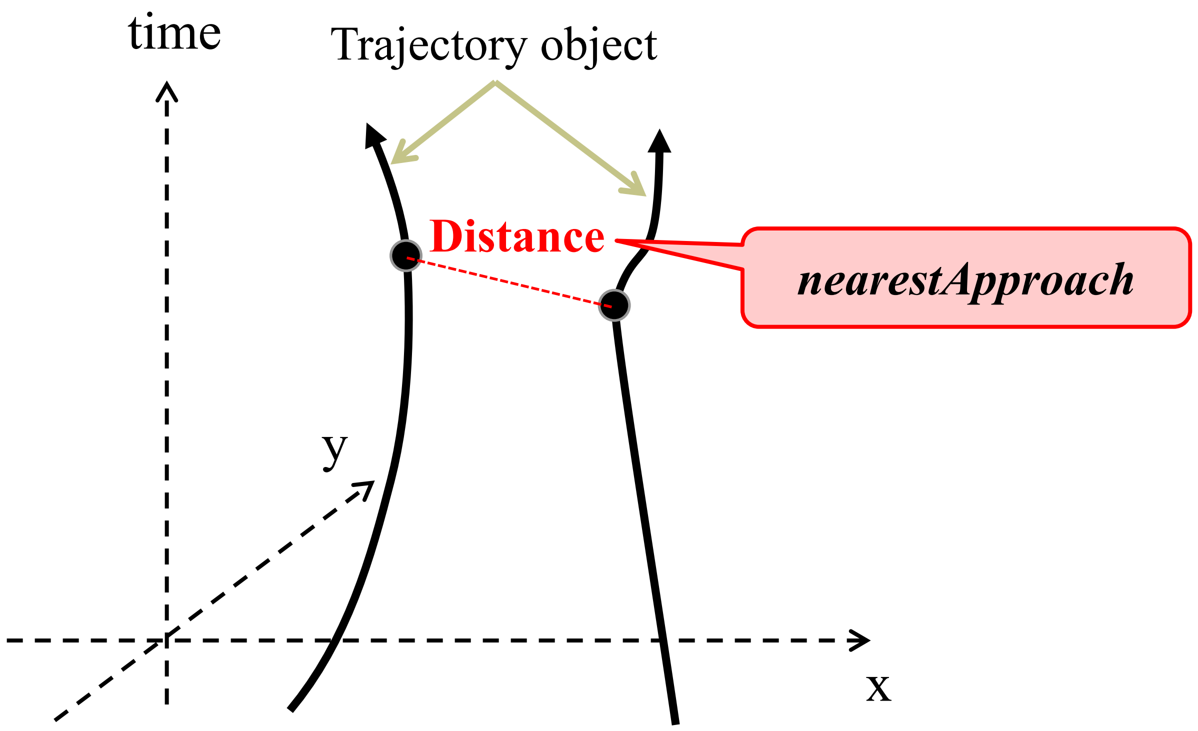


OGC®

Moving Feature Access: Two Trajectories

Operations between two trajectory objects

An example is to calculate a distance of the nearest approach of a trajectory to another trajectory.



Encodings

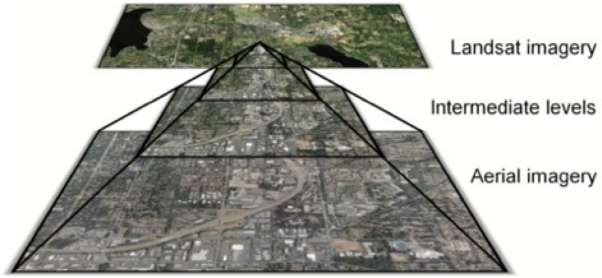
GeoPackage

- GeoPackage is a universal file format for geodata.
  – open, standards-based, application and platform independent, and self-describing.
  – Works on any desktop or mobile OS using SQLite
  – *Connected / limited / disconnected environment use*

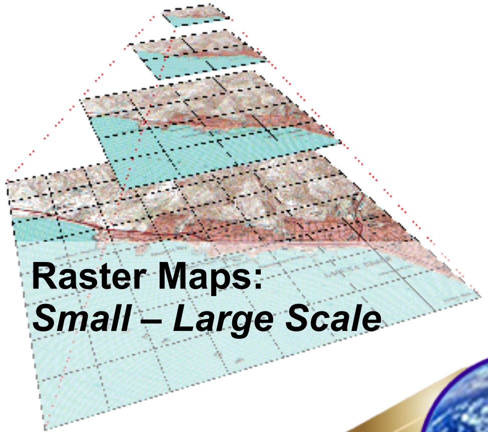
- GeoPackage - the modern alternative to formats like SDTS and vendor specific

- *Experience it here:* [*http://www.ogcnetwork.net/geopackage*](http://www.ogcnetwork.net/geopackage)

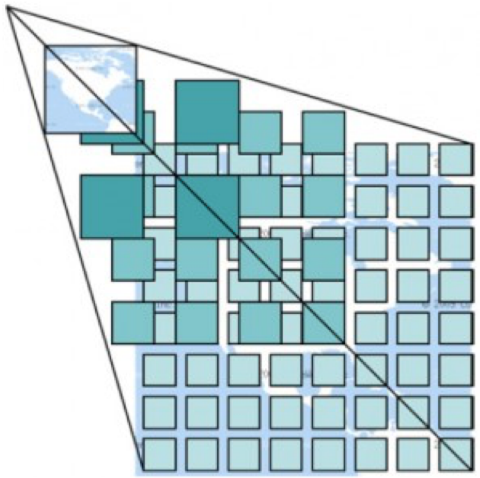
GeoPackage: Raster Maps, Images, Feature Data in One File

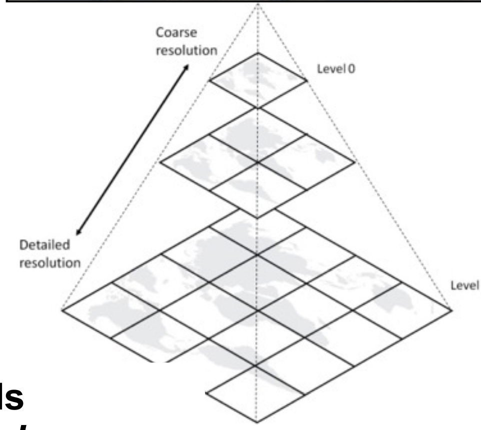


Imagery
Low – High Resolution

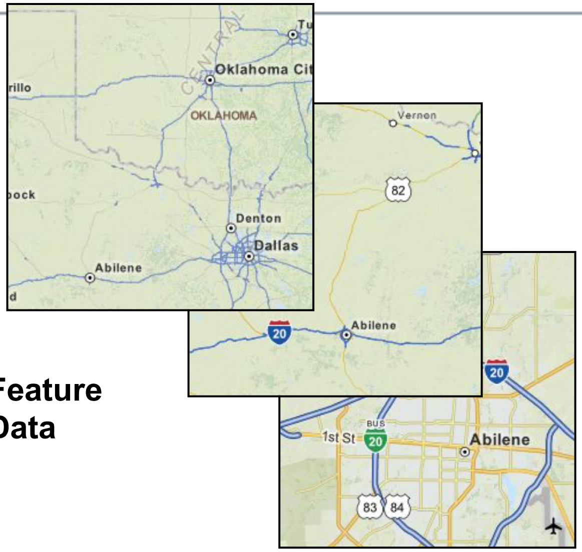


Raster Maps:
Small – Large Scale





Tile Pyramids
24 zoom levels



Feature Data

OGC®

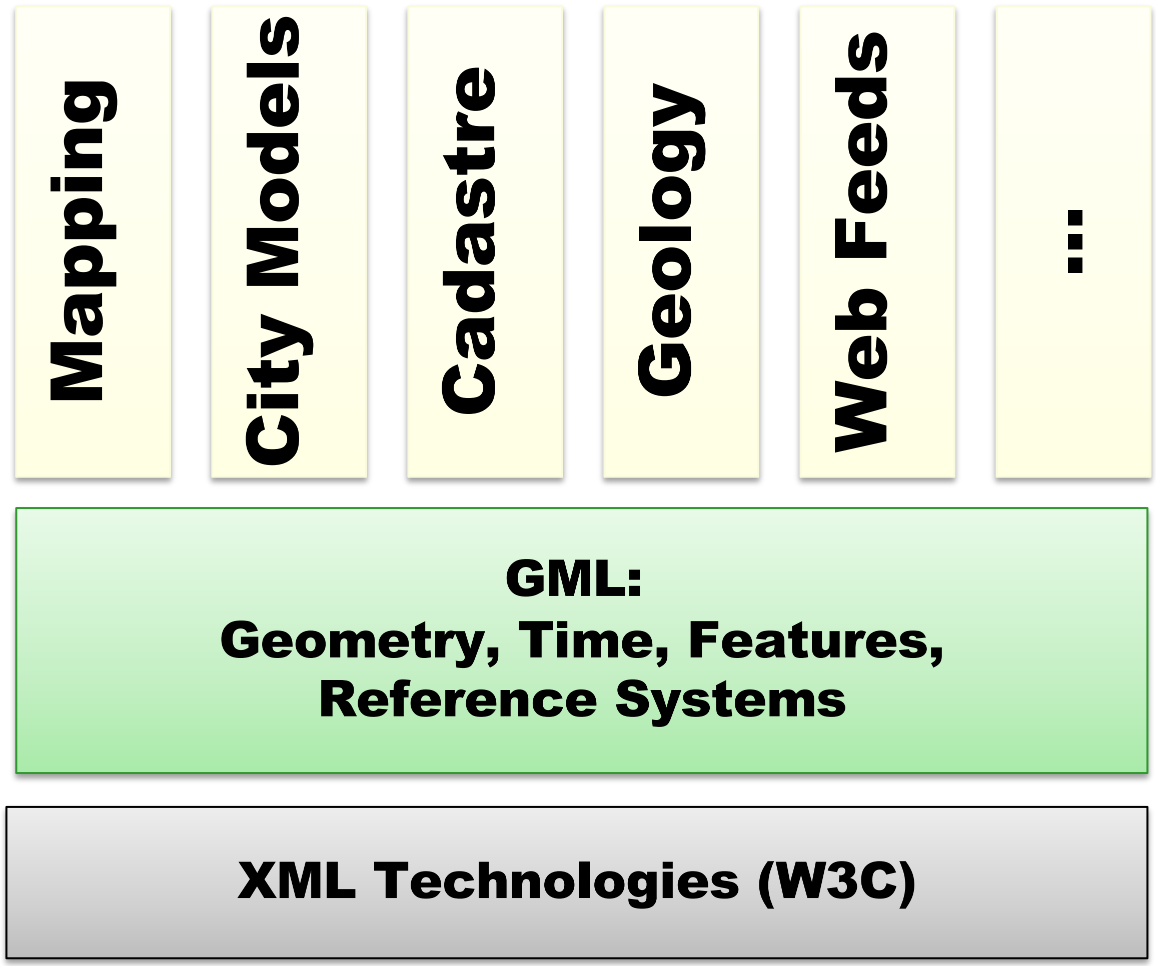
Single File Sqlite Database
containing all data for direct-use on mobile platforms & handheld devices

OGC Geography Markup Language (GML)

Two Different Use Patterns

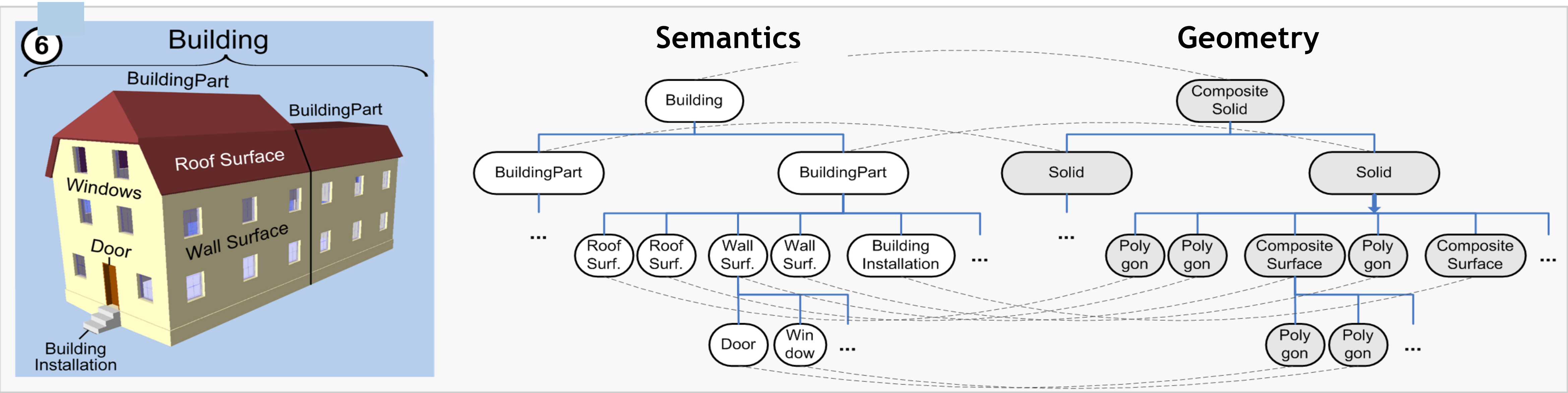
- **Thematic communities describe spatial datasets:** Cadastre, Topography, Geology, Hydrography, Meteorology, Aviation, City Models, etc.

- **Embed location in other XML grammars:** GeoRSS, GeoSPARQL (OGC), Geopriv (IETF), POI (W3C), Sensor Web (OGC), etc.



CityGML – Geometry and Semantics

CityGML: Simple to Complex objects with structured geometry



– Geometric entities know **WHAT** they are

– Semantic entities know **WHERE** they are and what their spatial extents are

OGC®

OGC GeoTIFF 1.1

- OGC standard for georeferencing information to be embedded in a TIFF file.
  - map projection, coordinate systems, ellipsoids, datums, etc. for spatial reference of data
- GeoTIFF compliant with TIFF
  - Software incapable of reading geo metadata will still open a GeoTIFF file
- Widely supported
  - including GDAL and libgeotiff

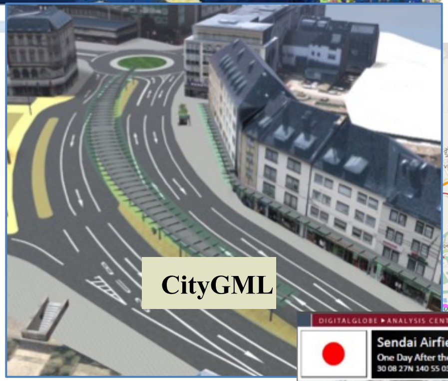
Geospatial Data and Processing in Apache Projects

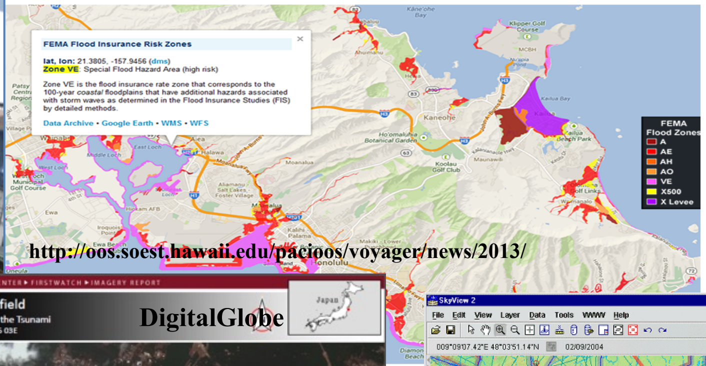
APIs and Web Services

OGC®

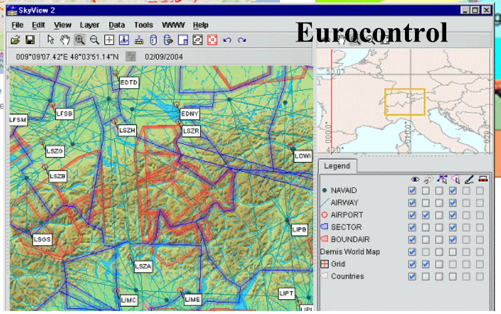
1000s of Services, 100Ks Datasets Implement OGC Web Services

Web Map Service (WMS)
Web Map Tile Service (WMTS)
Web Feature Service (WFS)
Web Coverage Service (WCS)
KML, GML, GeoPackage
GeoTIFF, NetCDF, HDF

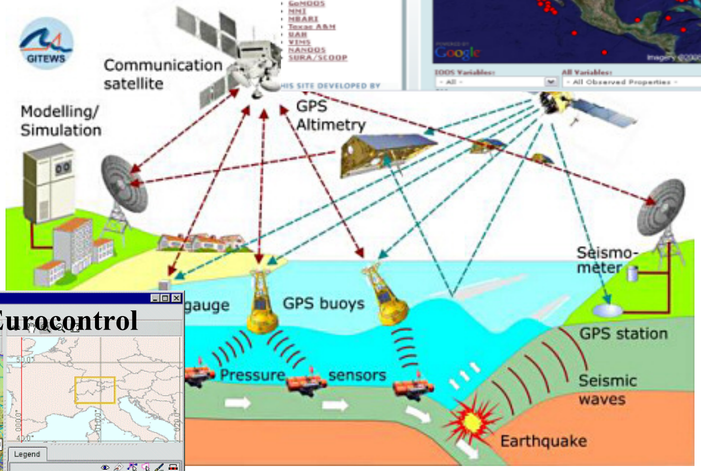




Emergency / Disaster Management



Aviation Flight Information / Safety



Meteorology, Hydrology, Ocean Monitoring

OGC®

OGC API Standards

Modular API building blocks to spatially enable Web APIs in a consistent way

- Modernized service architecture vs. OWS
- Leverages OpenAPI specification
- Current Web architecture and Spatial Data on the Web Best Practices,
- Focus on developer experience and usability
- Modular building blocks for fine-grained access to spatial data that can be used in data APIs,
- Open development; Public GitHub repo, Early implementations, In-depth validation, Slow release

OGC API - Features

OGC API - Coverages

OGC API - Map Tiles

OGC API - Processing

OGC API – Features – Part 1: Core

Resources of the Web API

Table 1. Overview of resources, applicable HTTP methods and links to the document sections

Resource	Path	HTTP method	Document reference
Landing page	`/`	GET	7.2 API landing page
Conformance declaration	`/conformance`	GET	7.4 Declaration of conformance classes
Feature collections	`/collections`	GET	7.12 Feature collections
Feature collection	`/collections/{collectionId}`	GET	7.13 Feature collection
Features	`/collections/{collectionId}/items`	GET	7.14 Features
Feature	`/collections/{collectionId}/items/{featureId}`	GET	7.15 Feature

information about the API

a dataset with a sub-division into named collections of features

the features

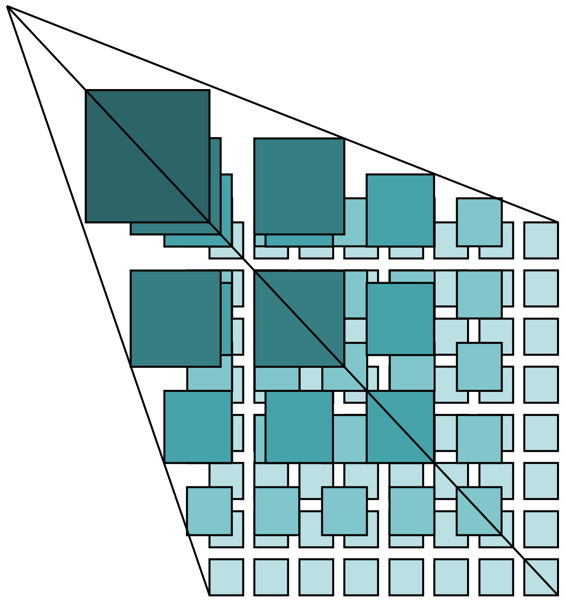
http://docs.opengeospatial.org/DRAFTS/17-069r1.html#tldr

OGC®

OGC API - Maps Tiles

Indexes 2D space as regular grids with scales in a projected CRS

- Path = /collections
  - Returns metadata describing \ collections at this API
- Path = /collections/{collectionId}
  - Returns metadata describing {collectionId}
- Path = /tileMatrixSet
  - Returns all available tile matrix sets
- Path = /tileMatrixSet/{tileMatrixSetId}
  - Returns a tiling scheme by id



OGC®

OGC API - Coverages

- Geospatial Coverage: “spatial function” or “field”, Spatial domain to Values Range

- OpenAPI = URL-oriented request language, aligned across OGC Coverage standards
  - *download coverage c001*

    http://acme.com/oapi/collections/{collectionid}/coverages/c001

  - *lat/long cutout, time slice t=2009-11-06T23:20:52”*

    http://acme.com/oapi/collections/{collectionid}/coverages/c001?
    SUBSET=Lat(40,50)&SUBSET=Long(10,20)

  - *coverage c001, in GeoTIFF*”: either http accept header, or:

    http://acme.com/oapi/collections/{collectionid}/coverages/c001?
    F=“image/tiff“

Coverage Analytics Sprint
January 2020

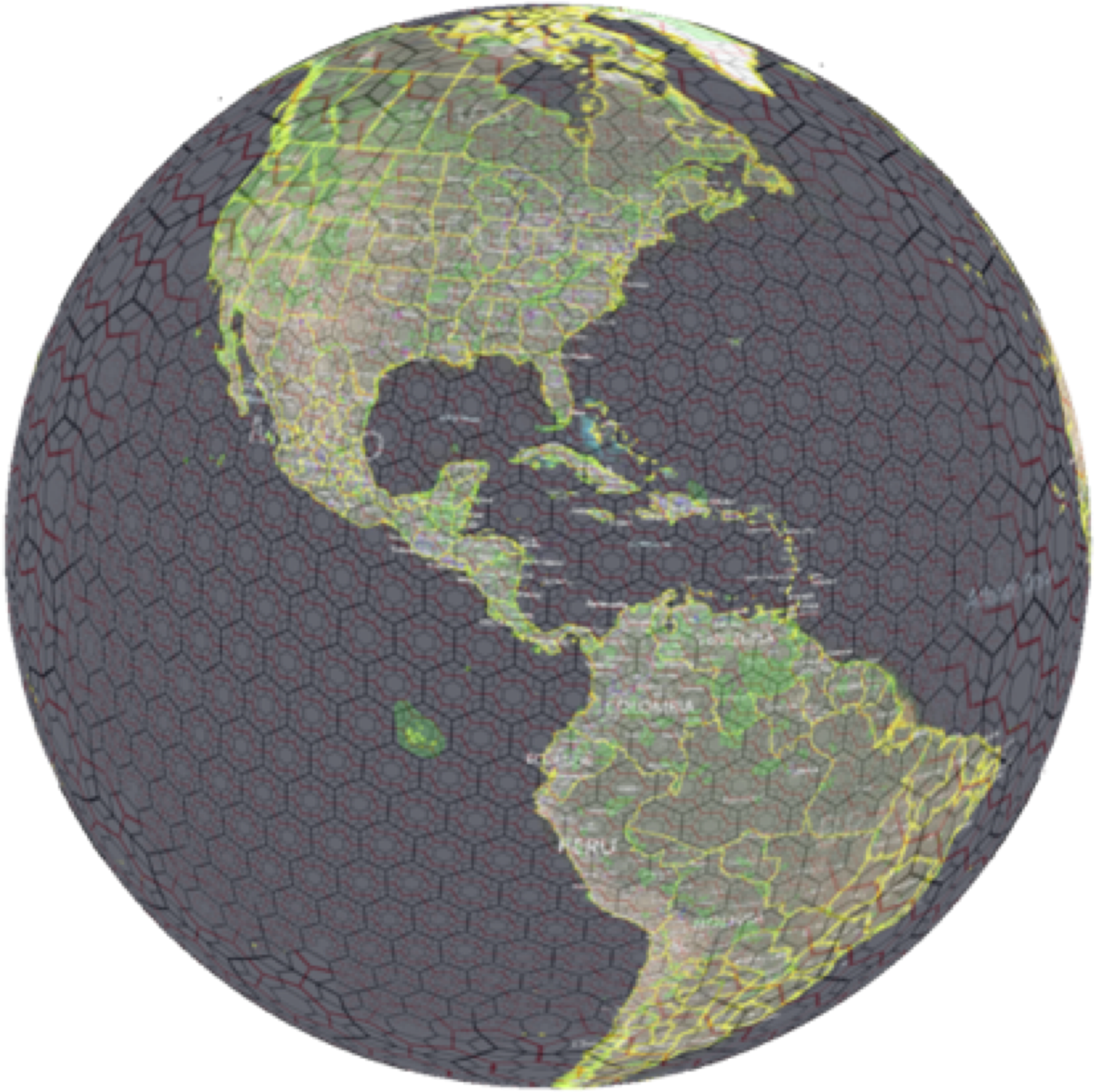
OGC®

Geospatial Data and Processing in Apache Projects

Discrete Global Grid Systems

OGC®

Discrete Global Grids



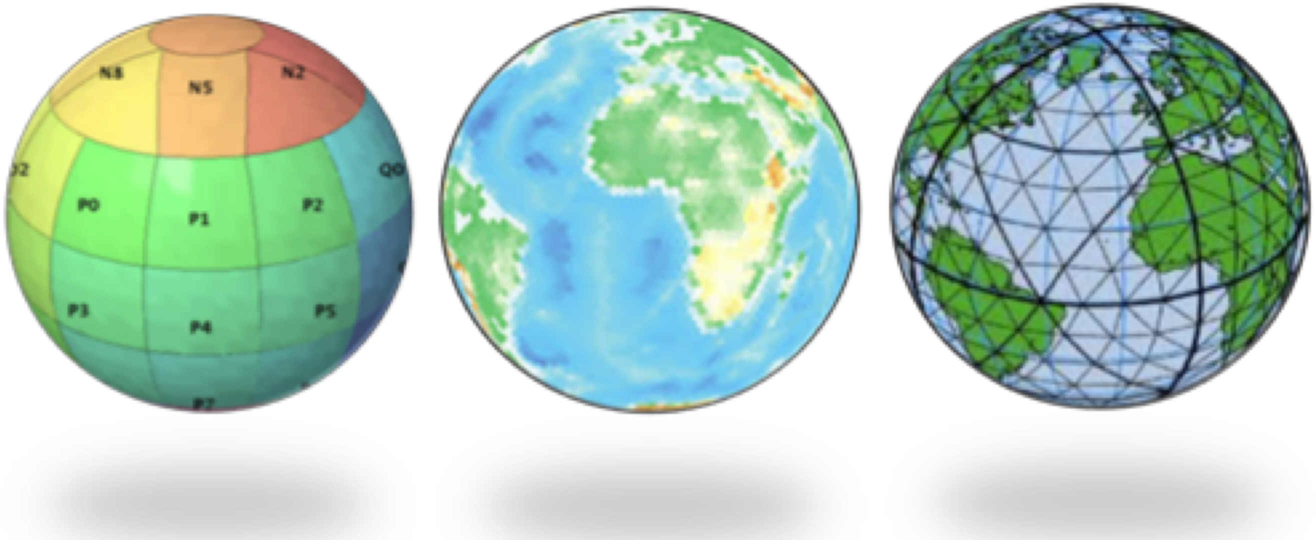
Slide source:
Matthew B. J. Purss,
Robert Gibb, Faramarz
Samavati, Perry
Peterson, Jin Ben,
Roger Lott

OGC®

Discrete Global Grid Systems

"…a spatial reference system that uses a hierarchical tessellation of cells to partition and address the globe. DGGS are characterized by the properties of their cell structure, geo-encoding, quantization strategy and associated mathematical functions."

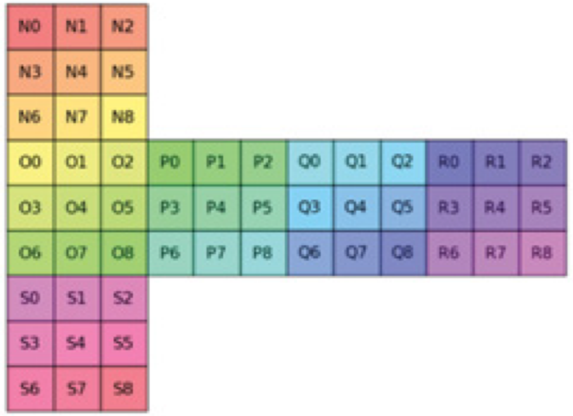
– OGC DGGS Standard



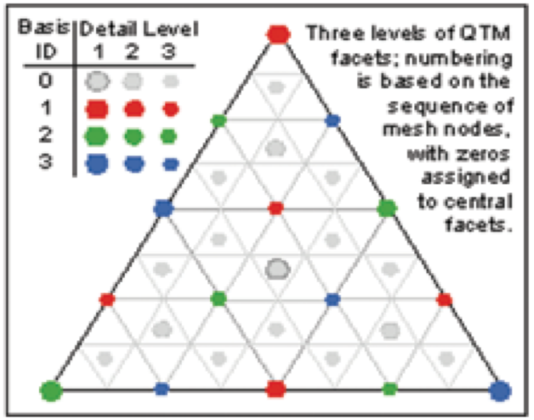
OGC®

Standardising Discrete Global Grid Systems

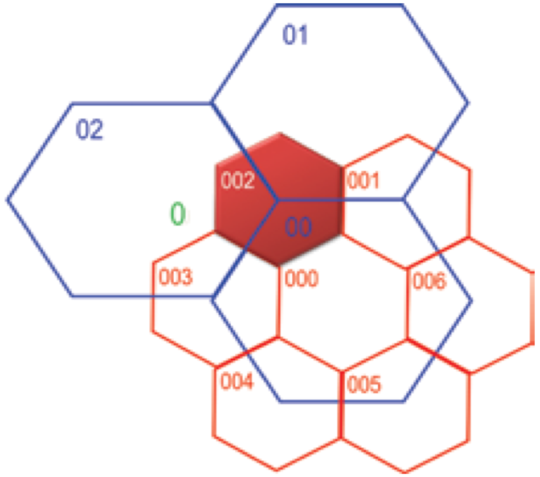
Different Cell Shapes



Square = Familiar



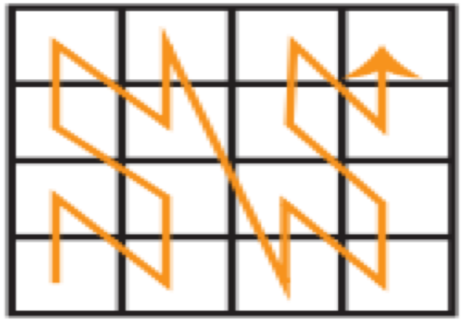
Triangular = Fast



Hexagonal = Fineness of Fit

Unique Cell Indices

- *Hierarchy-based, Space-filling Curve, Axes-based or Encoded Address*



11	13	31	33
10	12	30	32
01	03	21	23
00	02	20	22

nD Spatial Analyses

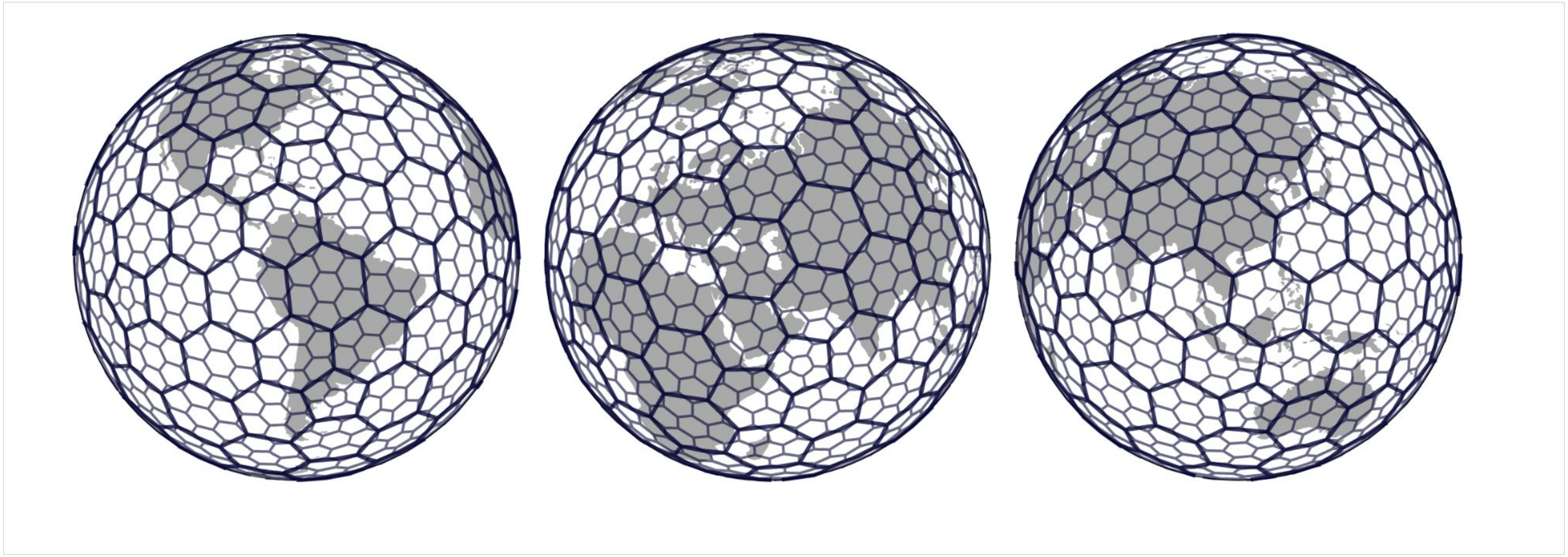
↓

1D Array Processes

00	01	02	03	10	11	12	13	20	21	22	23	30	31	32	33

OGC®

Hexagons!



Geospatial and Temporal Forecasting at Uber

September 09, 2019 Apachecon

Chong Sun, Brian Tang

Geospatial Data and Processing in Apache Projects

Geovisualization

OGC®

3D Geospatial Visualization

Berlin with 3D and Textures for Visualization

New York City portayal of attributes

The approach: Support multiple 3D data formats

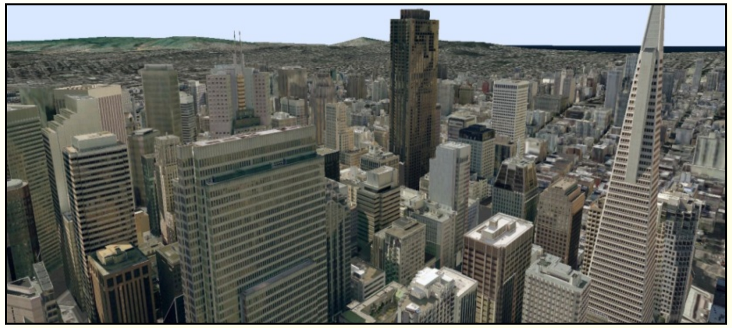
X3D	glTF	3D Tiles	I3S	KML	**Content Delivery**
OGC 3D Portrayal Service					**Query**

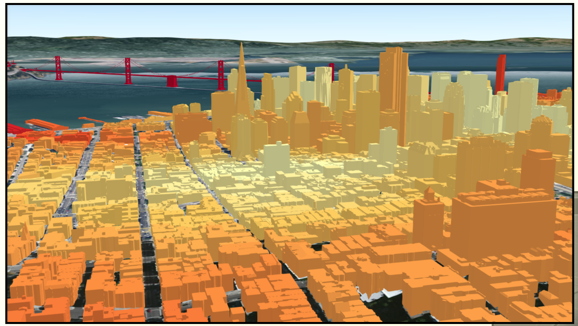
OGC®

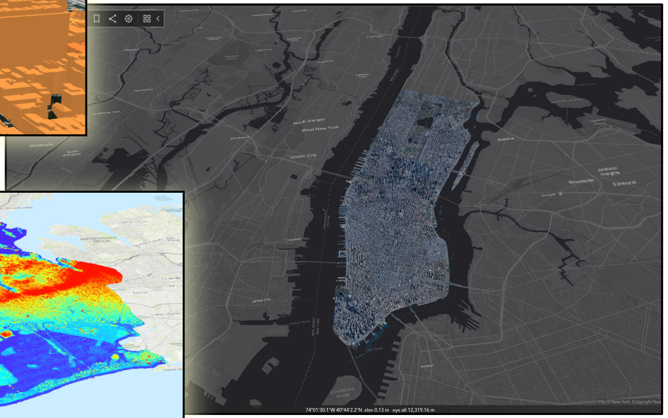
Indexed 3D Scene Layer (I3S)

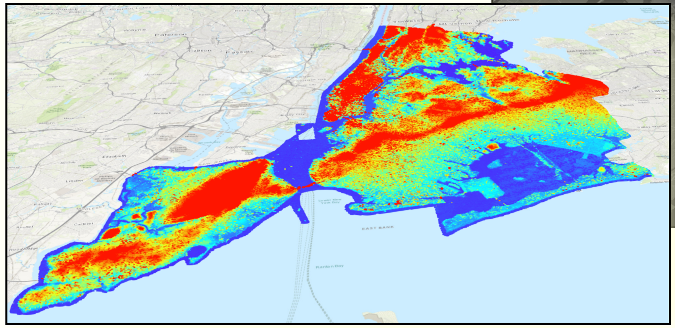
- Storage and transmission of large, heterogeneous 3D geospatial data sets
- 3D geospatial content, various coordinate systems along with a rich set of layer types
- Expandable to accommodate new data types and access patterns

- Developed by Esri now also an OGC Community Standard









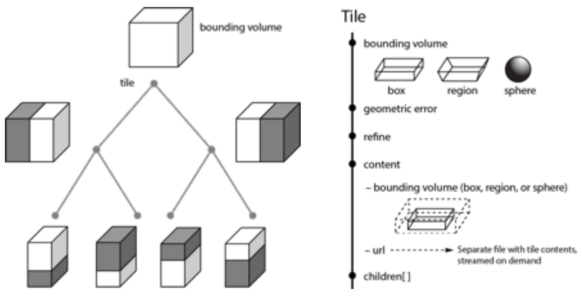
- 3D Objects
- Points
- Integrated Meshes
- Point Clouds

OGC®

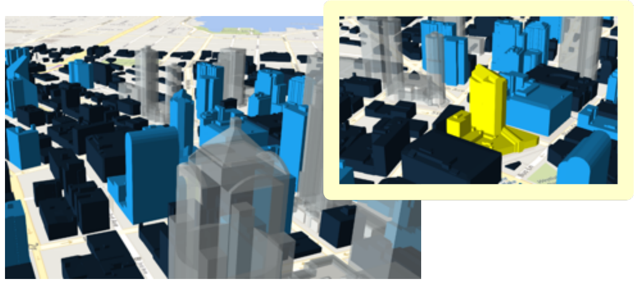
OGC Community Standard

Massive heterogeneous 3D geospatial datasets with semantics

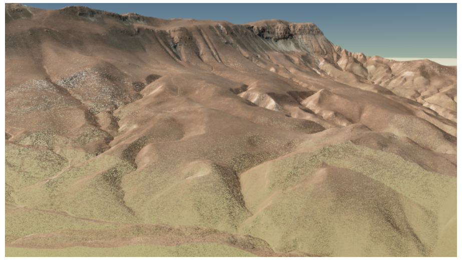
Spatial data structures + glTF + Styling + Metadata




```
{
  "show" : "${Area} > 0",
  "color" : {
    "conditions" : {
      "${Height} < 60" : "color('#13293D')",
      "${Height} < 120" : "color('#1B98E0')",
      "true" : "color('#E8F1F2', 0.5)"
    }
  }
}
```

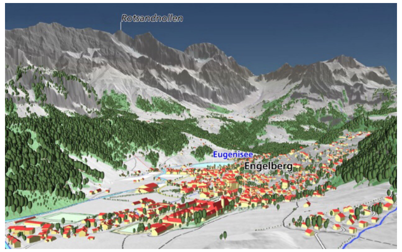


Photogrammetry



Point clouds

3D buildings

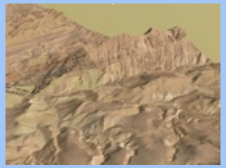


Terrain

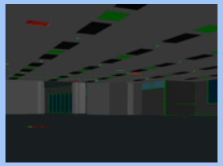
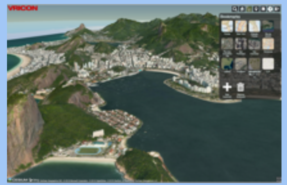
3DTiles Ecosystem

Exporters / Tilers



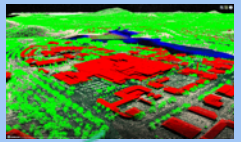



AGI CESIUM COMPOSER

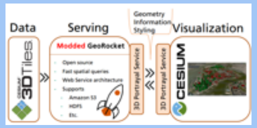

Bentley ContextCapture Bentley MicroStation VRICON entwine






virtualcitySYSTEMS GAMESIM DATA 61 VirtualGIS SAFE SOFTWARE FME

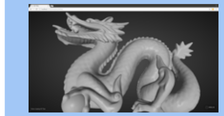


Fraunhofer IGD LOPoCS @mattshax Fraunhofer IGD cityzenith

Visualization engines

CESIUM

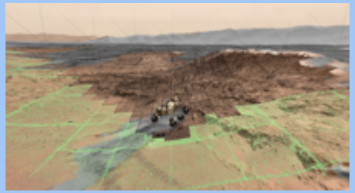


OSG.JS

three.js prototype

SiteSee

NASA considering unity



Validator in-progress

Built on glTF KHRONOS GROUP Google Microsoft Adobe NVIDIA Oculus VR

OGC®

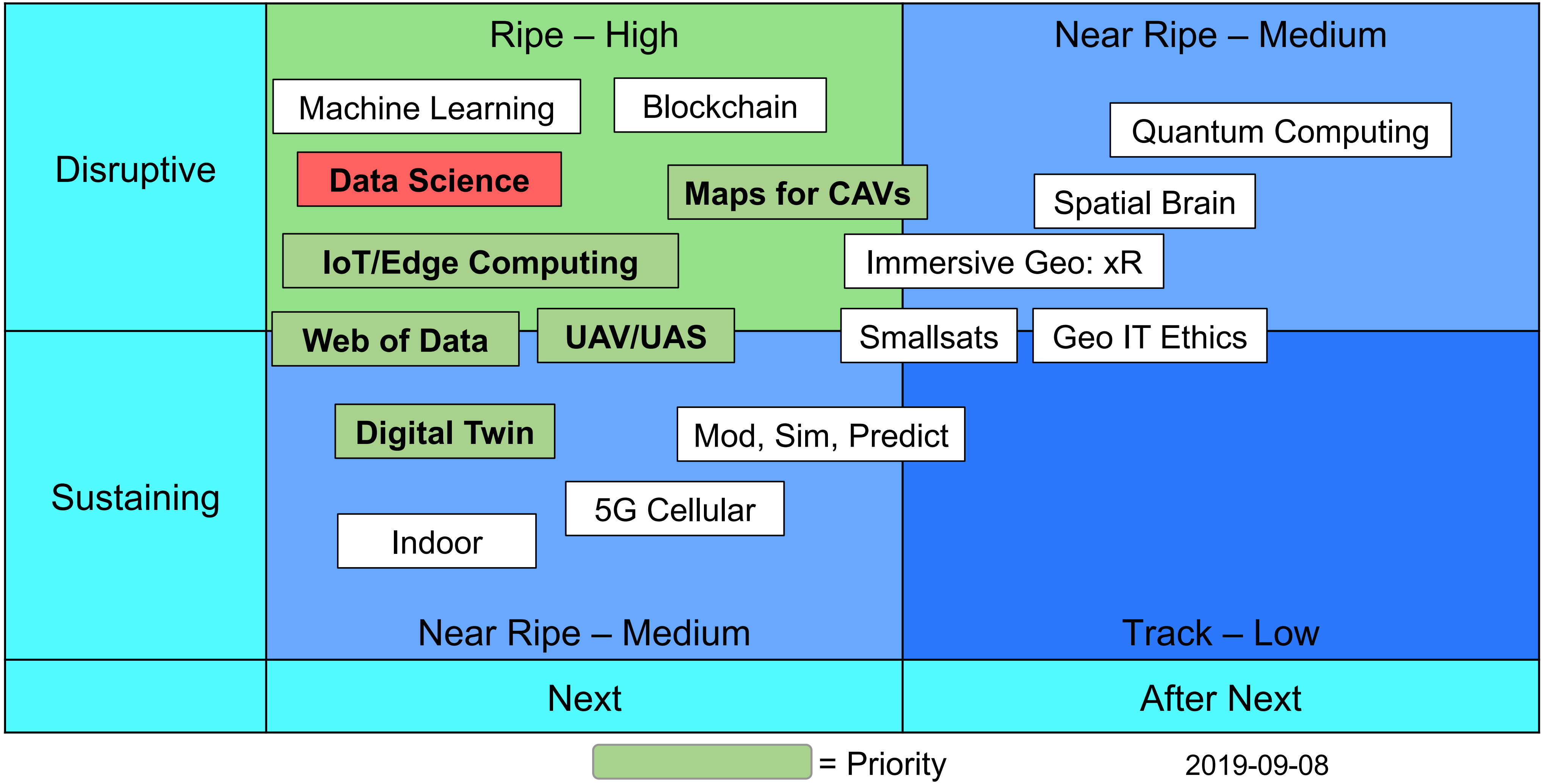
Slide source: Cesium at OGC TC

Geospatial Data and Processing in Apache Projects

Geospatial Trends

OGC®

Geospatial Tech Trends

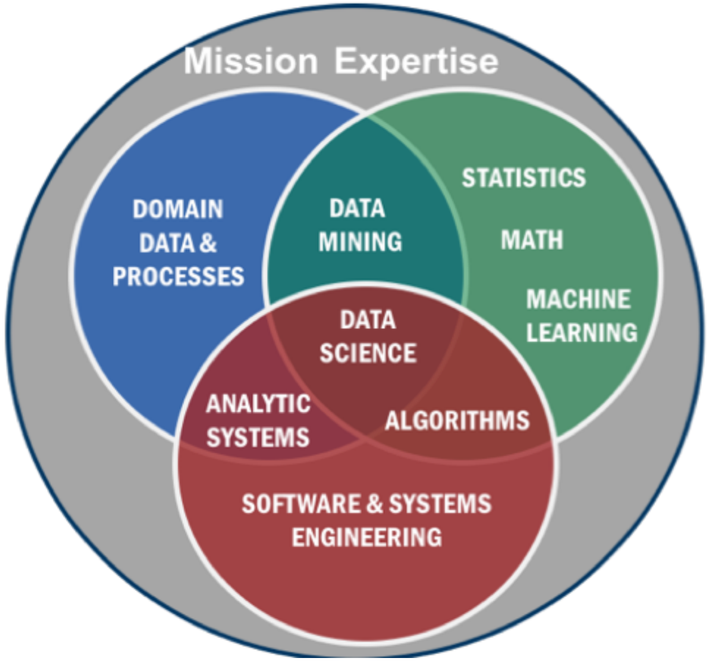


https://github.com/opengeospatial/OGC-Technology-Trends/

OGC®

Location Powers: Data Science

13th & 14th November 2019
Google, Mt View, CA



Data Science – source NIST

- Sessions
  1. Foundations
  2. Analytics
  3. Ripe Trends
  4. Outcomes
  5. Actions to Take
- Explosive availability of data on every aspect of human activity; with revolutionary advances in computing technologies is transforming geospatial data science.

http://locationpowers.net

OGC®

GeoSpark

Google "GeoSpark ASU"

Arizona State University

http://datasystemslab.github.io/GeoSpark/

In production!

mobike

GYANA

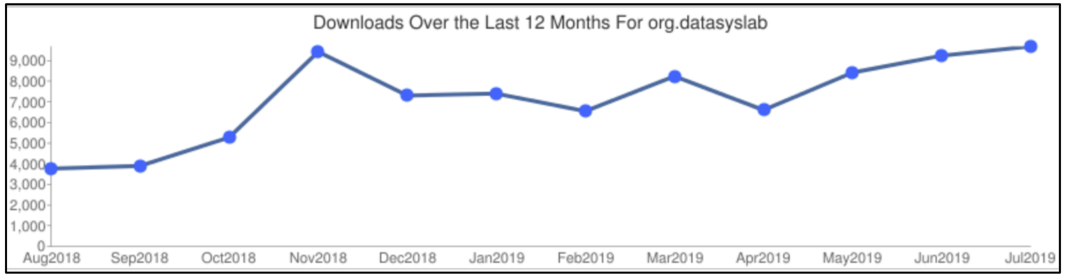
Blue DME

- Geospatial DBMS on top of Apache Spark
  - Spatial RDD, Spatial SQL, Spatial DataFrame
  - Distributed geospatial visualization
  - Interactive SQL/map visualization (with Apache Zeppelin)

"GeoSpark comes close to a complete spatial analytics system. It also exhibits the best performance in most cases."

"How Good Are Modern Spatial Analytics Systems?" Varun Pandey, Andreas Kipf, Thomas Neumann, Alfons Kemper, PVLDB 2018

8K - 10K monthly downloads



Geospatial Track - ApacheCon NA 2019

0900 **Geospatial Data and Processing - Reusable Building Blocks**
- George Percivall, OGC

1000 **Geospatial Data Management in Apache Spark**
- Jia Yu, ASU & Mohamed Sarwat, ASU

1130 **Apache Science Data Analytics Platform Apache (SDAP)**
- Frank Greguska, JPL

1400 **Using GeoMesa on top of Accumulo, HBase, Cassandra, and big data file formats for massive geospatial data - a LocationTech Project**
- James Hughes, CCRI

1500 **Geospatial Indexing and Search at Scale with Apache Lucene**
- Nick Knize, Elastic

1630 **Geospatial and Temporal Forecasting in Uber Marketplace**
- Chong Sun and Brian Tang, Uber

1730 **Realtime Geospatial Analytics with GPUs, RAPIDS, and Apache Arrow**
- Josh Patterson, NVIDIA

Geospatial BoF
Monday, 9th Sep,
19:00 - 20:50
Red Rock VI-VIII

Join geospatial@apache.org

OGC®