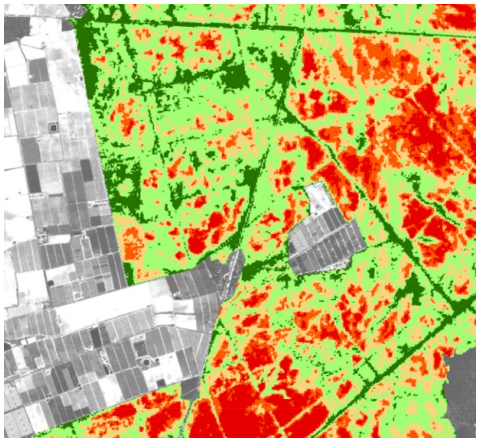
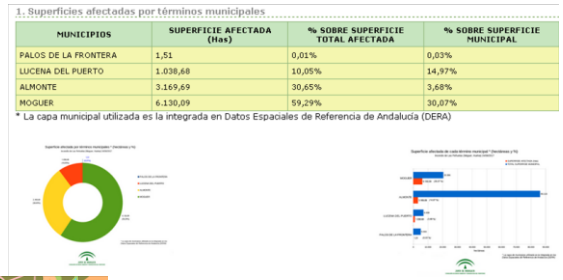
Cloud-IA:"E-infraestructura de Información Ambiental para el Desarrollo de Soluciones Basadas en la Integración Espacial Normalizada de Variables"

OBJETO: Creación de una e-infraestructura de información ambiental integrada espacialmente, con gran capacidad de almacenamiento, gestión, análisis y difusión web en un entorno cloud, que permita la generación de soluciones a medida para gestores, investigadores y ciudadanía no expertos en TIGs, en ámbitos relacionados con la gestión ambiental y territorial.

SOLUCIONES: Indicadores, estadísticas, modelizaciones, geovisores web, cuadros de mando (dashboard)….





1. Superficies afectadas por términos municipales

MUNICIPIOS	SUPERFICIE AFECTADA (Has)	% SOBRE SUPERFICIE TOTAL AFECTADA	% SOBRE SUPERFICIE MUNICIPAL
PALOS DE LA FRONTERA	1,51	0,01%	0,03%
LUCENA DEL PUERTO	1.038,68	10,05%	14,97%
ALMONTE	3.169,69	30,65%	3,68%
MOGUER	6.130,09	59,29%	30,07%

* La capa municipal utilizada es la integrada en Datos Espaciales de Referencia de Andalucía (DERA)

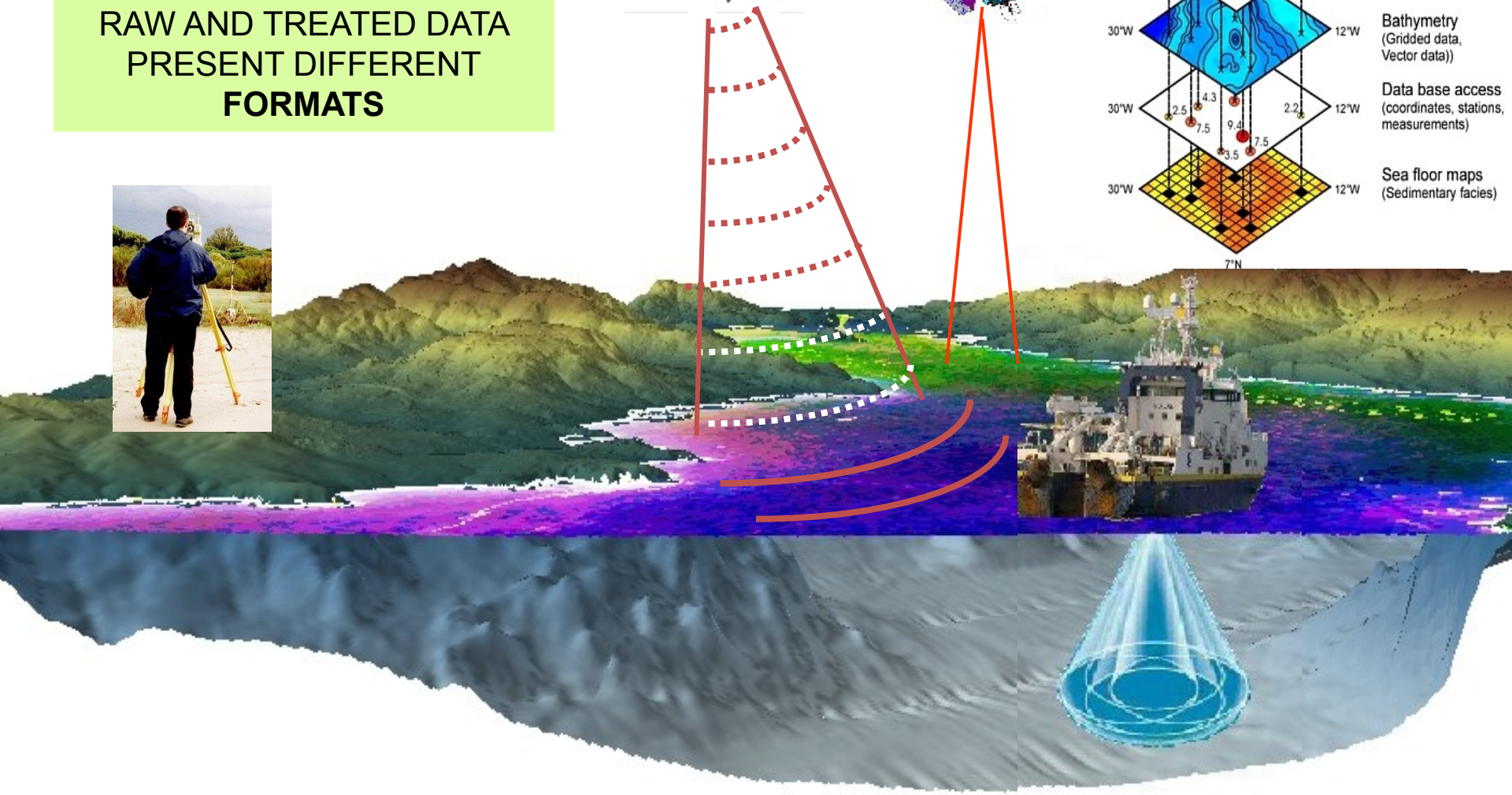
Cloud-IA:"E-infraestructura de Información Ambiental para el Desarrollo de Soluciones Basadas en la Integración Espacial Normalizada de Variables"

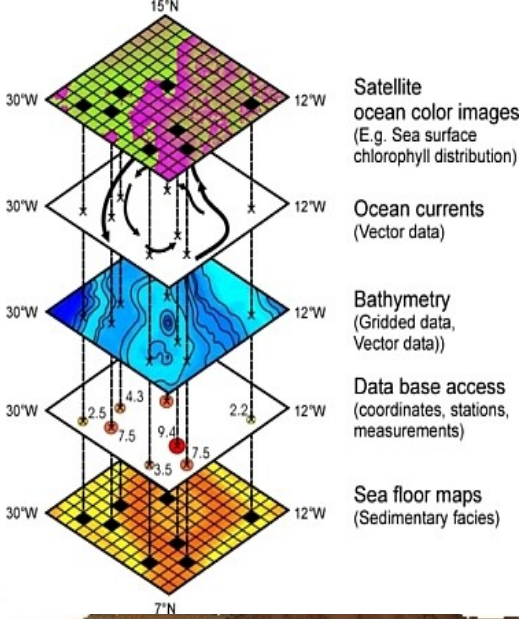
NECESIDAD DE DISPONER DE **VARIALES AMBIENTALES YA INTEGRADAS ESPACIALMENTE**

Si bien los avances en los últimos años (INSPIRE, OPEN DATA) han permitido la **normalización de las variables ambientales individualmente** (normas, modelos, especificaciones, etc.), **sigue pendiente** la **necesaria integración espacial de las mismas de forma normalizada**. Así, a día de hoy, esta integración se lleva a cabo, solo por especialistas en TIGs, y de forma particular para cada problemática, adoptándose diferentes soluciones: integración en formatos raster, utilización de rejillas grid, etc. **La necesidad de una propuesta innovadora y homogénea para la integración espacial de las variables ambientales** más demandadas por la administración ambiental y la sociedad **es una clara necesidad no cubierta**

TIF, JPG, ECW, SID,BIL, BIP,IMG, SHP, DXF, DGN, DBF, SID, KML, ETC,,…

RAW AND TREATED DATA PRESENT DIFFERENT **FORMATS**





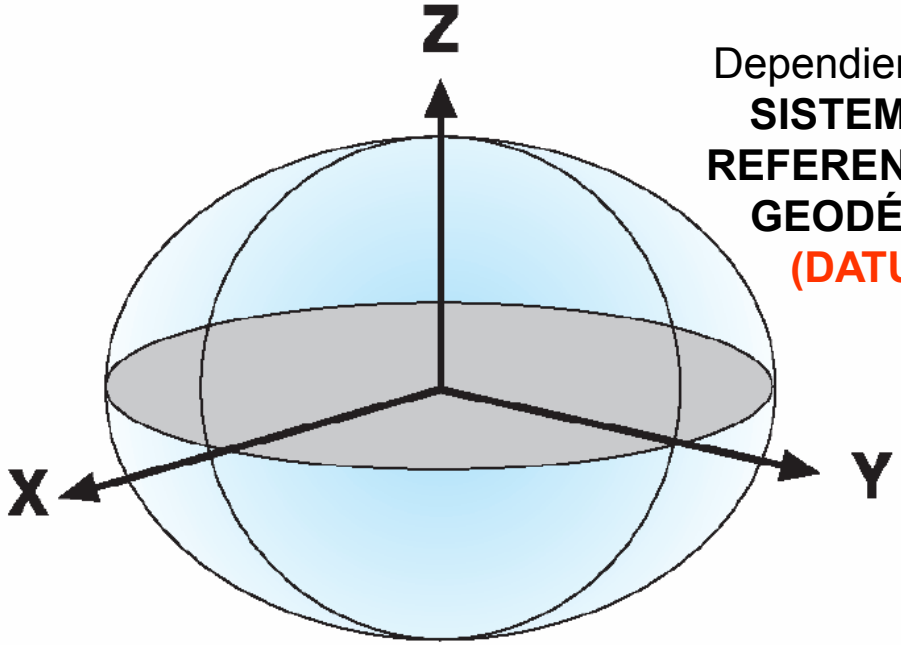
LA NECESARIA "INTEGRACIÓN ESPACIAL" DE LA GEOINFORMACIÓN MULTITEMÁTICA AMBIENTAL EN LA ACTUALIDAD.

LA INTEGRACIÓN ESPACIAL Y LA NORMALIZACIÓN DE DATOS MEDIOAMBIENTALES Y SOCIO-ECONÓMICOS

Integración y normalización espacial o geométrica

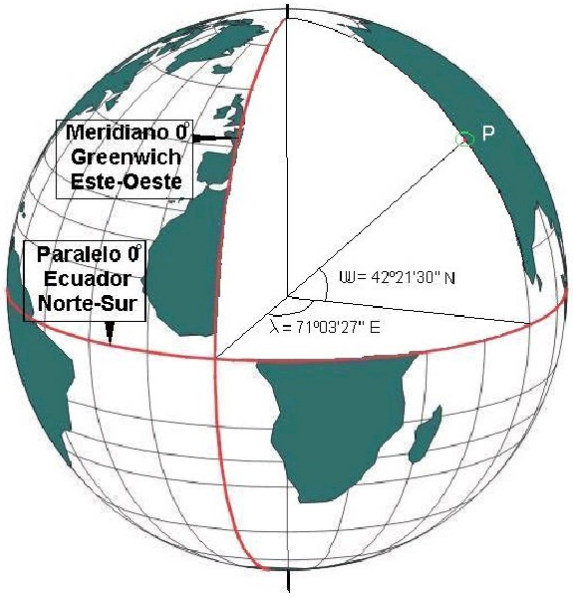
Normalización de la componente temática

CARTERSIANAS TRIDIMENSIONALES

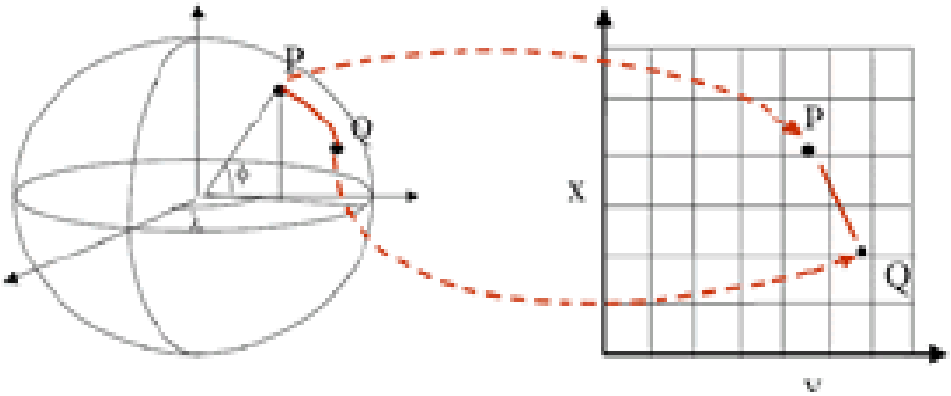


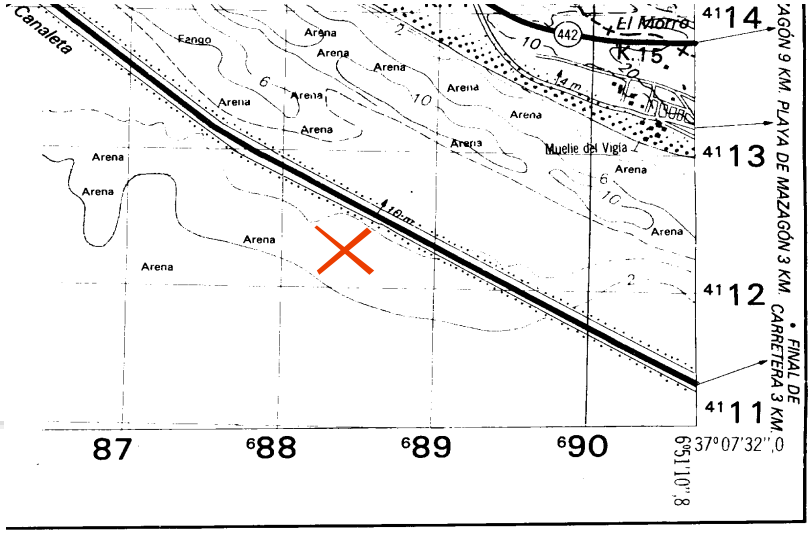
Dependientes del
SISTEMA DE REFERENCIA DE GEODÉSICO
(DATUM)

GEOGRÁFICAS/GEODÉSICAS

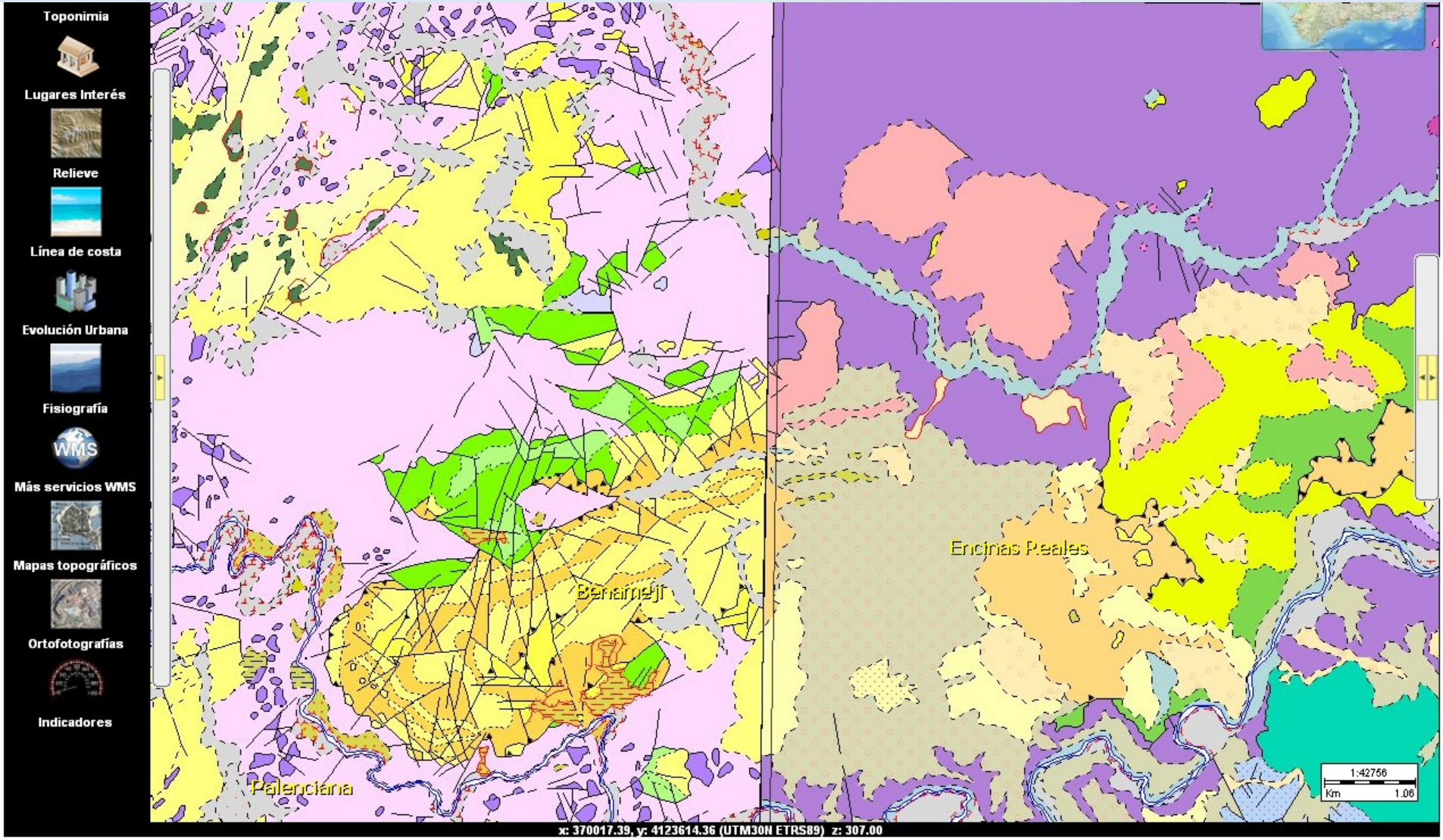


PROYECTADAS/PLANAS



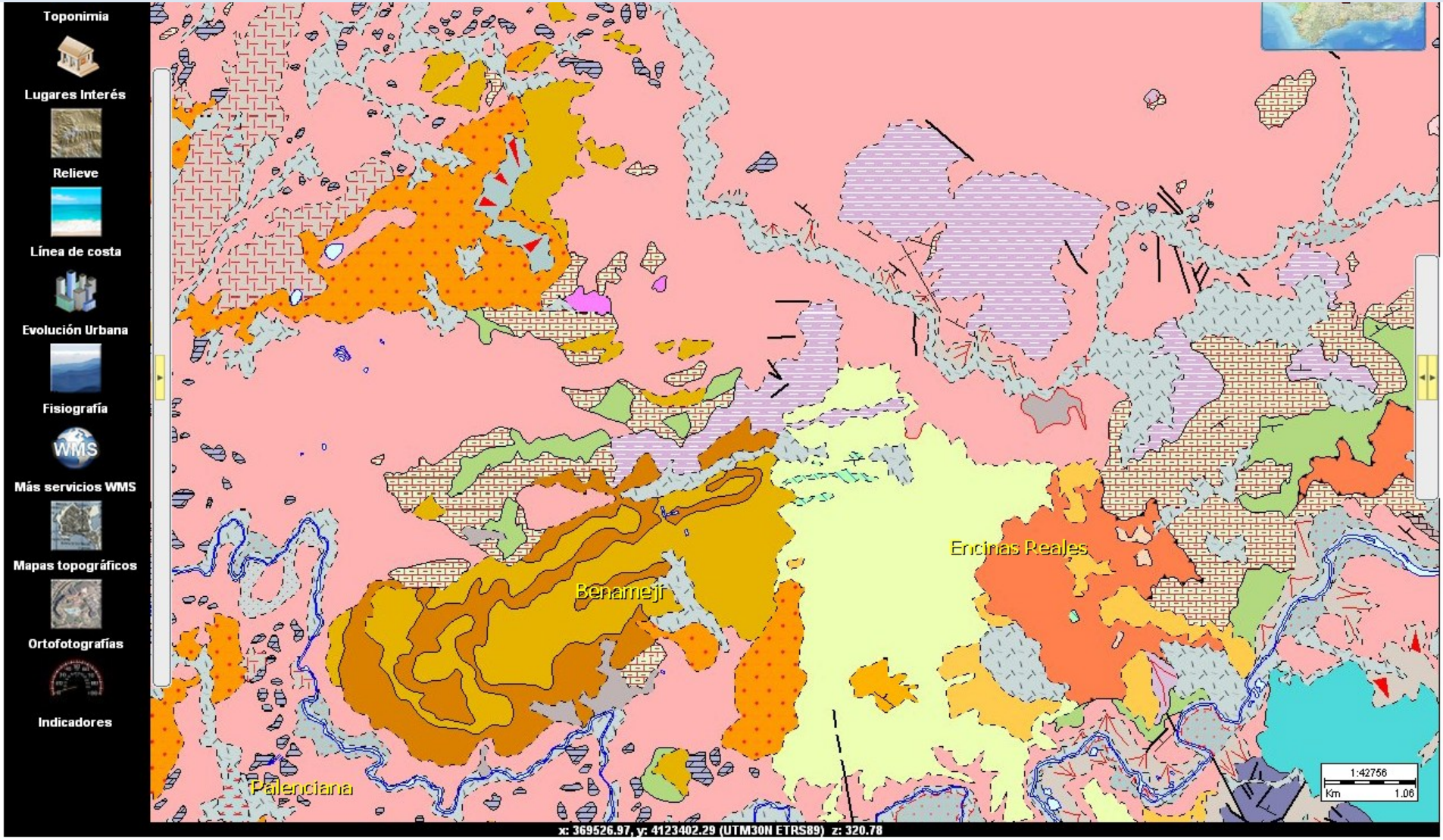


LA NECESARIA "INTEGRACIÓN ESPACIAL" DE LA GEOINFORMACIÓN MULTITEMÁTICA AMBIENTAL EN LA ACTUALIDAD.



Inicio CONF_COSTAS videos power_clase ORDENACIÓN... Presentación1... viewer.ppt European Atla... copernico.cica... ES 20:37

LA NECESARIA "INTEGRACIÓN ESPACIAL" DE LA GEOINFORMACIÓN MULTITEMÁTICA AMBIENTAL EN LA ACTUALIDAD.



Inicio CONF_COSTAS videos power_clase ORDENACIÓN... Presentación1... viewer.ppt European Atla... copernico.cica... ES 20:43

LA NECESARIA "INTEGRACIÓN ESPACIAL" DE LA GEOINFORMACIÓN MULTITEMÁTICA AMBIENTAL EN LA ACTUALIDAD.

LA INTEGRACIÓN ESPACIAL Y LA NORMALIZACIÓN DE DATOS MEDIOAMBIENTALES Y SOCIO-ECONÓMICOS

Integración y normalización espacial o geométrica

Normalización de la componente temática

Integración vector/raster (ficheros digitales)

Se realiza de forma particular (no normalizada) para cada problemática (necesidad expertos TIG)

LA NECESARIA "INTEGRACIÓN ESPACIAL" DE LA GEOINFORMACIÓN MULTITEMÁTICA AMBIENTAL EN LA ACTUALIDAD.

LA INTEGRACIÓN ESPACIAL Y LA NORMALIZACIÓN DE DATOS MEDIOAMBIENTALES Y SOCIO-ECONÓMICOS

Integración y normalización espacial o geométrica

Normalización de la componente temática

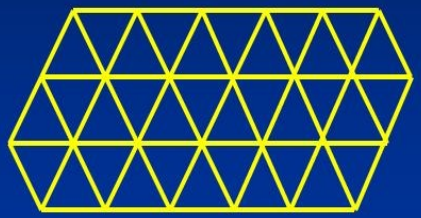
Integración vector/raster (ficheros digitales)

→ **Estructuras de rejillas o teselas multiescalares (GRID)**
Spatial tessellations

Tessellation Data Structures

Squares

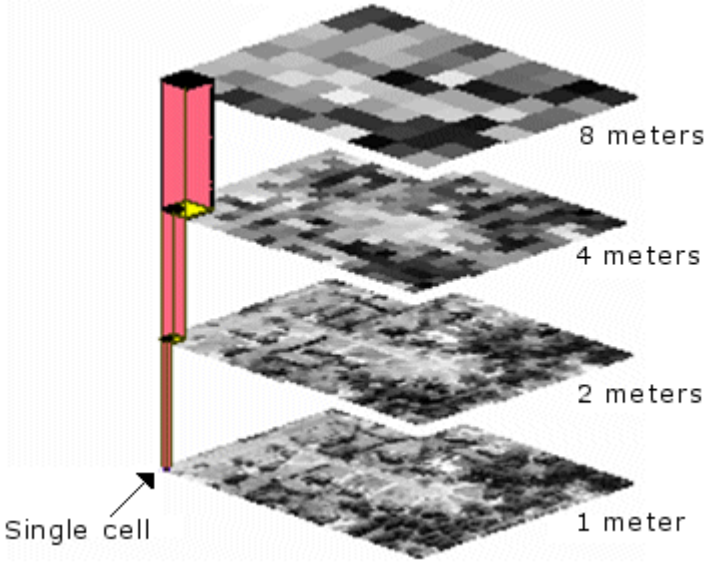
Triangles

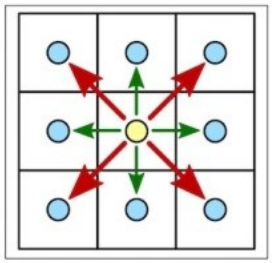


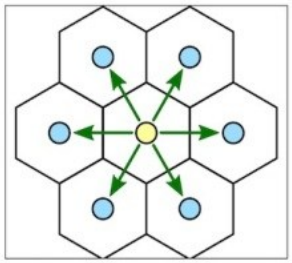
Hexagons

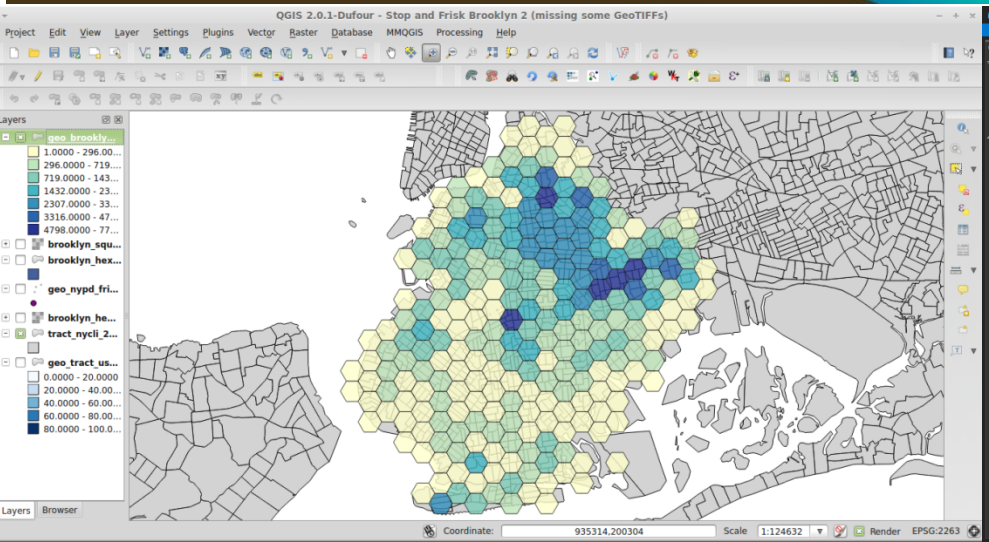
Properties:

- Orientation consistency
- Recursively subdividable?
- Radial symmetry
- Compatible with array data structures and a cartesian coorninate system.









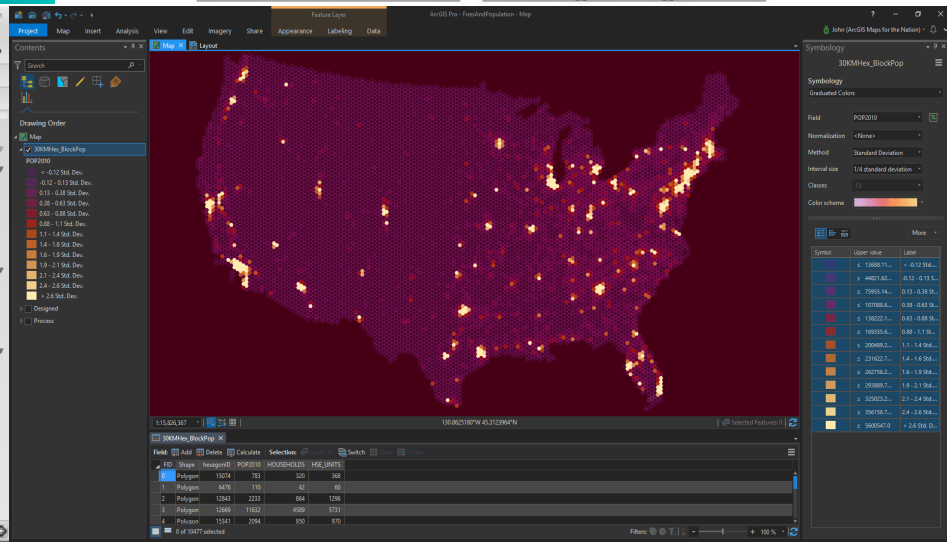
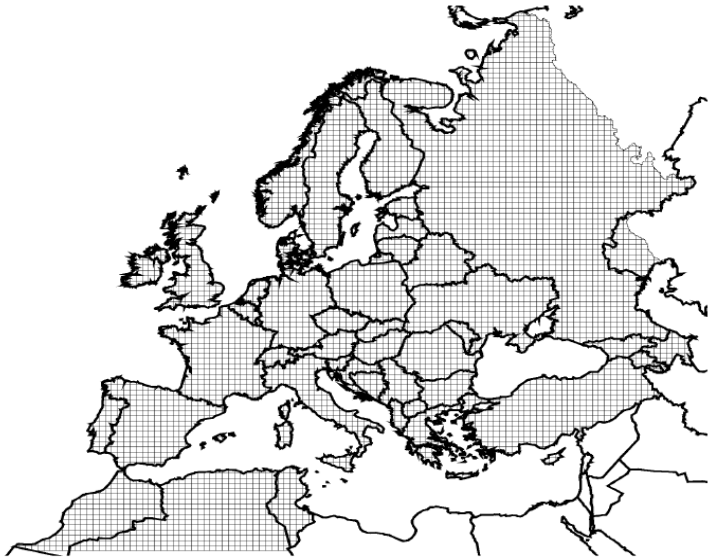


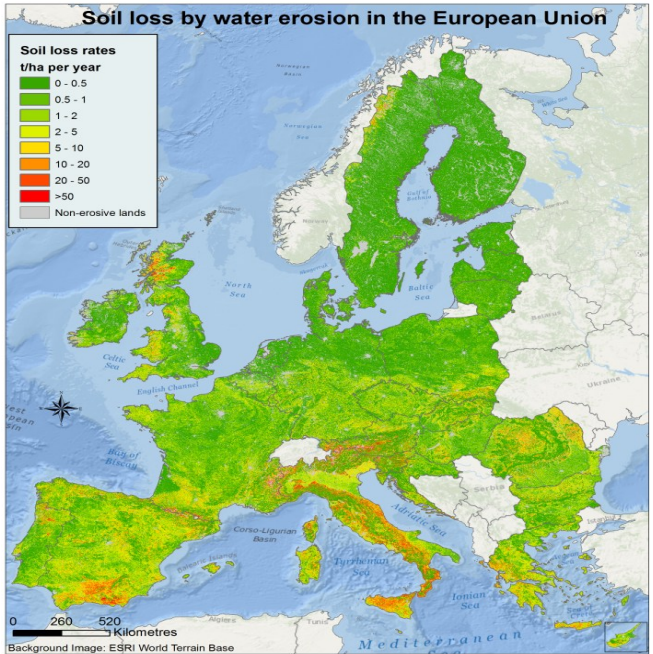
Figure 2 - CGMS 50x50 Km grid system

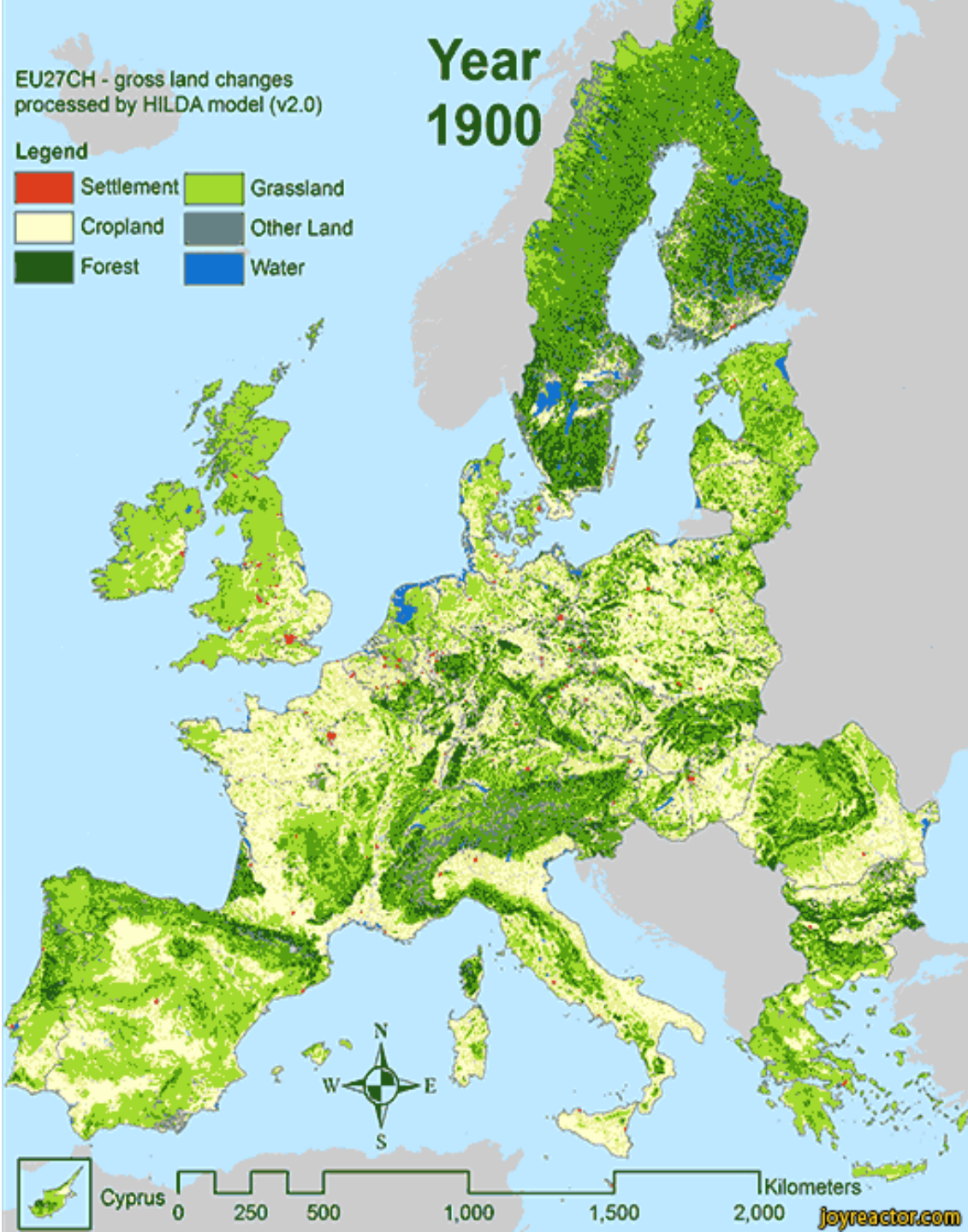


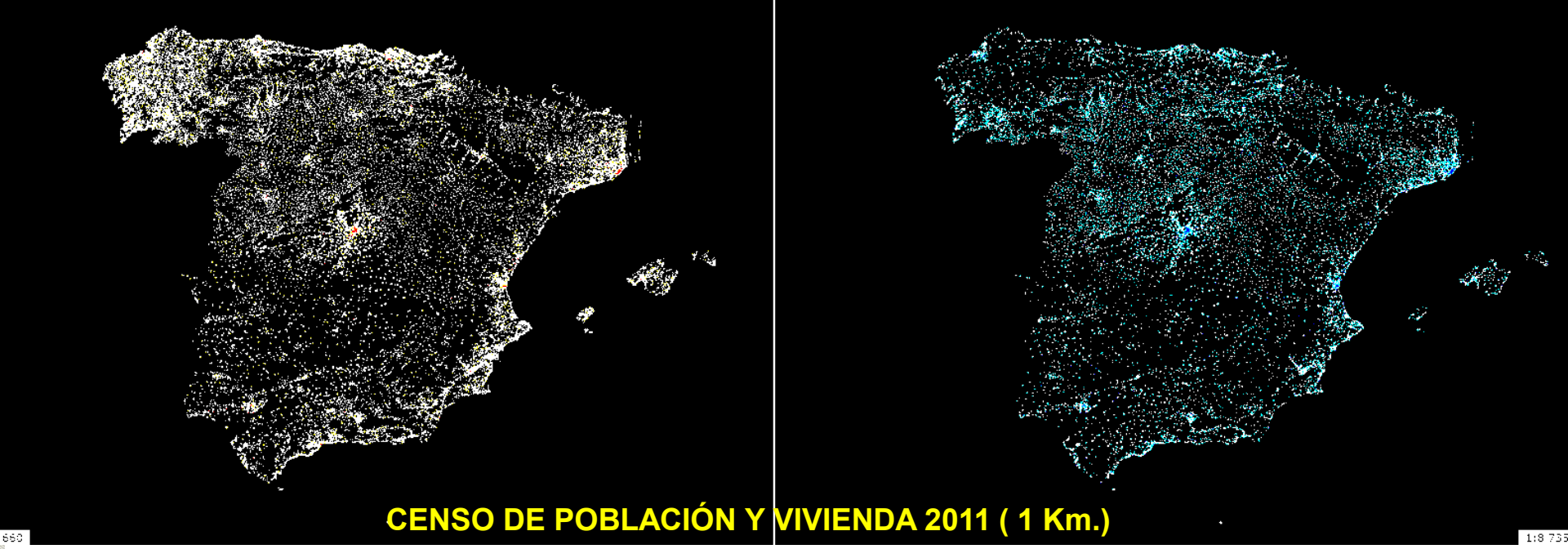
The CGMS Grid (from release 2.3 to current) was geographically defined according to the following parameters:

GRID projection and DATUM

- Projection: LAMBERT_AZIMUTH
- Units: METERS
- Spheroid: SPHERE
- Parameters 6378388







Instituto de Estadística y Cartografía de Andalucía
CONSEJERÍA DE ECONOMÍA Y CONOCIMIENTO

Calendario
Contacto

Está en: Inicio - Georreferenciación

Instituto | Estadística | Georreferenciación | Cartografía | Servicios

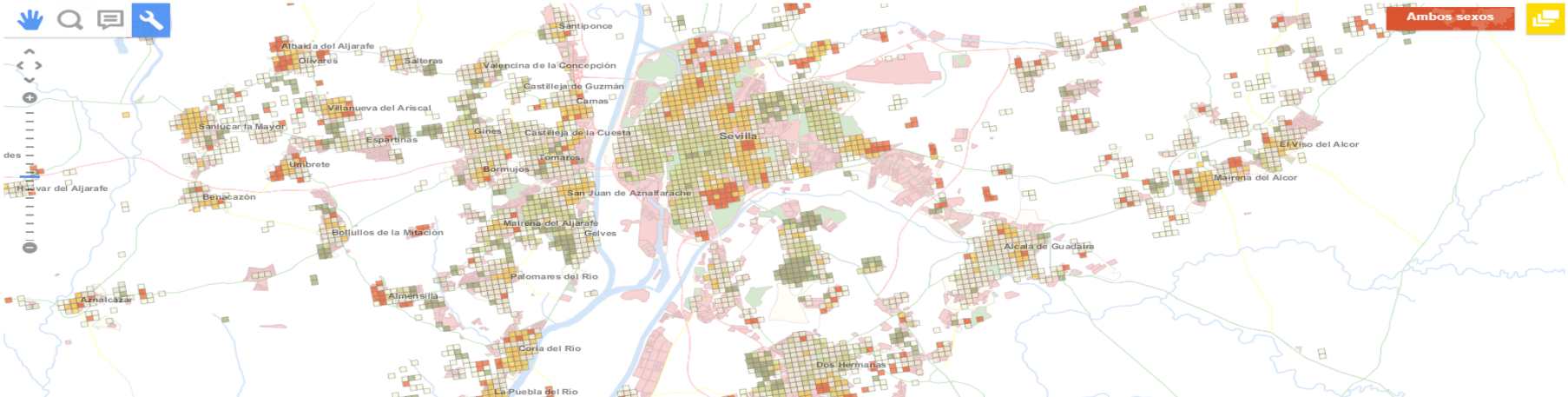
¿Qué información está buscando? Buscar

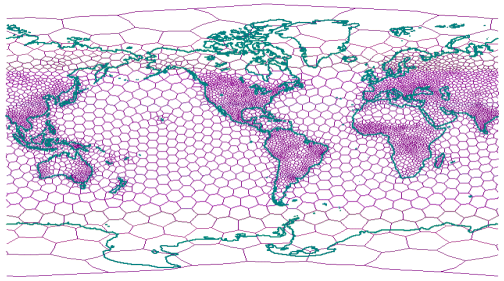
Razón de mortalidad general. Celdas de 250 metros

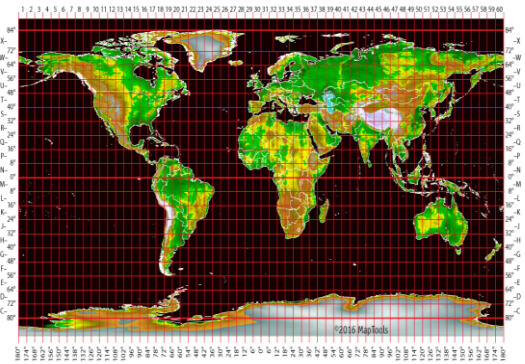
Andalucía se mueve con Europa

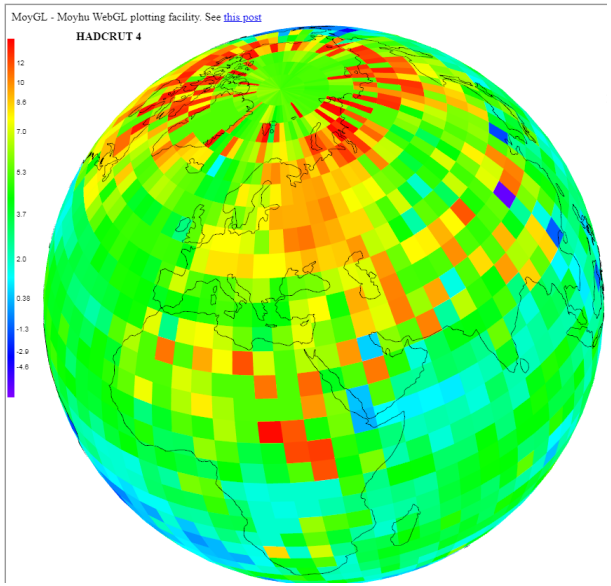
Estadísticas Longitudinales de Supervivencia y Longevidad en Andalucía, 2002-2013

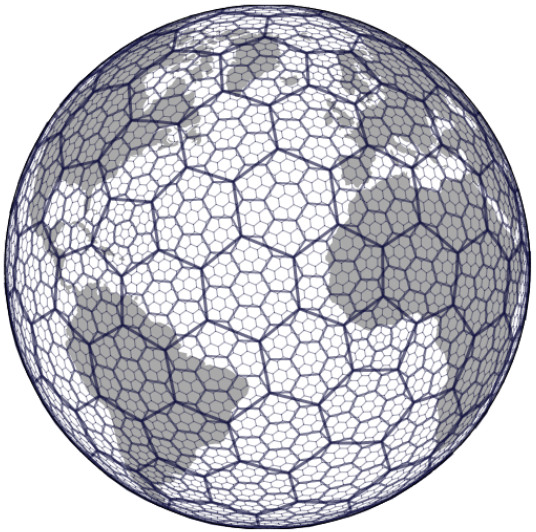
[Página del producto]

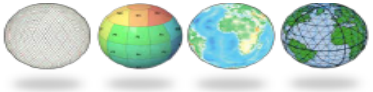




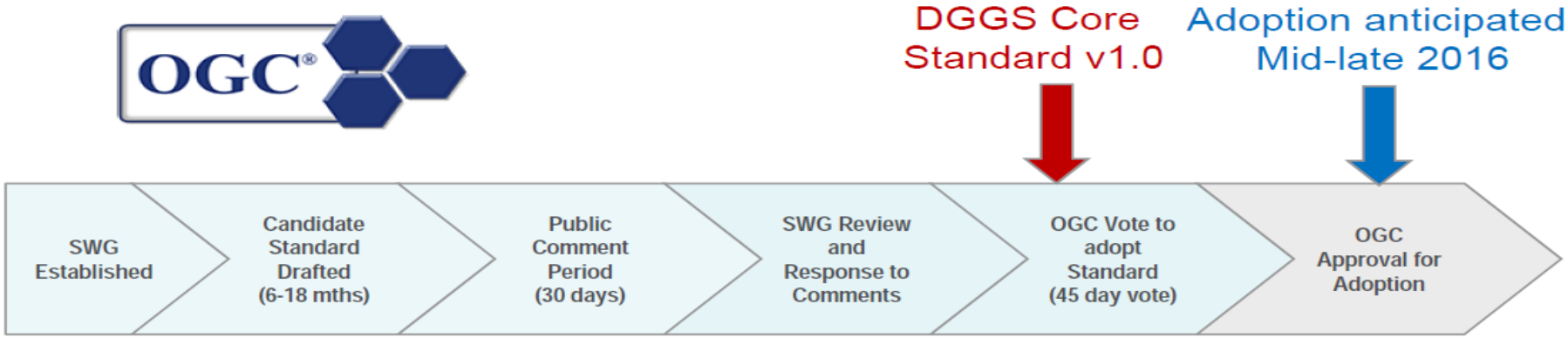








Standardising Discrete Global Grid Systems



OGC Standards Development Pipeline

Subscribe to our Press Release

Subscribe here to receive future Press Releases by email: Subscribe

2017

OGC announces a new standard that improves the way information is referenced to the earth

Contact:
info@opengeospatial.org

Release Date:
Tuesday, 24 October 2017 UTC

CHALLENGES OF GEOSPATIAL INTEGRATION

Geoscientific data has already exceeded the petabyte-scale barrier and is rapidly heading toward the exabyte-scale barrier. Converting this massive amount of data into timely information and decision support products is dependent on the capacity of the scientist to rapidly analyze this data in a transparent and repeatable fashion. The challenges of high velocity, high volume (> a terabyte per day) data is requiring those focused on combining and using these large data sources to rethink the way they store data in order to make best use of it. These challenges will only grow as requirements to combine these to produce near-real-time decision support information increases.

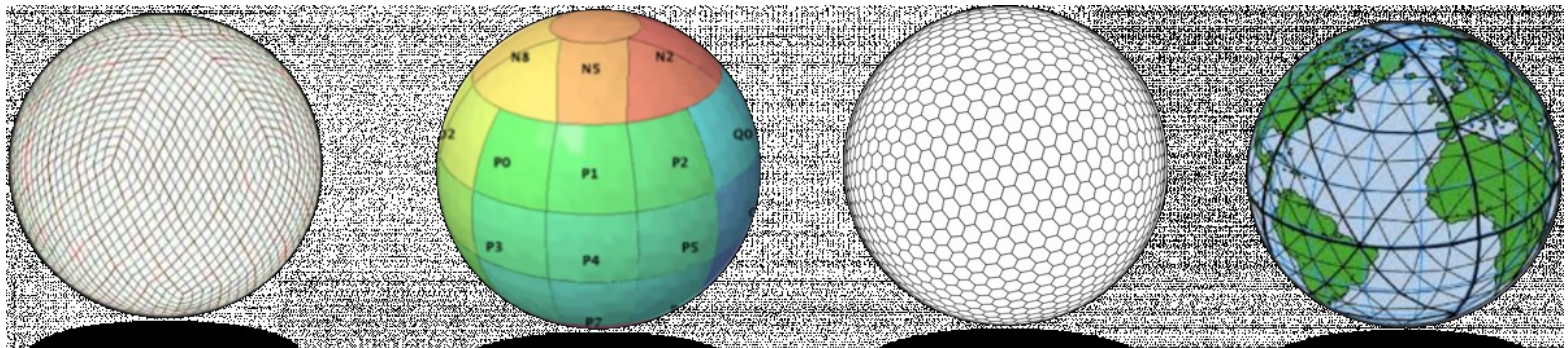
A new generation of decision-makers are also expecting systems that are not constrained by middle data integrators who post products in anticipation to their questions. These decision-makers have grown in sophistication – navigating time and space, mining the web for interesting information, and sharing their insights with others is everyday business. They are always connected; they are experts that demand choices and control over their own experience; they expect all the information now.

There remains a gap between expectation and the present reality [4]. Combining multiple sources of geospatial information - a necessary key step in the geospatial-intelligence cycle – is a very hard problem. Geospatial data integration on-demand is a grand challenge.

DISCRETE GLOBAL GRID AS A SOLUTION

A solution can only be achieved through the conversion of traditional data archives into standardized data architectures that support parallel processing in distributed and/or high performance compute environments. A common framework is required that will link very large multi-resolution and multi-domain datasets together and to enable the next generation of analytic processes to be applied. A solution must be capable of handling multiple data streams rather than being explicitly linked to a sensor or data type [5].

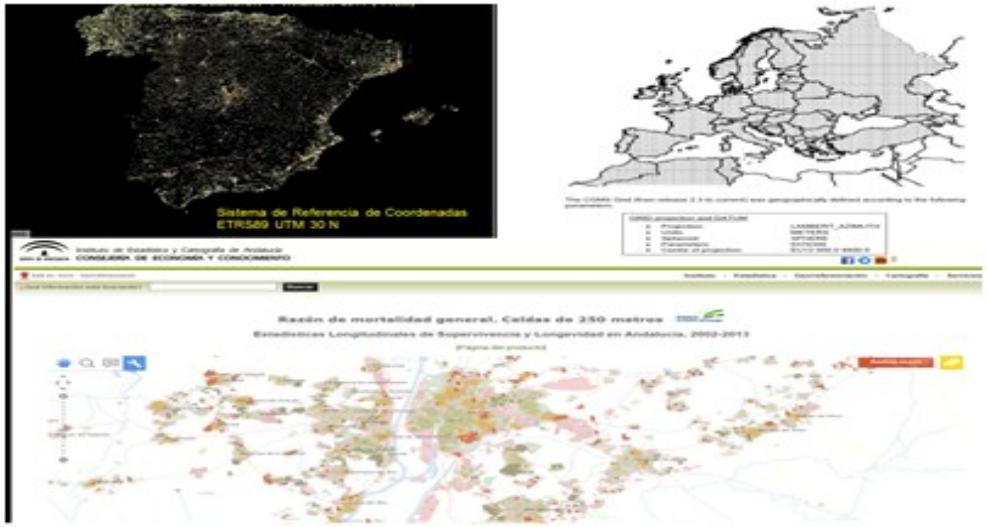
Success has been achieved using a framework called a discrete global grid system (DGGS). A DGGS is a form of Earth reference that, unlike its established counterpart the coordinate reference system that represents the Earth as a continual lattice of points, represents the Earth with a tessellation of nested cells [6]. Generally, a DGGS will exhaustively partition the globe in closely packed hierarchical tessellations, each cell representing a homogenous value, with a unique identifier or indexing that allows for linear ordering, parent-child operations, and nearest neighbour algebraic operations.



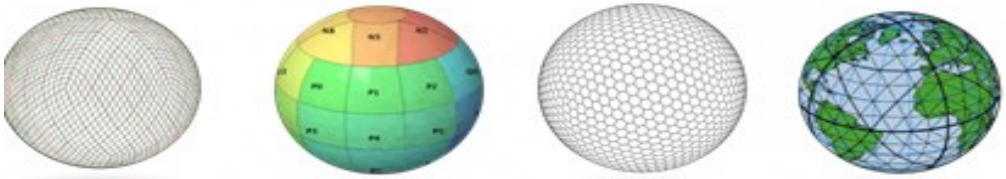
VENTAJAS **rejillas multiescalares/multidimensionales «Spatial tessellations»**

- Los datos originales **integrados espacialmente** en las **teselas** o **rejillas,** ofrecerán una **<u>estructura única y homogénea para todas las variables</u>** que facilitará subsiguientes procesos ligados con la **construcción de indicadores** y **análisis** basados en **Inteligencia Artificial y Machine Larning,** así como su **visualización y difusión web** en las diferentes **soluciones** a través de **clientes ligeros**.

- bien sobre **superficies proyectadas para regiones o espacios concretos (región, país, UE,,)** adecuando la proyección al espacio a teselar para garantizar que las teselas multiescalares difieran en superficie dentro de unos límites aceptables.



- o bien utilizando estructuras multiescalares (*nested*) grid a **escala global** (***Discrete Global Grid System*** –DGGS–), cuando se trate de problemáticas que necesiten integrar y operar con teselas uniformes en área a **escala global.**



Cloud-IA “"E-infraestructura de Información Ambiental para el Desarrollo de Soluciones Basadas en la Integración Espacial Normalizada de Variables"

NECESIDAD DE DISPONER DE **VARIALES AMBIENTALES YA INTEGRADAS ESPACIALMENTE**

Si bien los avances en los últimos años (INSPIRE, OPEN DATA) han permitido la **normalización de las variables ambientales individualmente** (normas, modelos, especificaciones, etc.), sigue pendiente la **necesaria integración espacial de las mismas de forma normalizada.** Así, a día de hoy, esta integración se lleva a cabo, solo por especialistas en TIGs, y de forma particular para cada problemática, adoptándose diferentes soluciones: integración en formatos raster, utilización de rejillas grid, etc. **La necesidad de una propuesta innovadora y homogénea para la integración espacial de las variables ambientales** más demandadas por la administración ambiental y la sociedad **es una clara necesidad no cubierta**

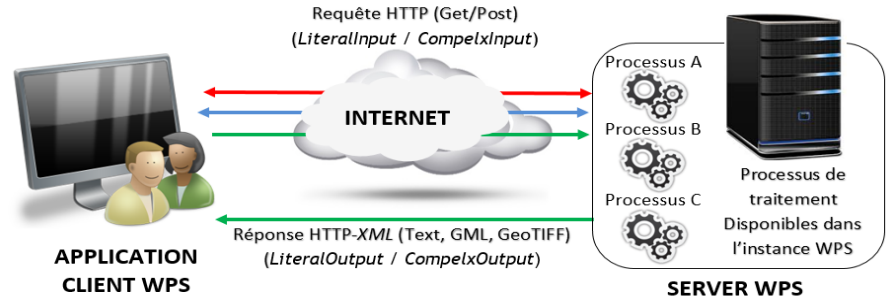
NECESIDAD DE **HERRAMIENTAS INFORMATICAS FLEXIBLES Y AMIGABLES EN UN ENTORNO CLOUD**

La Administración utiliza, para gestionar los problemas propios de la gestión ambiental**, datos en varios formatos, vectoriales y raster, así como herramientas informáticas sofisticadas** (programas SIG, gestores de bases de datos espaciales, ..) que hacen necesaria **la continua asistencia de los especialistas en los mismos**. El **desarrollo de herramientas informáticas más flexibles y amigables que se centren en soluciones bien definidas, personalizables para todos los usuarios, y que sean fácilmente utilizables por técnicos de la administración no especialistas es otra demanda no cubierta** y se entiende que **una e-infraestructura cloud innovadora** puede proporcionar las soluciones adecuadas a la Consulta

A new infrastructure for e-Science

- cloud computing
- unlimited database
- collaboration platform

OGC Web Processing Service (WPS)

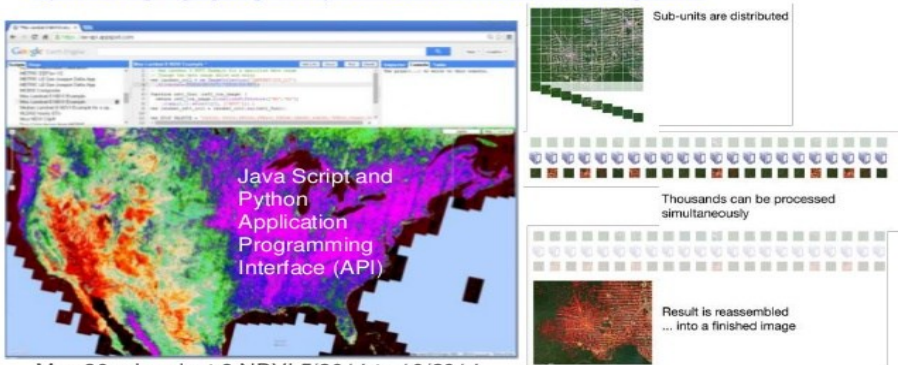


Légende :

- Le client envoi une requête ***GetCapabilities*** et le server répond (Liste des processus A, B, C)
- Le client sélectionne un traitement *processus C*, une requête ***DescribeProcess*** est envoyée et le server répand (Entrées & sorties du *processus C*)
- L'utilisateur choisie les entrées nécessaires et lance le traitement, une requête ***Execute*** est envoyé et le server répond par un document XML initial.

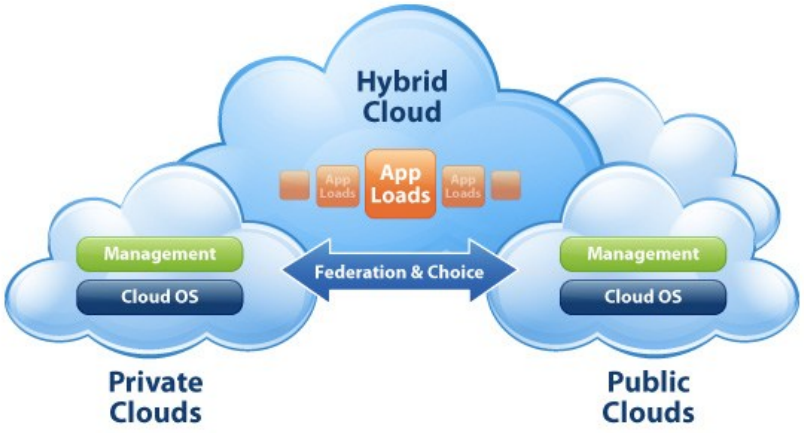
Google Earth Engine Cloud Computing

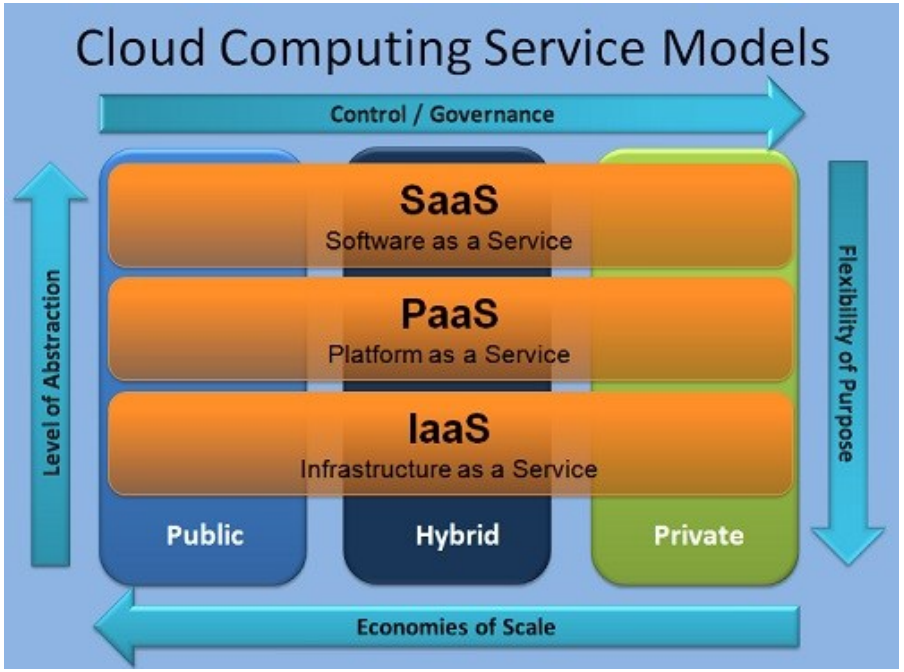
- Google has the entire archive of Landsat and MODIS imagery and CFSR, NLDAS, and downscaled NLDAS gridded weather data available for massive parallel processing in the cloud
- This technology has changed the paradigm of how we process and analyze satellite imagery and gridded weather data
- https://earthengine.google.org/#timelapse/v=40.86687,-117.50682,8.988,latLng&t=2.85



Max 30m Landsat 8 NDVI 5/2014 to 10/2014

MODELOS DE NUBE



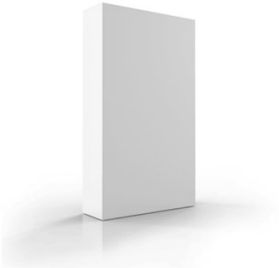


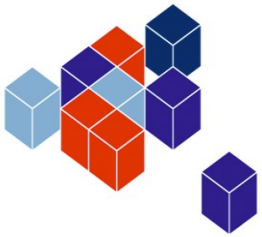
MODELOS DE COMPUTACIÓN CLOUD

A nivel de infraestructura (Openstack, EC2, Compute Engine)

Basado en contenedores (Openshift, ECS, Kubernetes Engine)

Basado en funciones (OpenWhisk,Kubeless, Lambda, Cloud Functions)



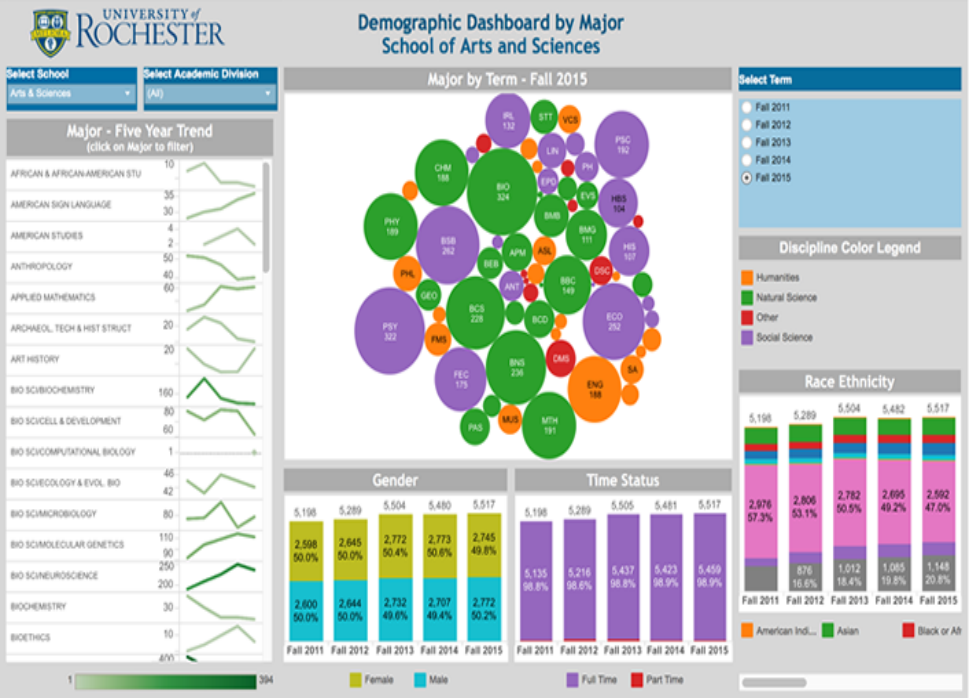


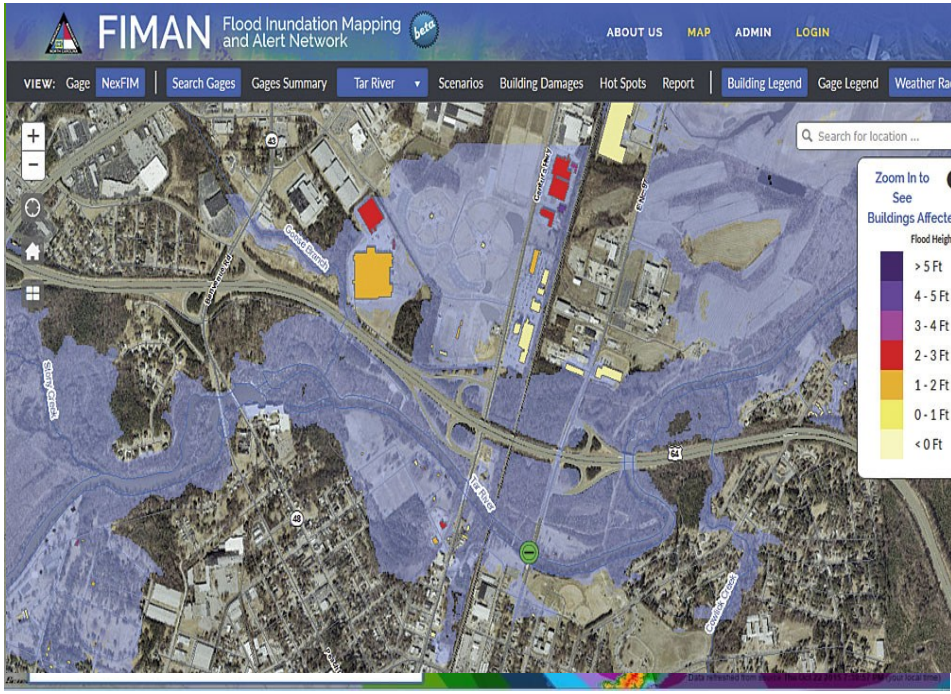
MONOLÍTICO

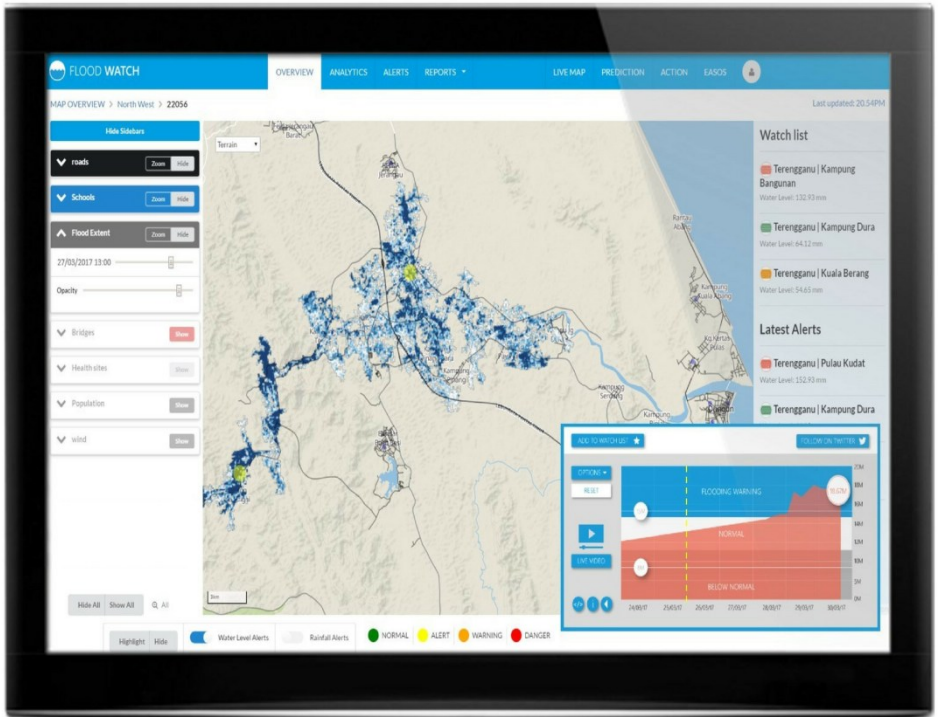
MICROSERVICIOS

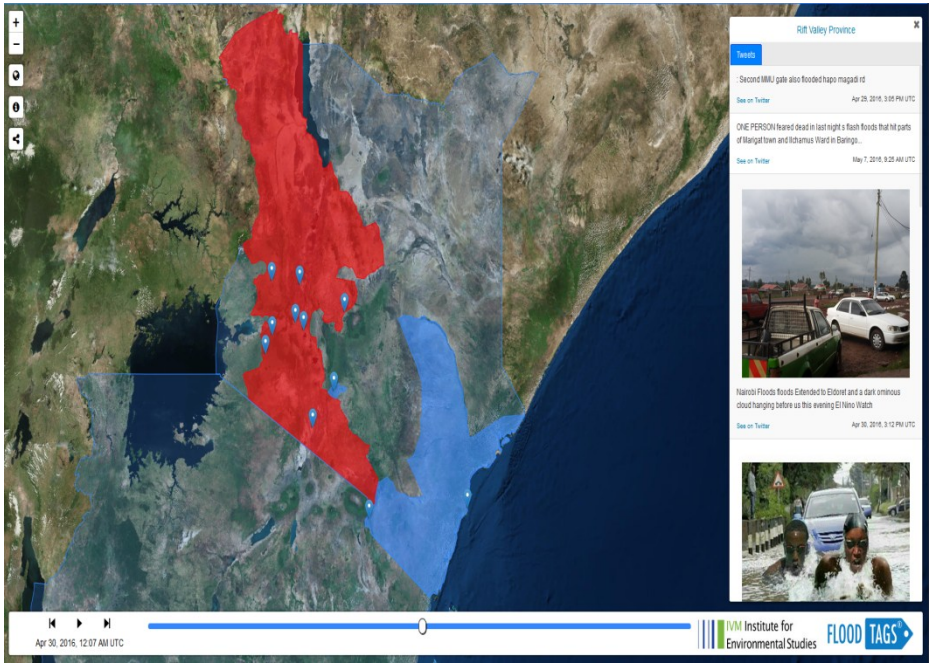
Una **"arquitectura de microservicios"** es un enfoque para desarrollar una aplicación software como una serie de pequeños servicios, cada uno ejecutándose de forma autónoma y comunicándose entre sí, por ejemplo, a través de peticiones HTTP a sus API.

Además cada uno es independiente y su código debe poder ser desplegado sin afectar a los demás. Incluso cada uno de ellos puede escribirse en un lenguaje de programación diferente, ya que solo exponen la API (una interfaz común, a la que le da igual el lenguaje de programación en la que el microservicio esté programado por debajo) al resto de microservicios.


UNIVERSITY of ROCHESTER
Demographic Dashboard by Major
School of Arts and Sciences
Select School
Arts & Sciences
Select Academic Division
(All)
Major - Five Year Trend
(click on Major to filter)
Major by Term - Fall 2015
Select Term
Fall 2011
Fall 2012
Fall 2013
Fall 2014
Fall 2015
Discipline Color Legend
Humanities
Natural Science
Other
Social Science
Race Ethnicity
Gender
Time Status
Female
Male
Full Time
Part Time


FIMAN
Flood Inundation Mapping and Alert Network
ABOUT US
MAP
ADMIN
LOGIN
VIEW: Gage
NexFIM
Search Gages
Gages Summary
Tar River
Scenarios
Building Damages
Hot Spots
Report
Building Legend
Gage Legend
Search for location ...
Zoom In to See Buildings Affecte
> 5 Ft
4 - 5 Ft
3 - 4 Ft
2 - 3 Ft
1 - 2 Ft
0 - 1 Ft
< 0 Ft


FLOOD WATCH
OVERVIEW
ANALYTICS
ALERTS
REPORTS
LIVE MAP
PREDICTION
ACTION
EASOS
Watch list
Terengganu | Kampung Bangunan
Terengganu | Kampung Dura
Terengganu | Kuala Berang
Latest Alerts
Terengganu | Pulau Kudat
Terengganu | Kampung Dura
FLOODING WARNING
NORMAL
BELOW NORMAL


Rift Valley Province
Tweets
Apr 30, 2016, 12:07 AM UTC
IVM Institute for Environmental Studies
FLOOD TAGS

Cloud-IA “"E-infraestructura de Información Ambiental para el Desarrollo de Soluciones Basadas en la Integración Espacial Normalizada de Variables"

OBJETO: Creación de una e-infraestructura de información ambiental integrada espacialmente, con gran capacidad de almacenamiento, gestión, análisis y difusión web en un entorno cloud, que permita la generación de soluciones a medida para gestores, investigadores y ciudadanía no expertos en TIGs, en ámbitos relacionados con la gestión ambiental y territorial.

- **REQUERIMIENTOS TEMÁTICOS DE LOS DATOS Y FUNCIONALIDADES DE LAS APLICACIONES A DESARROLLAR**

❑ REQUERIMIENTOS TEMÁTICOS DE LOS DATOS Y FUNCIONALIDADES DE LAS APLICACIONES A DESARROLLAR

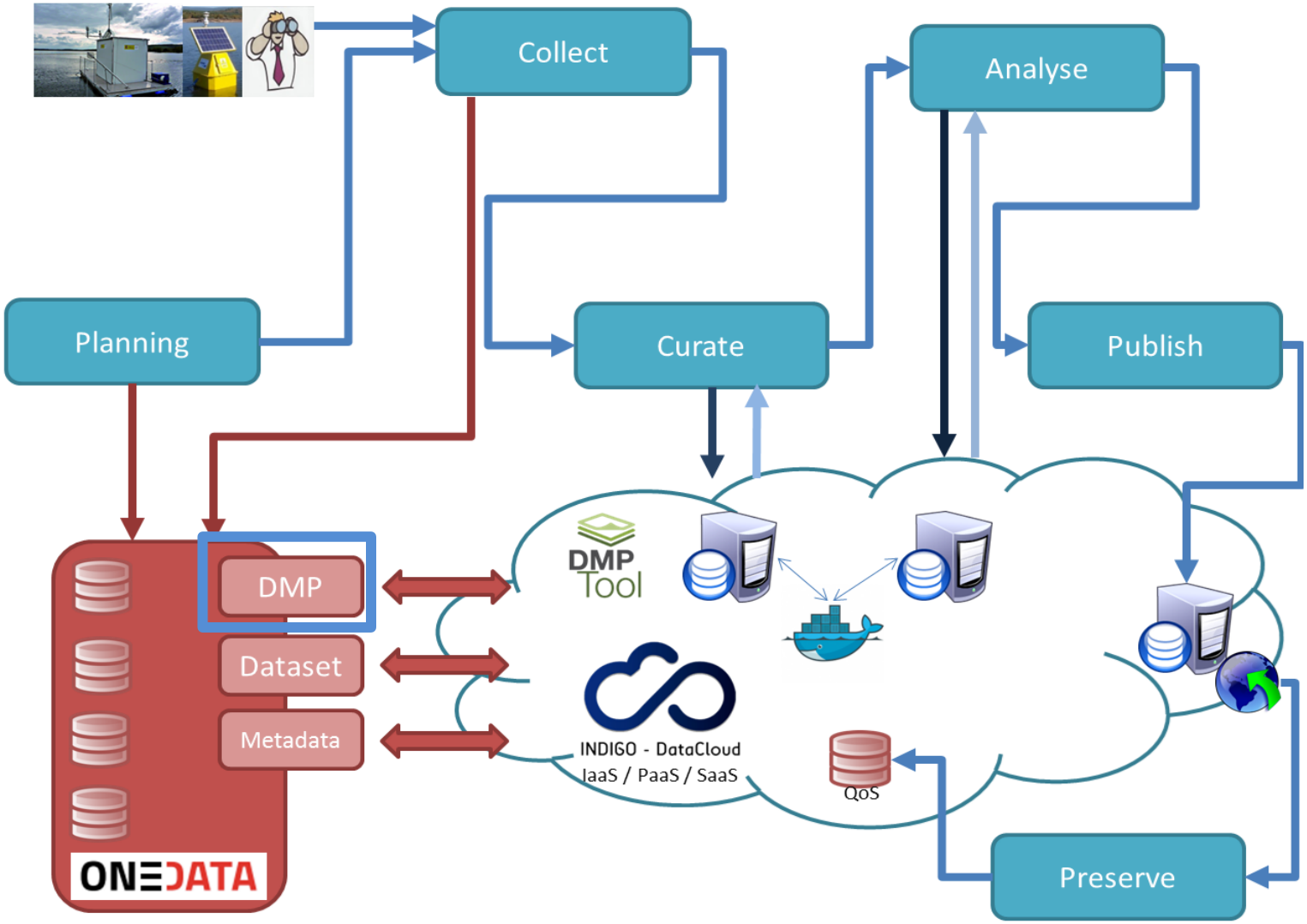
A la e-infraestructura a desarrollar en la "nube" **para el almacenamiento masivo de datos ambientales normalizados espacialmente, su acceso y difusión web, así como el desarrollo de las "soluciones",** se incorporarán **datos de muy diversa índole**:

- ✓ **DATOS ORIGINALES de fuentes externas (generalmente públicas)** para cada una de las variables medioambientales (usos de suelos, catastro, población, etc.).
- ✓ **DATOS NUEVOS que se generen:**
  - **por la transformación de los anteriores** para adscribir **las variables temáticas a una estructura de teselas/rejillas**
  - **por el análisis integrado de los datos anteriores** (indicadores, índices, clasificaciones estadísticas, etc..)
- ✓ Potencialmente, **DATOS QUE GENEREN LOS USUARIOS haciendo uso de los servicios del sistema** y que puedan ser incorporados al sistema.

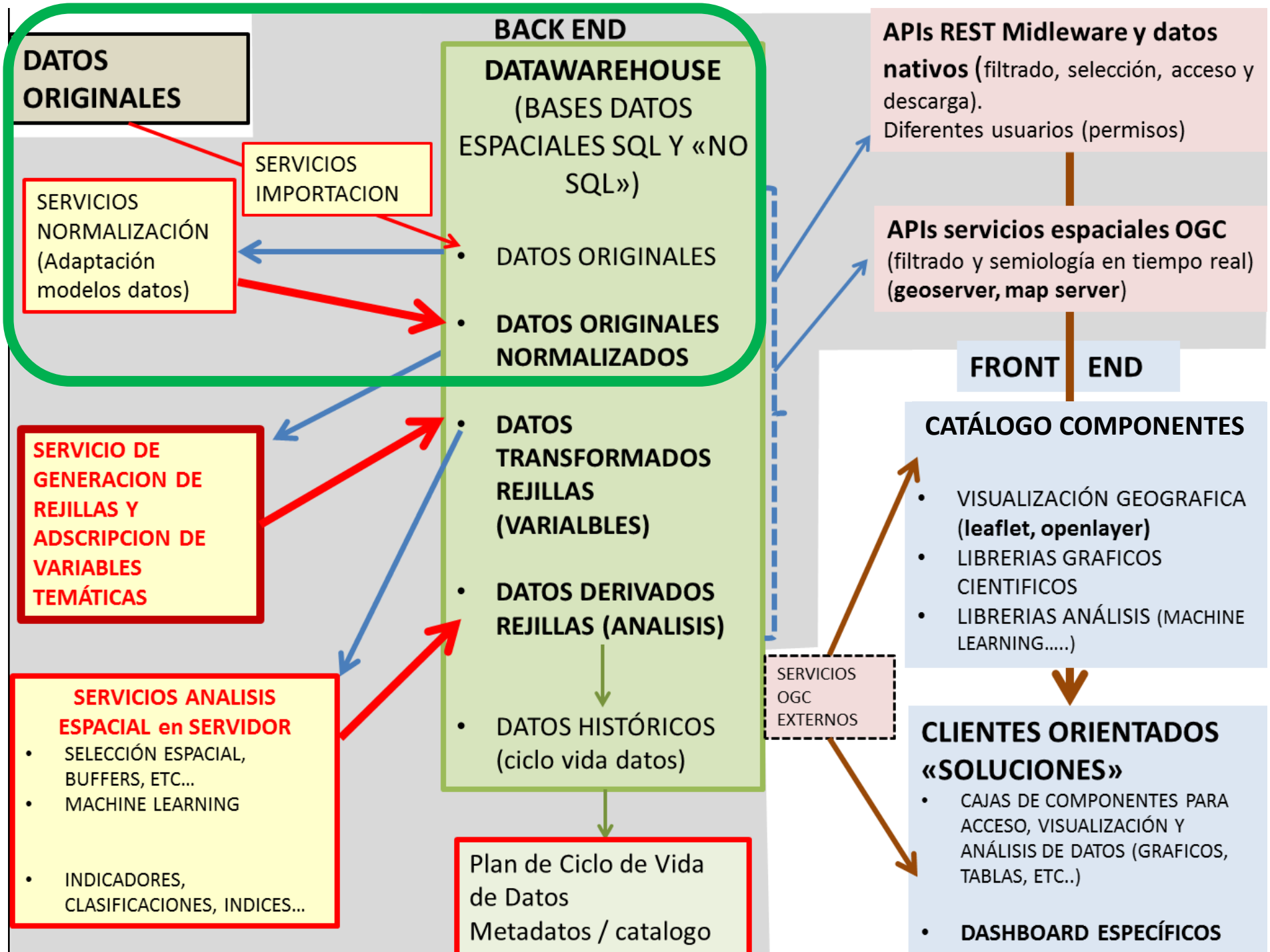
Por lo tanto, estos geodatos, al provenir mayoritariamente de la Administración, se deberán adaptar a las especificaciones de la **Open Knowledge Foundation (2012),** para la cual «OPEN DATA» son los datos que pueden ser **utilizados libremente, reutilizados y redistribuidos por cualquier persona**, permitiendo la creación productos derivados reconociendo su procedencia:

- **Directiva 2007/2/CE (INSPIRE)**
- ***Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno***
- **Ley 14/2010 sobre Infraestructuras y los Servicios de Información Geográfica**

Por otra parte, un conjunto de iniciativas europeas ("**European Open Science Cloud"**) están claramente relacionadas con los objetivos de este proyecto y enfatizan la apropiada **gestión del CICLO DE VIDA DEL DATO** (*plan, collect, curate, analyze, ingest-publish and preserve*) con el impulso de los **Planes de Gestión de Datos** (o Data Management Plans), así como la generación de **datos "FAIR"** (Encontrables, Accesibles, Interoperables y Reusables).



Aguilar Gomez, F. (2017): **Data Management in a Cloud Framework: application to the LifeWatch ESFRI.** Tesis doctoral. Universidad de Cantabria. Accesible en: https://repositorio.unican.es/xmlui/handle/10902/13073


BACK END
DATOS ORIGINALES
SERVICIOS IMPORTACION
SERVICIOS NORMALIZACIÓN (Adaptación modelos datos)
DATAWAREHOUSE (BASES DATOS ESPACIALES SQL Y «NO SQL»)
DATOS ORIGINALES
DATOS ORIGINALES NORMALIZADOS
DATOS TRANSFORMADOS REJILLAS (VARIALBLES)
DATOS DERIVADOS REJILLAS (ANALISIS)
DATOS HISTÓRICOS (ciclo vida datos)
SERVICIO DE GENERACION DE REJILLAS Y ADSCRIPCION DE VARIABLES TEMÁTICAS
SERVICIOS ANALISIS ESPACIAL en SERVIDOR
SELECCIÓN ESPACIAL, BUFFERS, ETC…
MACHINE LEARNING
INDICADORES, CLASIFICACIONES, INDICES…
Plan de Ciclo de Vida de Datos
Metadatos / catalogo
APIs REST Midleware y datos nativos (filtrado, selección, acceso y descarga).
Diferentes usuarios (permisos)
APIs servicios espaciales OGC (filtrado y semiología en tiempo real) (geoserver, map server)
FRONT END
CATÁLOGO COMPONENTES
VISUALIZACIÓN GEOGRAFICA (leaflet, openlayer)
LIBRERIAS GRAFICOS CIENTIFICOS
LIBRERIAS ANÁLISIS (MACHINE LEARNING…..)
SERVICIOS OGC EXTERNOS
CLIENTES ORIENTADOS «SOLUCIONES»
CAJAS DE COMPONENTES PARA ACCESO, VISUALIZACIÓN Y ANÁLISIS DE DATOS (GRAFICOS, TABLAS, ETC..)
DASHBOARD ESPECÍFICOS

SERVICIOS DE IMPORTACIÓN Y MODELIZACIÓN DE DATOS

- **1- Relieve.** Incluye información procedente de cobertura LIDAR: MDT, MDE, pendientes y orientaciones.
- **2- Catastro.** Información catastral procedente de los formatos *shapefile* y CAT: polígonos, parcelas, construcciones, número de plantas, etc. También incluye la información del formato Atom (INSPIRE) relativa a los edificios y usos principales de los mismos (buildings)
- **SIPNA (Sistema de Información sobre el Patrimonio Natural de Andalucía).** Incluye hábitats, vegetación, biogeografía, lusos del suelo, ocupación del suelo y el seguimiento del territorio desde 1956 a 2013
- **Catálogo de Montes Públicos.** Incluye información relativa a las parcelas, titularidad, etc…de los montes públicos de Andalucía.
- **Población.** Incluye información poblacional procedente del IECA: población total, por edades, por sexos, por origen, así como variables socioeconómicas espacializadas: afiliación a la seguridad social, pensiones contributivas de la SS, razón de mortalidad, longevidad, número de propietarios con segunda residencia, tasas de desempleo, nivel educativo, etc...
- **Empresas e infraestructuras energéticas.** Incluye información del Directorio de empresas y establecimientos con actividad económica de más de 50 asalariados. Empresas sometidas a directiva IPPC. Empresas Públicas Locales y Empresas agroalimentarias (DERA. Tejido económico productivo).
- Infraestructuras energéticas (DERA). Incluye información relativa a las infraestructuras básicas energéticas de la región : plantas de generación según tipologías, gaseoducots, etc…
- **Espacios Naturales Protegidos:** límites espaciales de todas las figuras de ptotección, zonificación PORNs, etc.
- **Límites administrativos y afecciones territoriales:** términos municipales, DPMT, DP Hidráulico, Líneas de base recta, línea de costa oficial, Zonificación urbanística y de Ordenación del Territorio, demarcaciones de servicios públicos (DERA G_17 divisiones administrativas).Cuadrícula Minera (REDIAM).
- **Servicios Públicos**. Información procedente del Inventario de sedes y equipamientos Junta de Andalucía recogidos en las capas DERA (G_16 Servicios).
- **Hidrografía:** red hidrográfica, inventario de balsas, humedales y balsas de agua, cuencas y subcuencas (oficial y natural), acuíferos. Infraestructuras hidráulicas (canales, embalses, depuradoras, etc.) (DERA).
- **11. Litoral.** Caracterizaciones de las diferentes lineas de costa, unidades geomorfológicas, tasas de erosión, etc.
- **12.-Clima.** Estaciones, Precipitación, temperaturas, proyecciones escenarios cambio climático, etc.
- **Geología, geomorfología, suelos.** Mapa geomorfoedáfico, suelos, cronología geológica, unidades geoestructurales, unidades litológicas, seguimiento anual d ellas pérdidas de suelo por erosión.
- **Paisaje. Unidades y catálogos.**
- **Asentamientos urbanos de Andalucía (DERA. IECA).**
- **Viario e infraestructuras de transporte.** Red de carreteras, caminos, cortafuegos, ferrocarril, vías pecuarias, CDAU, carriles bici, senderos. Puertos, aeropuertos (DERA y CDAU).
- **Instalaciones de Tratamiento de Residuos sólidos Urbanos.** DERA.

- **Riesgos.** Zonas inundación fluvial y escenarios, zonas inundación marina, serie histórica de áreas recorridas por el fuego, mapa de conatos e incendios del INFOCA, mapas de modelos de combustibles forestales, cortafuegos, …
- **Dehesas:** Censo de dehesas. Formaciones adehesadas.
- **Series temporales de parámetros obtenidos a partir de imágenes de satélites:** NDVI, SST, Clorofilas, Turbidez, etc.
- **Zonas urbanas y rurales. EUROSTAT.**
- **SIGPAC**, Parcelas del SIGPAC
- **Patrimonio:** catálogo de bienes patrimoniales en Andalucía, etc…

Servicios de importación (datos originales)

2. HIDROGRAFÍA

2.1. RED HIDROGRÁFICA SUPERFICIAL

MODELO DE DATOS	RESOLUCIÓN	FORMATOS	FUENTE	PRODUCTOR	WEB DESCARGA	WEB METADATOS	VOLUMEN APROXIMADO	ACTUALIZACION EN EL SISTEMA
VECTORIAL	*1:100000*	*SHAPEFILE*	*REDIAM*	*CMAYOT*	http://descargasrediam.cica.es/repo/s/RUR?path=%2F04_RECURSOS_NATURALES%2F04_AGUAS%2F01_SUPERFICIALES%2F00_SUPERFICIALES%2FRed_Hidrog_Agues_Sup	http://laboratorio.rediam.cica.es/geonetwork/srv/spa/search#|e260f066-5c4c-4737-aa1e-ac3c927467c3	*28,2 MB*	*VARIABLE (SEGÚN ACTUALIZACIÓN EN REDIAM)*

OBSERVACIONES

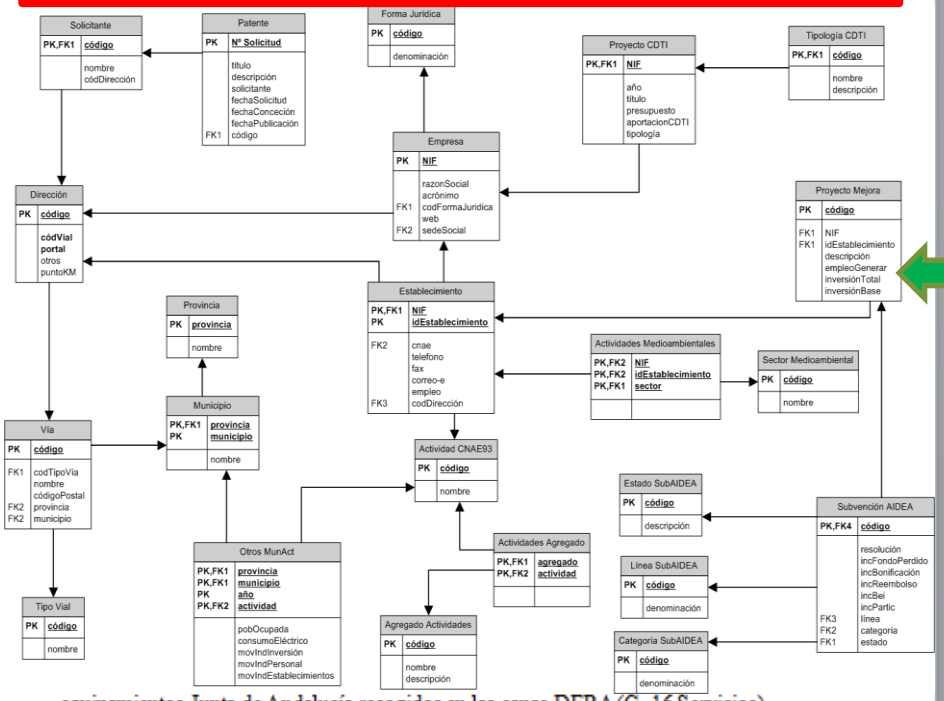
Red hidrográfica procedente del Mapa Topográfico de Andalucía 1:100.000 y adaptada para su integración en la base de datos de Información General de Aguas Superficiales. Año 2008.

Las capas que contiene son: red_hidrografica, Red_tramos. La primera representa las corrientes fluviales continuas mientras que la segunda representa los tramos de dicha red y jerarquías.

Más información:
http://descargasrediam.cica.es/repo/s/RUR?path=%2F04_RECURSOS_NATURALES%2F04_AGUAS%2F01_SUPERFICIALES%2F00_SUPERFICIALES%2FRed_Hidrog_Agues_Sup%2FModeloDatos

SERVICIOS DE IMPORTACIÓN Y MODELIZACIÓN DE DATOS

Servicios de normalización (variables modeladas)



equipamientos Junta de Andalucía recogidos en las capas DERA (G_16 Servicios).

- **Hidrografía:** red hidrográfica, inventario de balsas, humedales y balsas de agua, cuencas y subcuencas (oficial y natural), acuíferos. Infraestructuras hidráulicas (canales, embalses, depuradoras, etc.) (DERA).
- **11. Litoral.** Caracterizaciones de las diferentes lineas de costa, unidades geomorfológicas, tasas de erosión, etc.
- **12.-Clima.** Estaciones, Precipitación, temperaturas, proyecciones escenarios cambio climático, etc.
- **Geología, geomorfología, suelos.** Mapa geomorfoedáfico, suelos, cronología geológica, unidades geoestructurales, unidades litológicas, seguimiento anual d ellas pérdidas de suelo por erosión.
- **Paisaje. Unidades y catálogos.**
- **Asentamientos urbanos de Andalucía (DERA. IECA).**
- **Viario e infraestructuras de transporte.** Red de carreteras, caminos, cortafuegos, ferrocarril, vías pecuarias, CDAU, carriles bici, senderos. Puertos, aeropuertos (DERA y CDAU).
- **Instalaciones de Tratamiento de Residuos sólidos Urbanos.** DERA.

- **Riesgos.** Zonas inundación fluvial y escenarios, zonas inundación marina, serie histórica de áreas recorridas por el fuego, mapa de conatos e incendios del INFOCA, mapas de modelos de combustibles forestales, cortafuegos, …
- **Dehesas:** Censo de dehesas. Formaciones adehesadas.
- **Series temporales de parámetros obtenidos a partir de imágenes de satélites:** NDVI, SST, Clorofilas, Turbidez, etc.
- **Zonas urbanas y rurales. EUROSTAT.**
- **SIGPAC.** Parcelas del SIGPAC
- **Patrimonio:** catálogo de bienes patrimoniales en Andalucía, etc…

Servicios de importación (datos originales)

2. HIDROGRAFÍA

2.1. RED HIDROGRÁFICA SUPERFICIAL

MODELO DE DATOS	RESOLUCIÓN	FORMATOS	FUENTE	PRODUCTOR	WEB DESCARGA	WEB METADATOS	VOLUMEN APROXIMADO	ACTUALIZACION EN EL SISTEMA
VECTORIAL	*1:100000*	*SHAPEFILE*	*REDIAM*	*CMAYOT*	http://descargasrediam.cica.es/repo/s/RUR?path=%2F04_RECURSOS_NATURALES%2F04_AGUAS%2F01_SUPERFICIALES%2F00_SUPERFICIALES%2FRed_Hidrog_Aguas_Sup	http://laboratorio.rediam.cica.es/geonetwork/srv/spa/search#|e260f066-5c4c-4737-aa1e-ec3c927467c3	*28,2 MB*	*VARIABLE (SEGÚN ACTUALIZACIÓN EN REDIAM)*

OBSERVACIONES

Red hidrográfica procedente del Mapa Topográfico de Andalucía 1:100.000 y adaptada para su integración en la base de datos de Información General de Aguas Superficiales. Año 2008.

Las capas que contiene son: red_hidrografica, Red_tramos. La primera representa las corrientes fluviales continuas mientras que la segunda representa los tramos de dicha red y jerarquías.

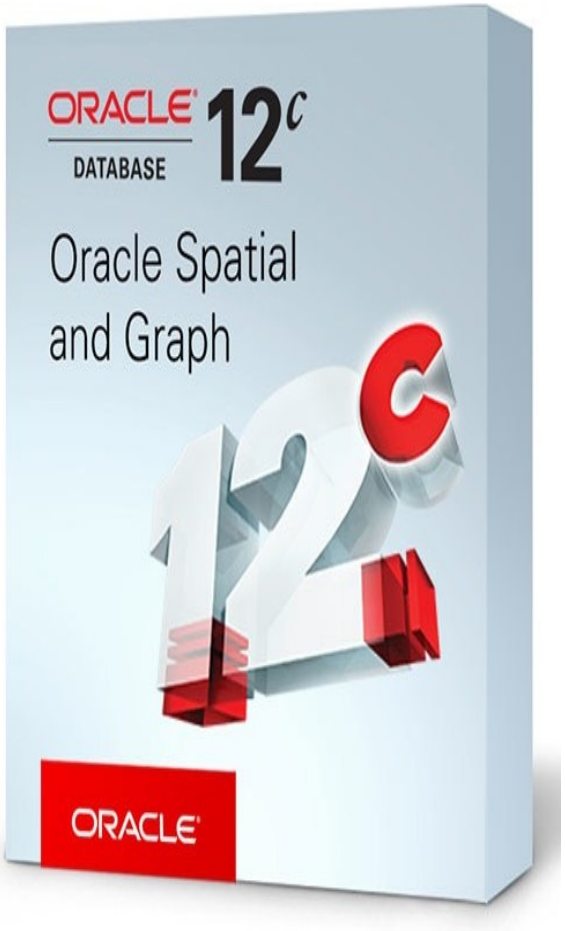
Más información:
http://descargasrediam.cica.es/repo/s/RUR?path=%2F04_RECURSOS_NATURALES%2F04_AGUAS%2F01_SUPERFICIALES%2F00_SUPERFICIALES%2FRed_Hidrog_Aguas_Sup%2FModeloDatos

LOS SISTEMAS GESTORES DE BASES DE DATOS: «ESPACIALES» Y «NoSQL»

PostgresSQL y PostGIS

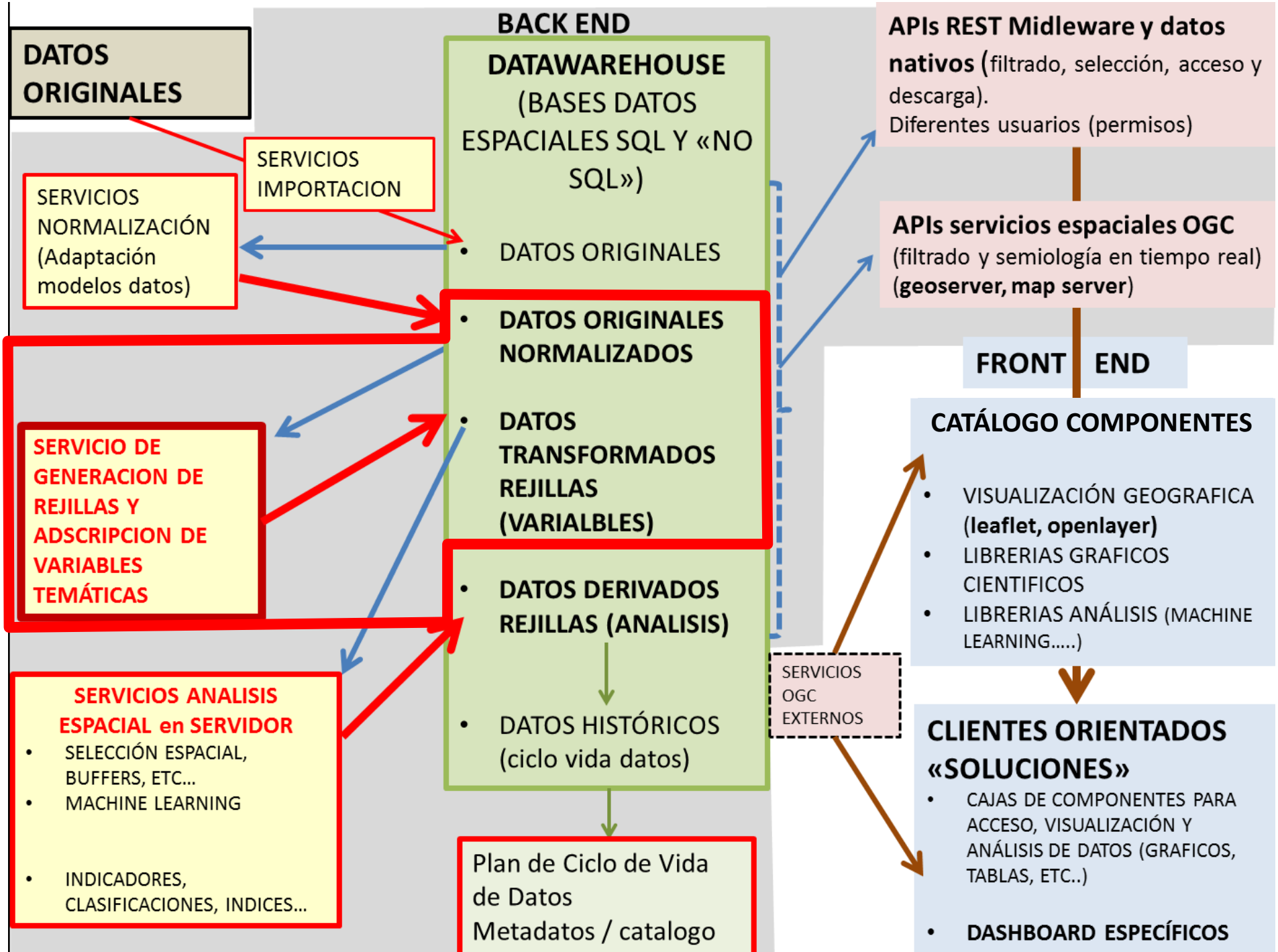
- Es un **módulo** que añade soporte de objetos geográficos a la base de datos objeto-relacional PostgreSQL, convirtiéndola en una **base de datos espacial** para su utilización en SIG.

- PostGIS ha sido certificado en 2006 por el Open Geospatial Consortium (OGC) lo que garantiza la interoperabilidad con otros sistemas .



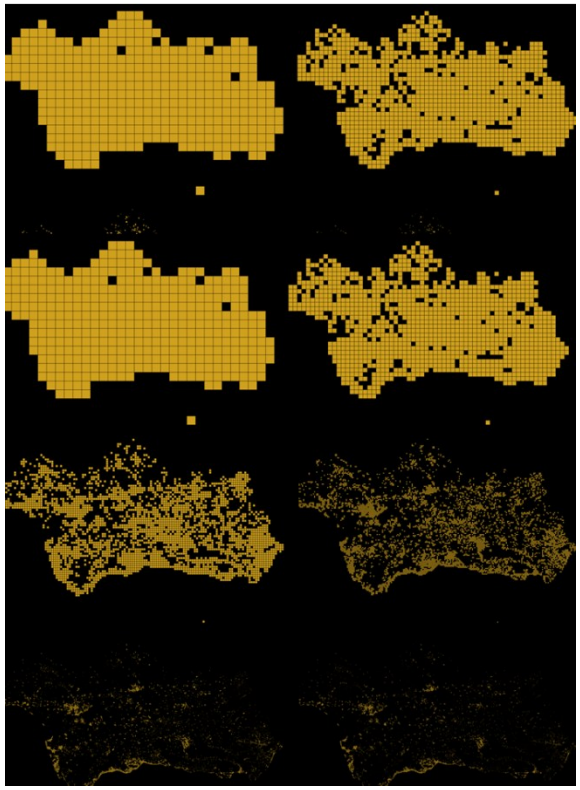
Tipos de bases de datos NoSQL

Documentales	Datos semi-estructurados en documentos (XML, YAML, JSON y BSON)	elastic, mongoDB
Grafo	Datos estructurados como nodos relacionados entre si	Neo4j, OPENLINK SOFTWARE
Clave / valor	Datos estructurados como clave / valor	redis, memCached
Familia de columnas	Datos estructurados en columnas donde cada fila puede tener una configuración diferente	cassandra, hadoop

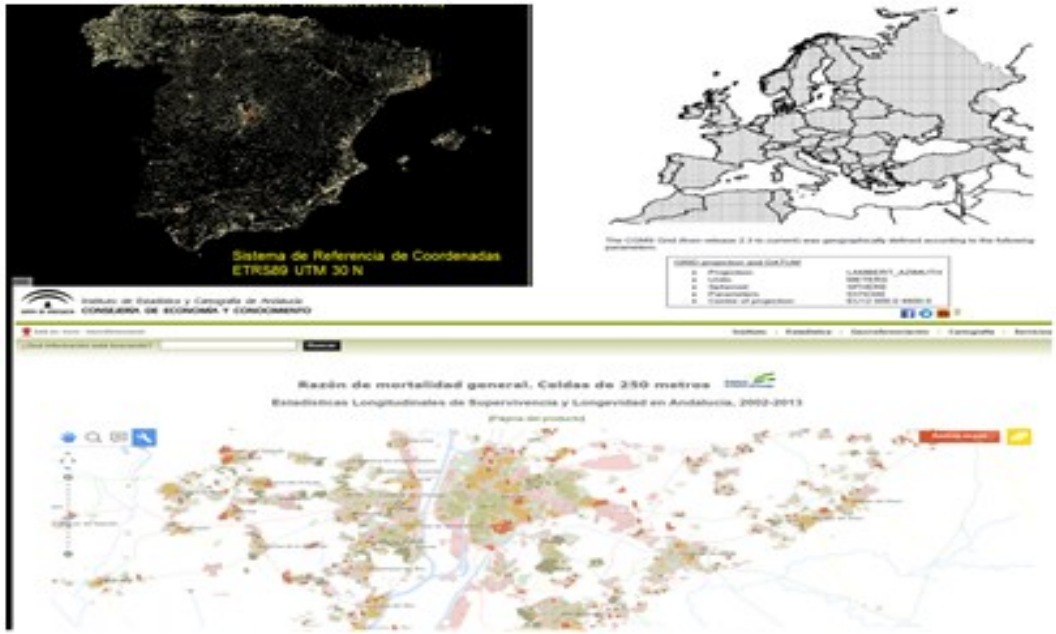

DATOS ORIGINALES
BACK END
DATAWAREHOUSE (BASES DATOS ESPACIALES SQL Y «NO SQL»)
SERVICIOS IMPORTACION
SERVICIOS NORMALIZACIÓN (Adaptación modelos datos)
DATOS ORIGINALES
DATOS ORIGINALES NORMALIZADOS
DATOS TRANSFORMADOS REJILLAS (VARIALBLES)
SERVICIO DE GENERACION DE REJILLAS Y ADSCRIPCION DE VARIABLES TEMÁTICAS
DATOS DERIVADOS REJILLAS (ANALISIS)
DATOS HISTÓRICOS (ciclo vida datos)
SERVICIOS ANALISIS ESPACIAL en SERVIDOR
SELECCIÓN ESPACIAL, BUFFERS, ETC…
MACHINE LEARNING
INDICADORES, CLASIFICACIONES, INDICES…
Plan de Ciclo de Vida de Datos
Metadatos / catalogo
APIs REST Midleware y datos nativos (filtrado, selección, acceso y descarga).
Diferentes usuarios (permisos)
APIs servicios espaciales OGC (filtrado y semiología en tiempo real) (geoserver, map server)
FRONT END
CATÁLOGO COMPONENTES
VISUALIZACIÓN GEOGRAFICA (leaflet, openlayer)
LIBRERIAS GRAFICOS CIENTIFICOS
LIBRERIAS ANÁLISIS (MACHINE LEARNING…..)
SERVICIOS OGC EXTERNOS
CLIENTES ORIENTADOS «SOLUCIONES»
CAJAS DE COMPONENTES PARA ACCESO, VISUALIZACIÓN Y ANÁLISIS DE DATOS (GRAFICOS, TABLAS, ETC..)
DASHBOARD ESPECÍFICOS

SERVICIOS DE GENERACIÓN DE REJILLAS Y ADSCRIPCIÓN DE VARIABLES

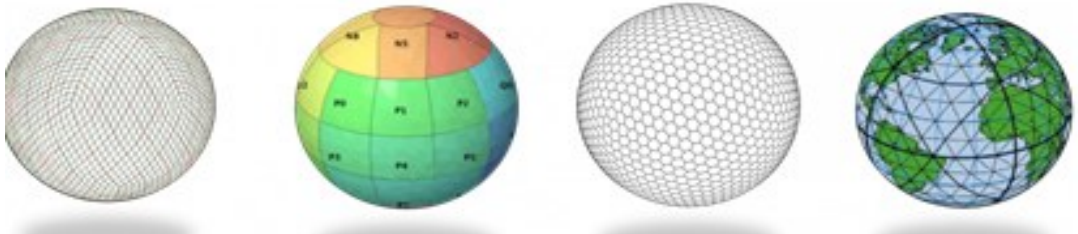
1.- Definición de la rejilla multiescalar a utilizar.



- bien sobre **superficies proyectadas para regiones o espacios concretos (región, país, UE,,)** adecuando la proyección al espacio a teselar para garantizar que las teselas multiescalares difieran en superficie dentro de unos límites aceptables.



- o bien utilizando estructuras multiescalares (*nested*) grid a **escala global** (***Discrete Global Grid System*** –DGGS–), cuando se trate de problemáticas que necesiten integrar y operar con teselas uniformes en área a **escala global.**

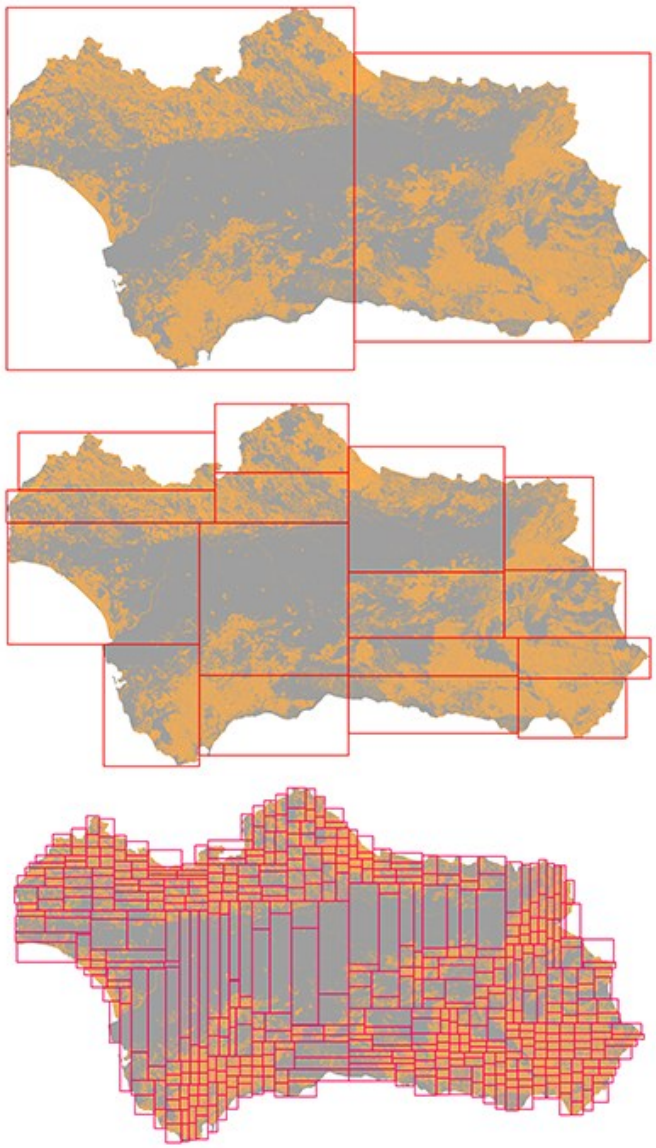


SERVICIOS DE GENERACIÓN DE REJILLAS Y ADSCRIPCIÓN DE VARIABLES

1.- Definición de la rejilla multiescalar a utilizar.

2.- Importación de datos normalizados de las fuentes seleccionadas para la elección de las variables a adscribir a las rejillas multiescalares.

3.- Reparto de los datos normalizados seleccionados en paquetes de trabajo para su procesado en paralelo.



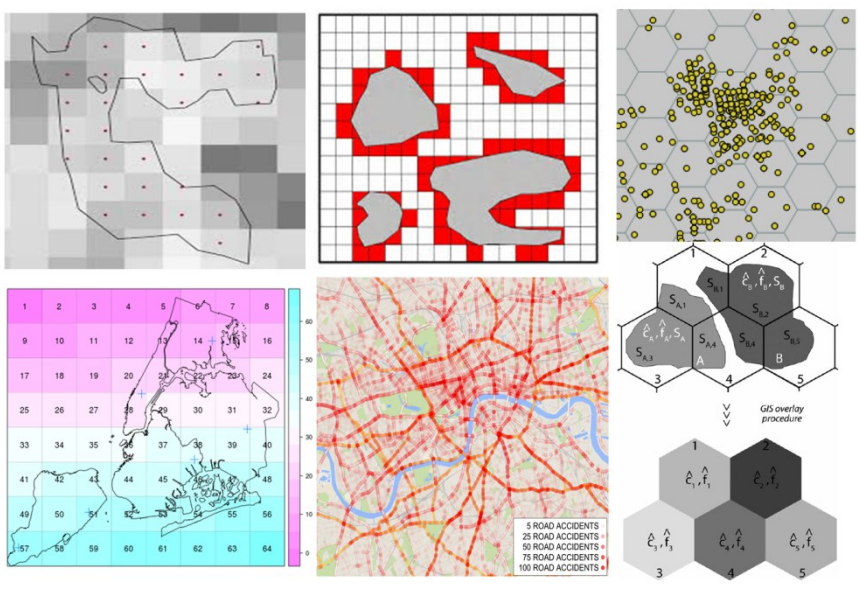
SERVICIOS DE GENERACIÓN DE REJILLAS Y ADSCRIPCIÓN DE VARIABLES

1.- Definición de la rejilla multiescalar a utilizar.

2.- Importación <u>de datos normalizados de las fuentes seleccionadas</u> para la elección de las variables a adscribir a las rejillas multiescalares.

3.- Reparto de los datos normalizados seleccionados en paquetes de trabajo para su <u>procesado en paralelo</u>.

4.- Elección de la variable temática <u>y método de adscripción.</u>



El primer grupo lo constituirían las **fuentes que siguen el modelo raster, pero con diferente granularidad de rejilla, origen y/o nivel de** detalle.

Las **fuentes que proporcionan información sobre <u>variables discretas</u> asociados a <u>puntos</u>** . El método de adscripción de estos datos no es compleja ya que se basaría en una relación espacial de inclusión en la rejilla de cada nivel escalar.

Las **fuentes que proporcionan información sobre <u>variables contínuas</u> asociados a <u>puntos</u>.** El método de adscripción de estos contemplaría una mayor complejidad si se quiere incorporar en el algoritmo de adscripción la información tras complejos procesos de interpolación.

Las **fuentes que proporcionan información sobre <u>variables asociadas a líneas.</u>** Son representaciones <u>unidimensionales</u> de un fenómeno que se extiende sobre el territorio. En este caso, una opción podría ser la adscripción de las variables temáticas en función de la longitud de cada categorías de líneas que intescectan la rejilla.

Las **fuentes que proporcionan información sobre <u>variables asociadas a polígonos</u>** Son representaciones <u>bidimensionales</u> de un fenómeno que se extiende sobre el territorio. En este caso, una opción podría ser la adscripción de las variables temáticas en función del área de cada categoría de polígonos que intescectna la rejilla.

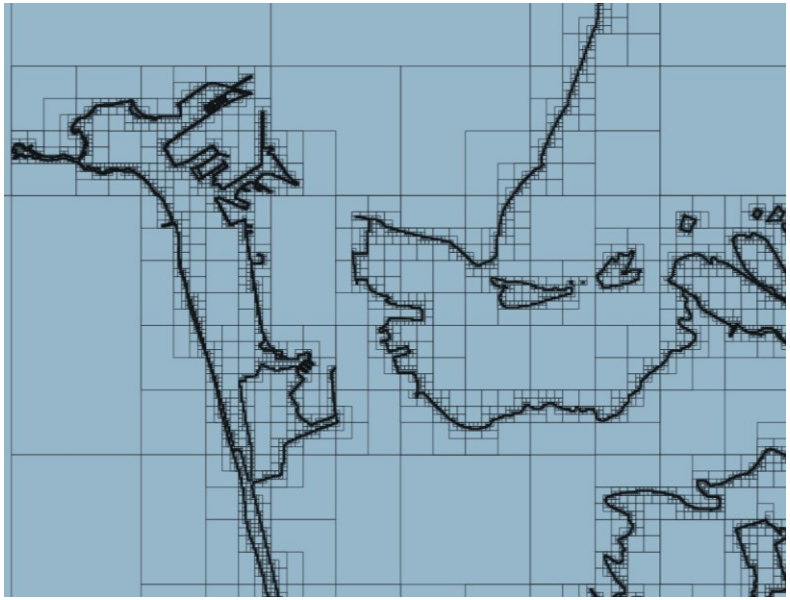
SERVICIOS DE GENERACIÓN DE REJILLAS Y ADSCRIPCIÓN DE VARIABLES

1.- Definición de la rejilla multiescalar a utilizar.

2.- Importación de datos normalizados de las fuentes seleccionadas para la elección de las variables a adscribir a las rejillas multiescalares.

3.- Reparto de los datos normalizados seleccionados en paquetes de trabajo para su procesado en paralelo.

4.- Elección de la variable temática y método de adscripción.



El primer grupo lo constituirían las **fuentes que siguen el modelo raster, pero con diferente granularidad de rejilla, origen y/o nivel de** detalle.

Las **fuentes que proporcionan información sobre variables discretas asociados a puntos** . El método de adscripción de estos datos no es compleja ya que se basaría en una relación espacial de inclusión en la rejilla de cada nivel escalar.

Las **fuentes que proporcionan información sobre variables contínuas asociados a puntos.** El método de adscripción de estos contemplaría una mayor complejidad si se quiere incorporar en el algoritmo de adscripción la información tras complejos procesos de interpolación.

Las **fuentes que proporcionan información sobre variables asociadas a líneas.** Son representaciones unidimensionales de un fenómeno que se extiende sobre el territorio. En este caso, una opción podría ser la adscripción de las variables temáticas en función de la longitud de cada categorías de líneas que intescectan la rejilla.

Las **fuentes que proporcionan información sobre variables asociadas a polígonos** Son representaciones bidimensionales de un fenómeno que se extiende sobre el territorio. En este caso, una opción podría ser la adscripción de las variables temáticas en función del área de cada categoría de polígonos que intescectna la rejilla.

SERVICIOS DE GENERACIÓN DE REJILLAS Y ADSCRIPCIÓN DE VARIABLES

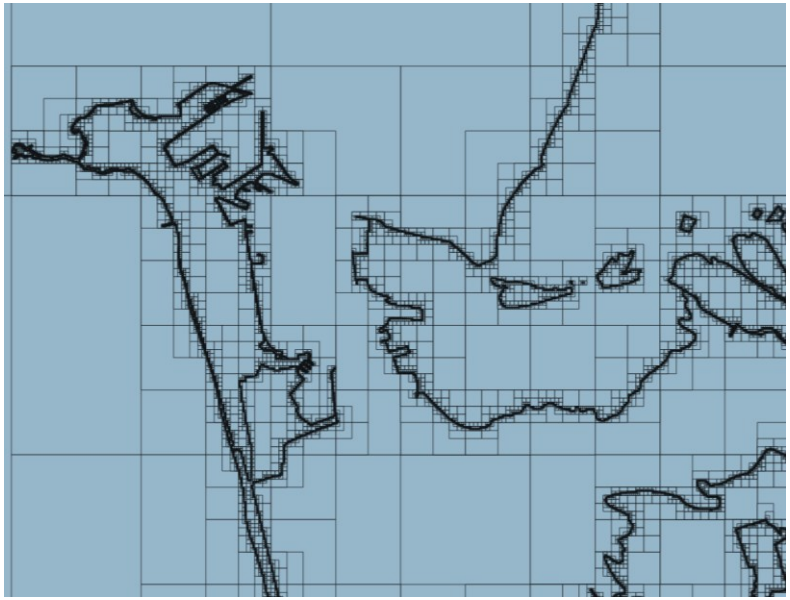
1.- Definición de la rejilla multiescalar a utilizar.

2.- Importación de datos normalizados de las fuentes seleccionadas para la elección de las variables a adscribir a las rejillas multiescalares.

3.- Reparto de los datos normalizados seleccionados en paquetes de trabajo para su procesado en paralelo.

4.- Elección de la variable temática y método de adscripción.

5.- Adscripción y almacenamiento.



El primer grupo lo constituirían las **fuentes que siguen el modelo raster, pero con diferente granularidad de rejilla, origen y/o nivel de** detalle.

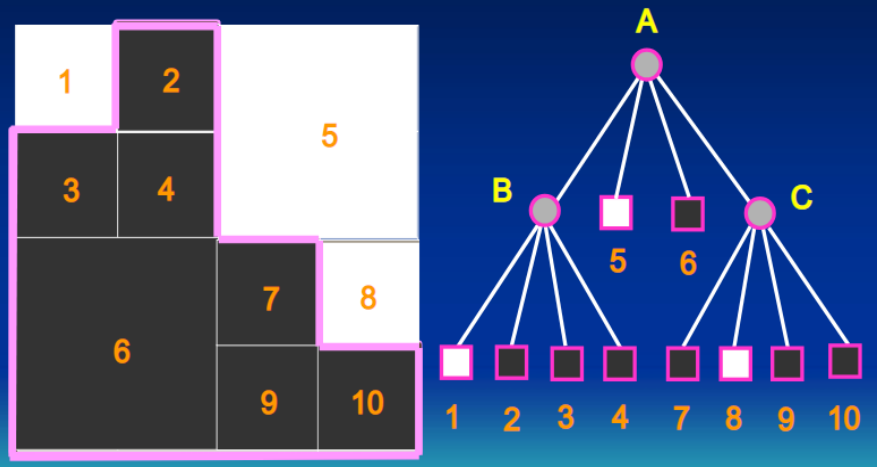
Las **fuentes que proporcionan información sobre variables discretas asociados a puntos** . El método de adscripción de estos datos no es compleja ya que se basaría en una relación espacial de inclusión en la rejilla de cada nivel escalar.

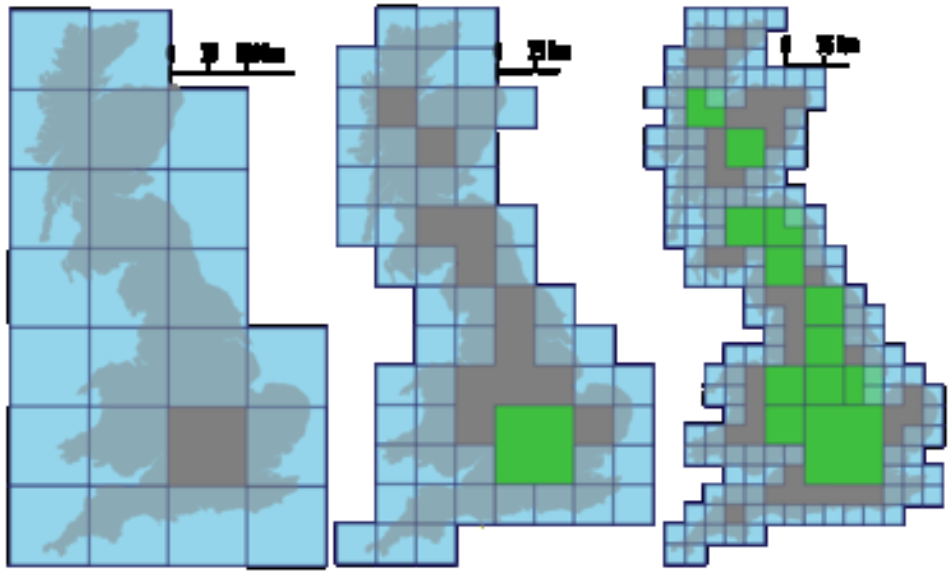
Las **fuentes que proporcionan información sobre variables contínuas asociados a puntos.** El método de adscripción de estos contemplaría una mayor complejidad si se quiere incorporar en el algoritmo de adscripción la información tras complejos procesos de interpolación.

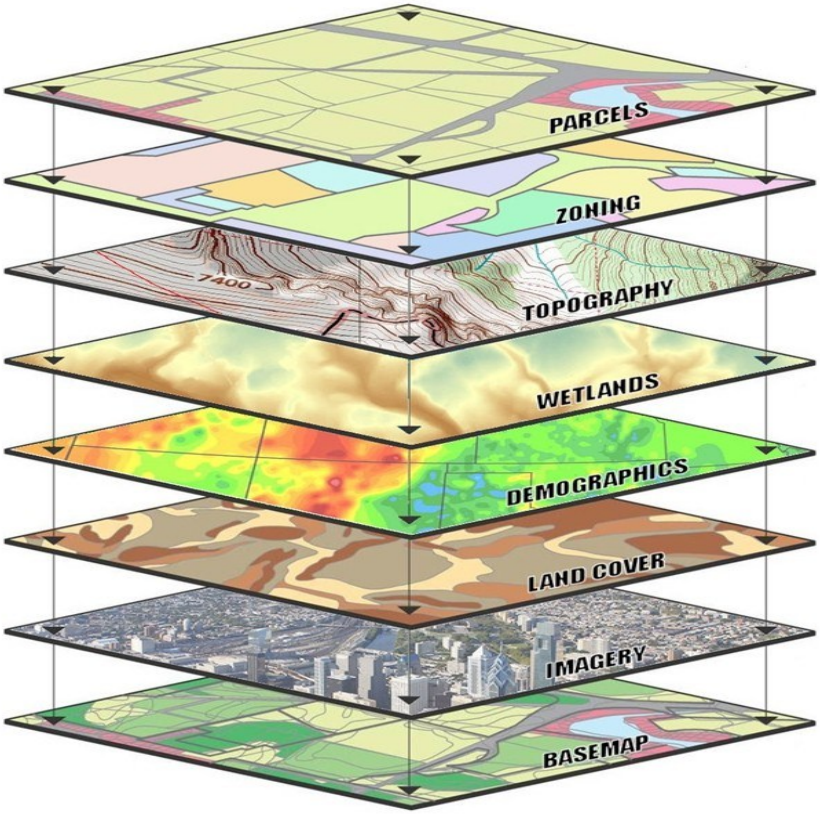
Las **fuentes que proporcionan información sobre variables asociadas a líneas.** Son representaciones unidimensionales de un fenómeno que se extiende sobre el territorio. En este caso, una opción podría ser la adscripción de las variables temáticas en función de la longitud de cada categorías de líneas que intescectan la rejilla.

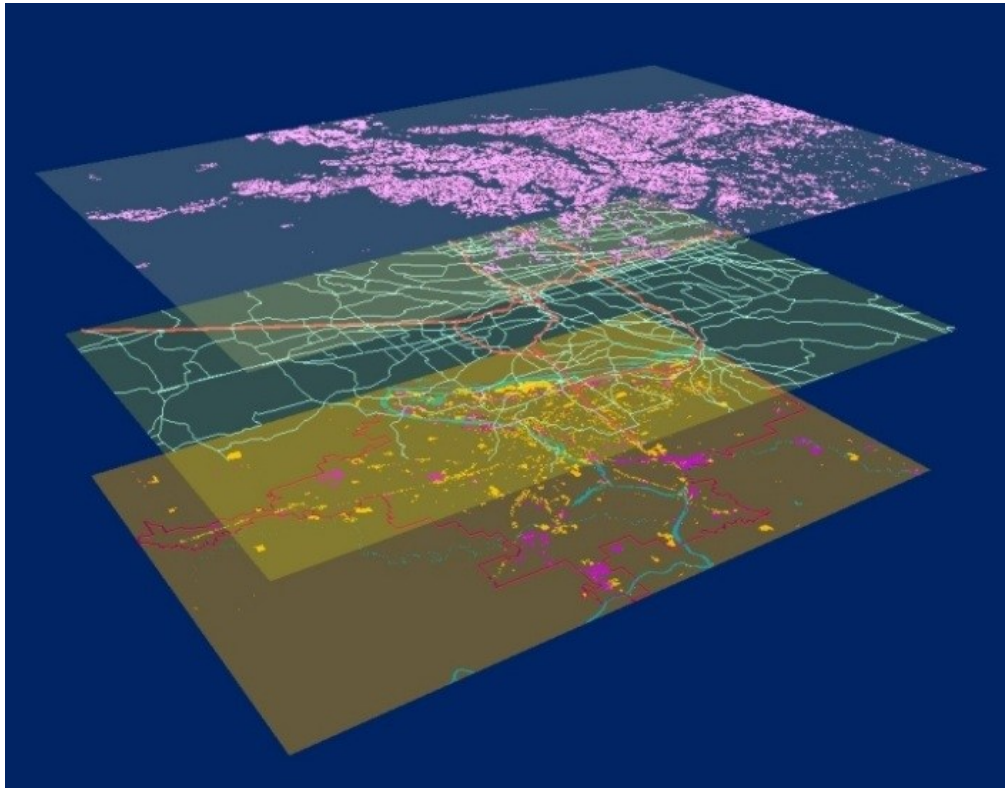
Las **fuentes que proporcionan información sobre variables asociadas a polígonos** Son representaciones bidimensionales de un fenómeno que se extiende sobre el territorio. En este caso, una opción podría ser la adscripción de las variables temáticas en función del área de cada categoría de polígonos que intescectna la rejilla.

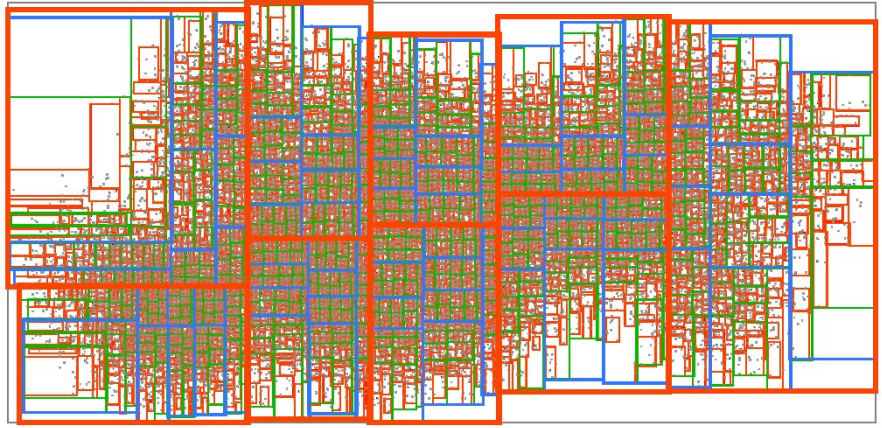
Quad-Tree Structure

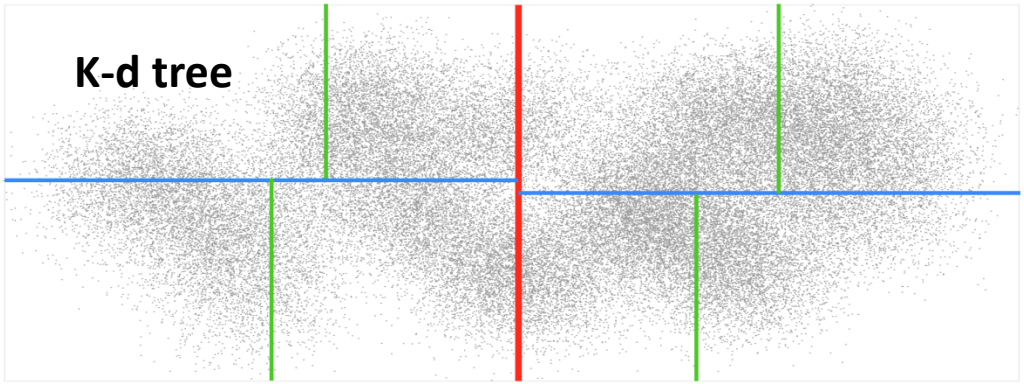


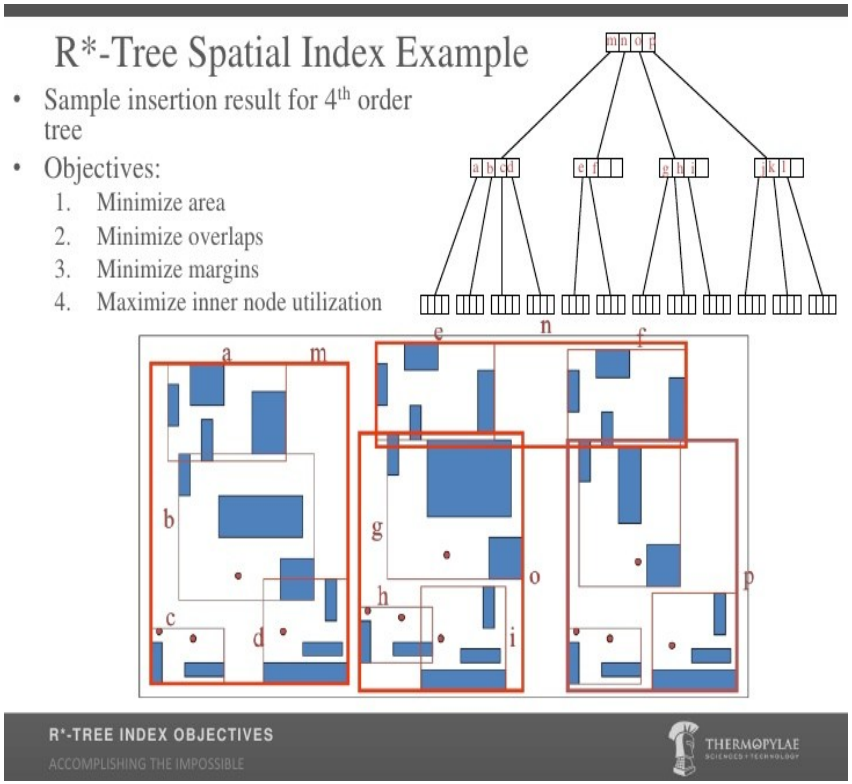


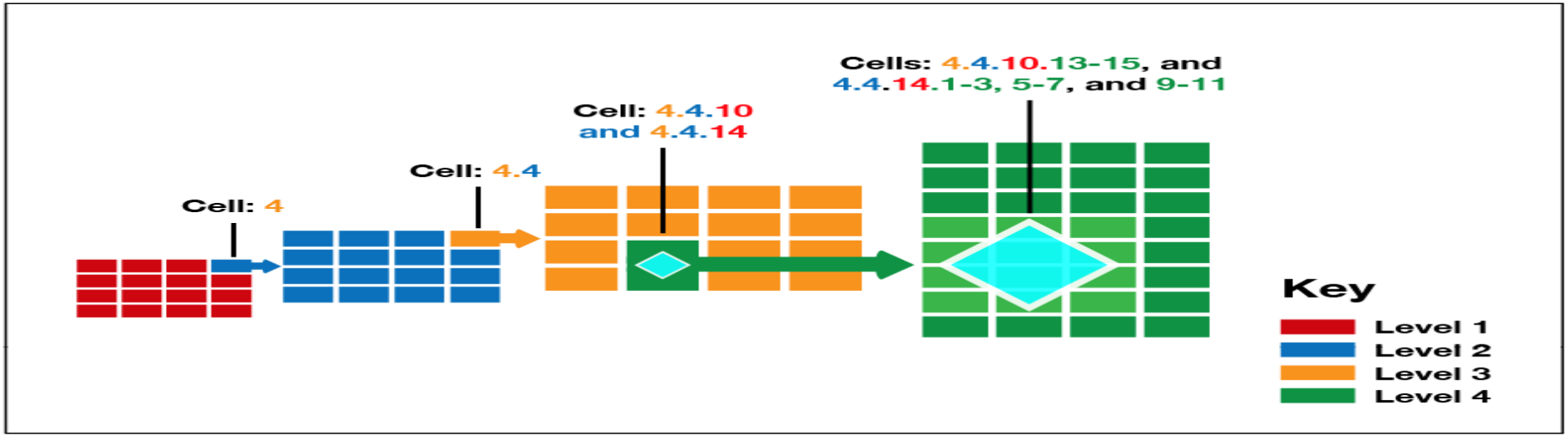






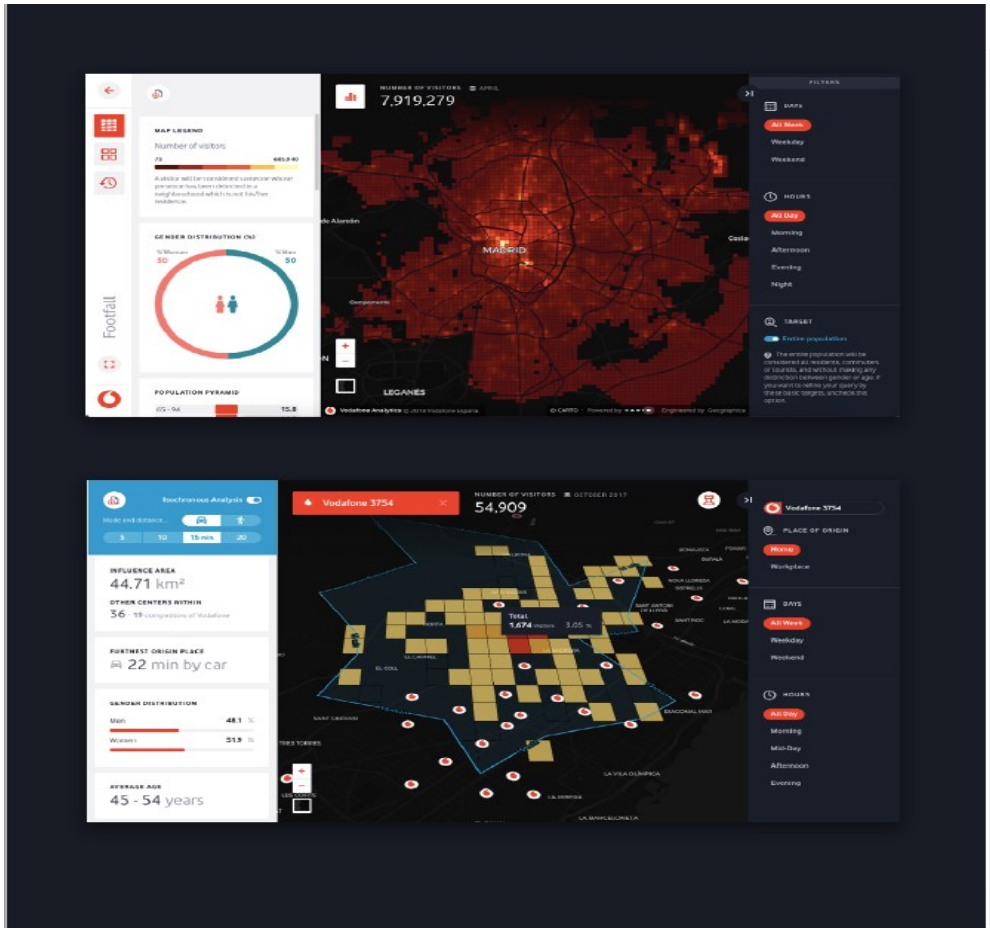

K-d tree

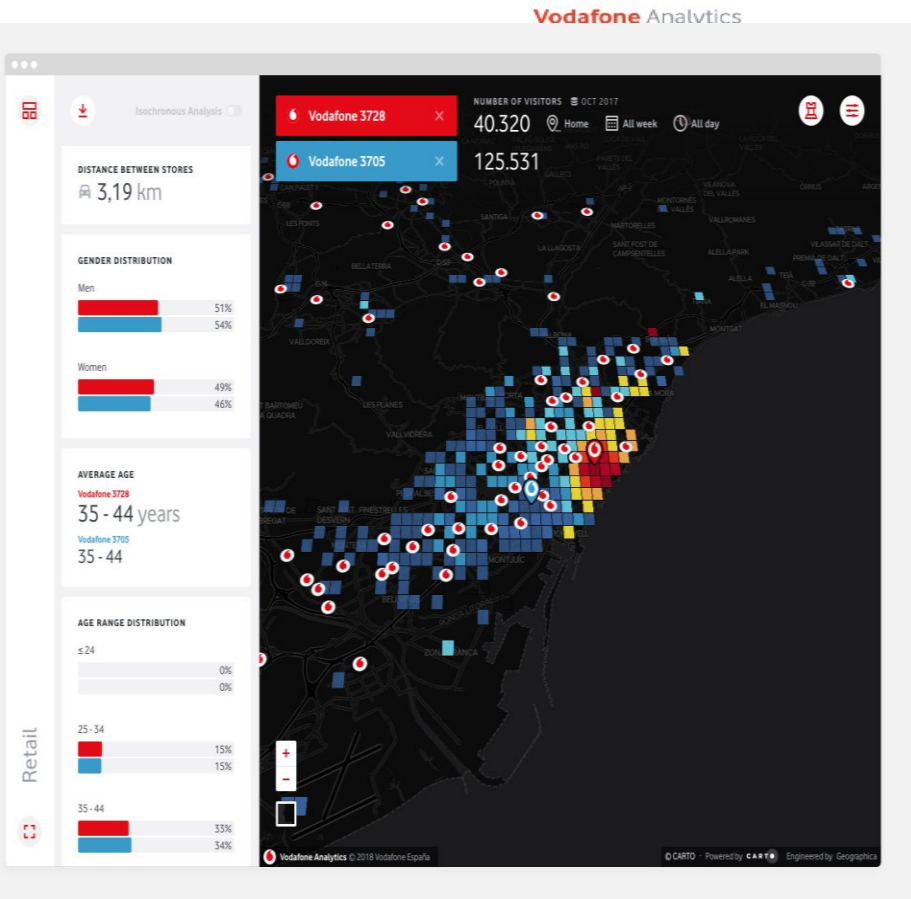

R*-Tree Spatial Index Example
• Sample insertion result for 4th order tree
• Objectives:
1. Minimize area
2. Minimize overlaps
3. Minimize margins
4. Maximize inner node utilization
a
m
e
n
f
b
g
o
p
h
c
d
i
R*-TREE INDEX OBJECTIVES
ACCOMPLISHING THE IMPOSSIBLE
THERMOPYLAE


Cell: 4
Cell: 4.4
Cell: 4.4.10
and 4.4.14
Cells: 4.4.10.13-15, and
4.4.14.1-3, 5-7, and 9-11
Key
Level 1
Level 2
Level 3
Level 4

VENTAJAS **rejillas multiescalares/multidimensionales «Spatial tessellations»**

Su estructura permite la **anonimización** y la **conservación del secreto estadístico para fuentes oficiales** y es también comúnmente utilizada en el procesado y análisis de **«big data»** por el mismo motivo (usuarios de móviles, por ejemplo)





SERVICIOS DE GENERACIÓN DE REJILLAS Y ADSCRIPCIÓN DE VARIABLES

1.- Definición de la rejilla multiescalar a utilizar.

2.- Importación de datos normalizados de las fuentes seleccionadas para la elección de las variables a adscribir a las rejillas multiescalares.

3.- Reparto de los datos normalizados seleccionados en paquetes de trabajo para su procesado en paralelo.

4.- Elección de la variable temática y método de adscripción.

5.- Adscripción y almacenamiento.

El microservicio/s debería proporcionar una serie de métodos de adscripción predefinidos para las fuentes de información seleccionadas, pero debería proporcionar a un **usuario avanzado (con los permisos de acceso y utilización correspondientes a nivel de API)** una suerte de **herramientas con las que desarrollar cualquier método que se precise,** no contemplado entre los incluidos de serie**. Estos nuevos métodos de adscripción (y los servicios en que se basen) pasarían a estar disponibles (catálogo de métodos y servicios de adscripción) para nuevos usuarios.**

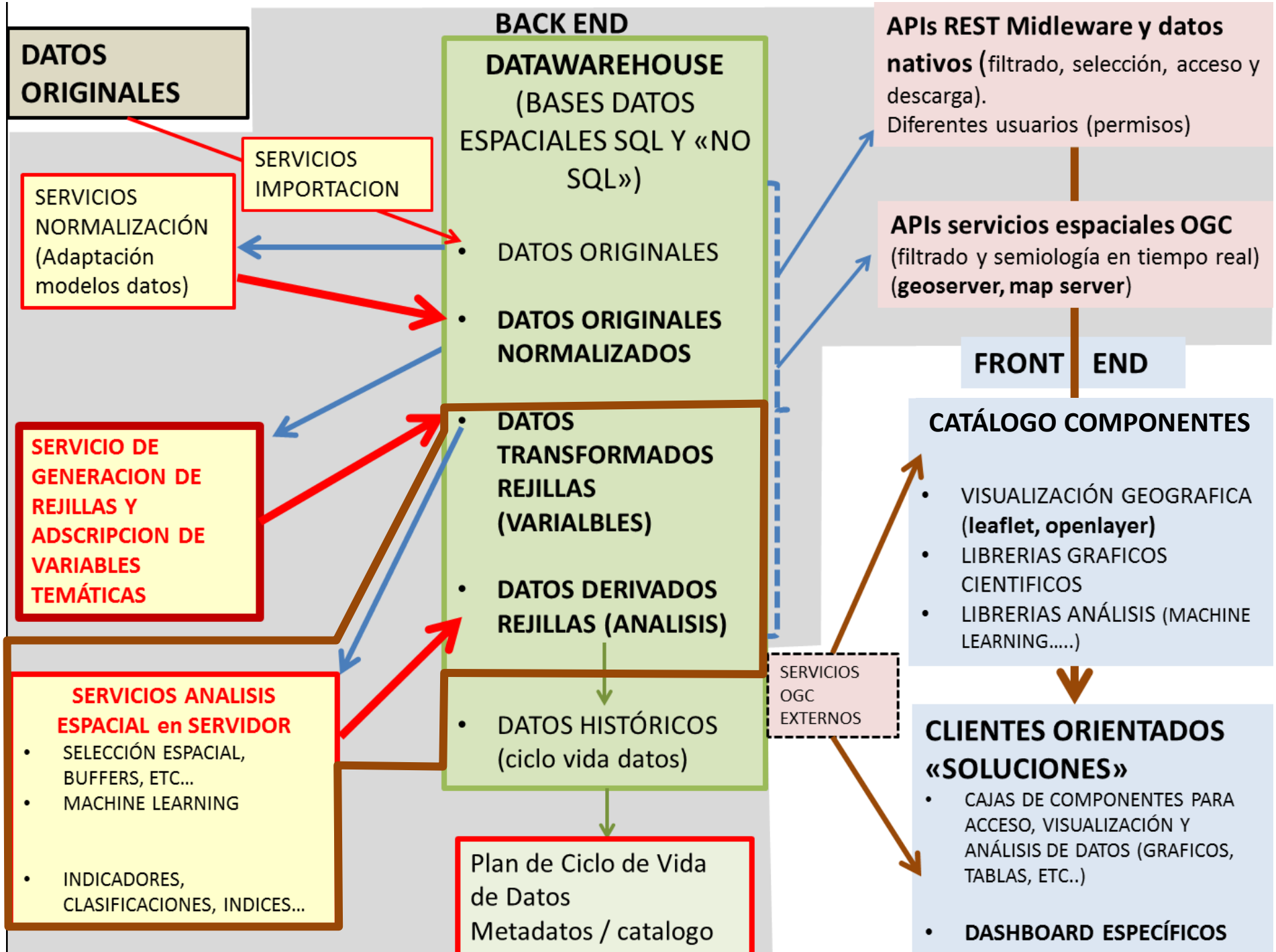
El primer grupo lo constituirían las **fuentes que siguen el modelo raster, pero con diferente granularidad de rejilla, origen y/o nivel de** detalle.

Las **fuentes que proporcionan información sobre variables discretas asociados a puntos** . El método de adscripción de estos datos no es compleja ya que se basaría en una relación espacial de inclusión en la rejilla de cada nivel escalar.

Las **fuentes que proporcionan información sobre variables contínuas asociados a puntos.** El método de adscripción de estos contemplaría una mayor complejidad si se quiere incorporar en el algoritmo de adscripción la información tras complejos procesos de interpolación.

Las **fuentes que proporcionan información sobre variables asociadas a líneas.** Son representaciones unidimensionales de un fenómeno que se extiende sobre el territorio. En este caso, una opción podría ser la adscripción de las variables temáticas en función de la longitud de cada categorías de líneas que intescectan la rejilla.

Las **fuentes que proporcionan información sobre variables asociadas a polígonos** Son representaciones bidimensionales de un fenómeno que se extiende sobre el territorio. En este caso, una opción podría ser la adscripción de las variables temáticas en función del área de cada categoría de polígonos que intescectna la rejilla.


BACK END
DATOS ORIGINALES
SERVICIOS IMPORTACION
SERVICIOS NORMALIZACIÓN (Adaptación modelos datos)
DATAWAREHOUSE (BASES DATOS ESPACIALES SQL Y «NO SQL»)
DATOS ORIGINALES
DATOS ORIGINALES NORMALIZADOS
DATOS TRANSFORMADOS REJILLAS (VARIALBLES)
DATOS DERIVADOS REJILLAS (ANALISIS)
DATOS HISTÓRICOS (ciclo vida datos)
SERVICIO DE GENERACION DE REJILLAS Y ADSCRIPCION DE VARIABLES TEMÁTICAS
SERVICIOS ANALISIS ESPACIAL en SERVIDOR
SELECCIÓN ESPACIAL, BUFFERS, ETC…
MACHINE LEARNING
INDICADORES, CLASIFICACIONES, INDICES…
Plan de Ciclo de Vida de Datos
Metadatos / catalogo
APIs REST Midleware y datos nativos (filtrado, selección, acceso y descarga).
Diferentes usuarios (permisos)
APIs servicios espaciales OGC (filtrado y semiología en tiempo real) (geoserver, map server)
FRONT END
CATÁLOGO COMPONENTES
VISUALIZACIÓN GEOGRAFICA (leaflet, openlayer)
LIBRERIAS GRAFICOS CIENTIFICOS
LIBRERIAS ANÁLISIS (MACHINE LEARNING…..)
SERVICIOS OGC EXTERNOS
CLIENTES ORIENTADOS «SOLUCIONES»
CAJAS DE COMPONENTES PARA ACCESO, VISUALIZACIÓN Y ANÁLISIS DE DATOS (GRAFICOS, TABLAS, ETC..)
DASHBOARD ESPECÍFICOS

VENTAJAS **rejillas multiescalares/multidimensionales «Spatial tessellations»**

- En relación los **procesos analíticos y construcción de indicadores** utilizando la información **multidimensional** dos **ventajas adicionales** podrían resaltarse:

  - Su estructura, al subdividir el espacio en **teselas de "cuasi equal-area"** facilitará su análisis con el **procesado distribuido (parallel computing**) ya que el área de estudio es fácilmente subdivisible y posteriormente los resultados del proceso analítico son fácilmente reensamblados espacialmente recuperando la topología implícita en la rejilla.

  - Por otra parte, el poseer cada tesela un **identificador único** en cada nivel escalar y tener asociados la totalidad de las variables ambientales y socioeconómicas que se definan, proporcionarán su información en una estructura parecida a un **vector n-dimensional** para cada tesela, que es una de las estructuras de datos que alimentan a los procesos analíticos de **Machine Learning (ML) e Inteligencia Artificial (AI)**

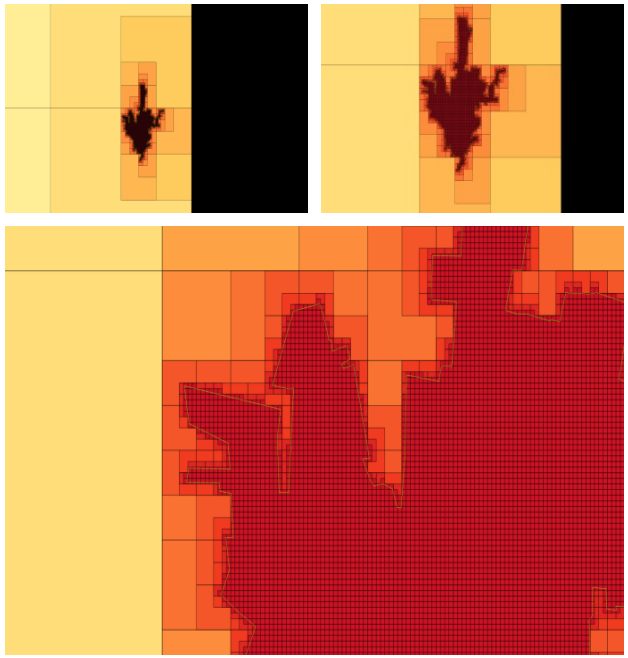
SERVICIOS DE ANÁLISIS ESPACIAL SOBRE DATOS EN EL SERVIDOR

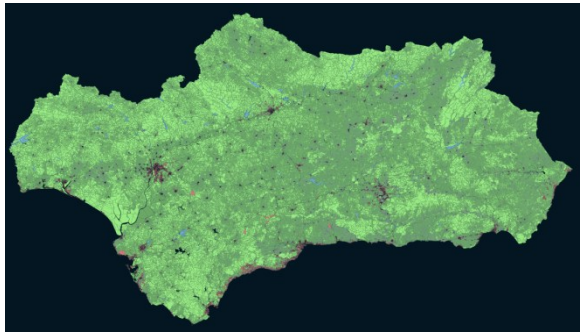
Para ello, los **servicios de análisis** podría hacer uso varias **aplicaciones y librerías** que proporcionarían el flujo de datos y servicios para realizar las tareas analíticas en la nube, que serán requisitos para el desarrollo de "algunas soluciones" . A modo de ejemplo podríamos citar las siguientes:

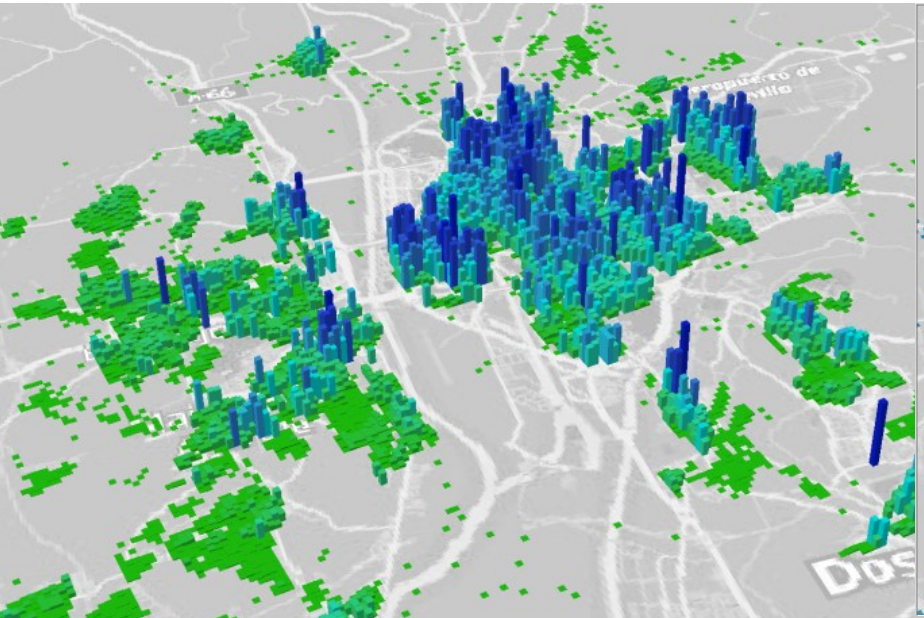
- Las librerias **GDAL (raster formats)** and **OGR (vector formats)** proporcionan funciones para la transformación de diferentes formatos de datos favoreciendo la interopertividad del sistem
- La librería **Proj4JS** proporciona funciones para la transformación de sistemas de referencia de coordenadas
- Las **bases de datos espaciales** (PostgreSQL/PostGis, por ejemplo) proporcionan un elevado conjunto de **funciones de análisis espacial** a través de sentencias SQL espaciales
- Para la **integración de flujos de trabajo** podría utilizarse **lenguajes de propósito general** (Python, etc.). El el caso de **Python,** este lenguaje es soportado por un amplio conjunto de librerías y aplicaciones orientadas a **funciones analíticas generales o específicas de análisis espacial**:
    - **Numpy, Pandas, Scipy, and Rasterio:** para operaciones científicas en el contexto de algebra de mapas para datos raster
    - **Fiona, Shapely:** para funciones analíticas en el ambito de los datos vectoriales.
    - **"R"**
    - **GRASS**
- **HADOOP o SPARK** proporcionan funcionalidades para ligadas al mundo **del Big data, Machine Learning o Inteligencia Artificial** que podrían ser muy interesantes (clasificaciones de datos, etc..) y para las que la estructura de datos de las rejillas multiescalares/multitemáticas ofrecen las comentadas ventajas para su ingesta nativa (vectores n-dimensionales).

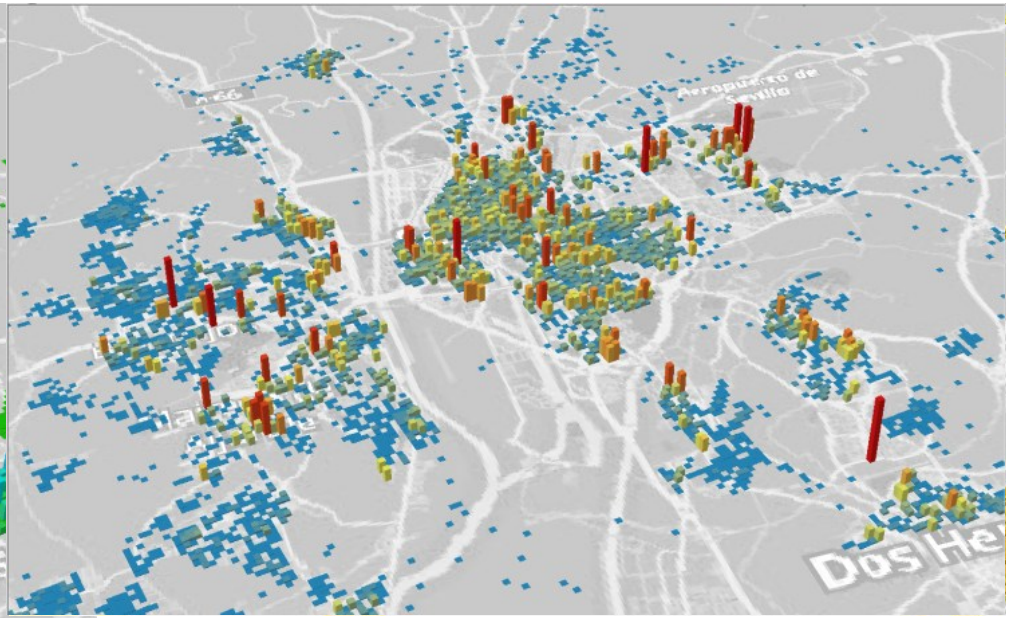
Dado que el uso de funciones analíticas en este contexto de **spatial big data** es muy exigente en tiempo de computación dicho servicio/s tendría que **poder ser desplegado fácilmente en una estrategia de computación en paralelo.**

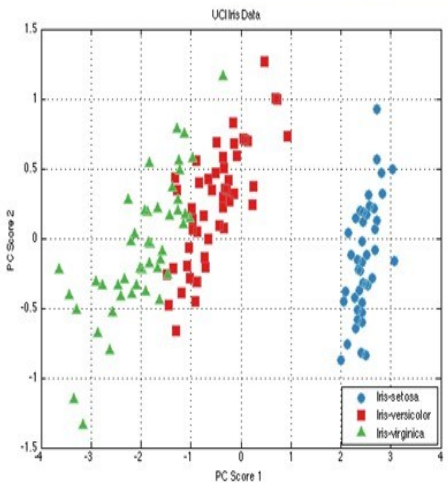
INTEGRACIÓN ESPACIAL DE VARIABLES SOCIO-AMBIENTALES EN ESTRUCTURAS DE REJILLAS MULTIESCALARES PARA SU DIFUSIÓN WEB











Técnicas de Machine Learning

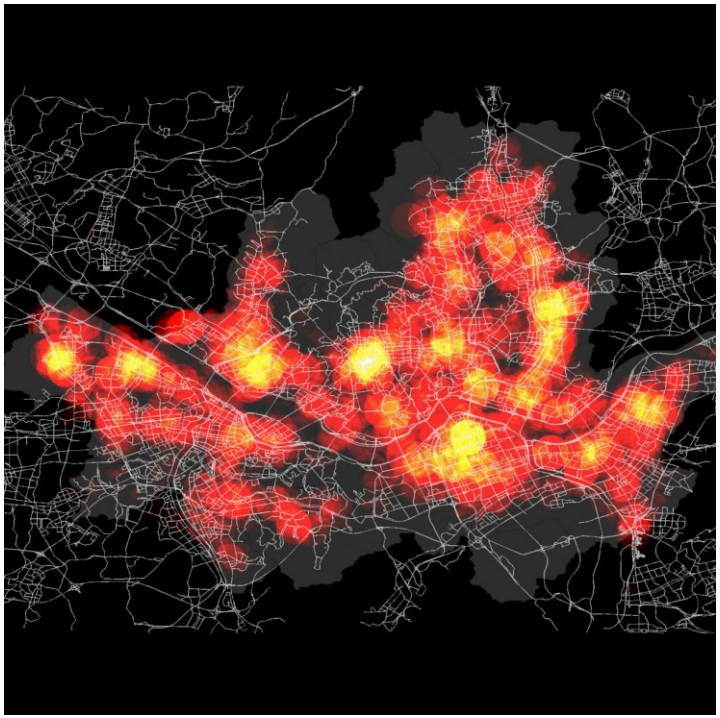
Clustering
Clasificación
Regresión

Básicamente los más comunes son:

- Clustering - agrupamiento
- Clasificación
- Regresión

Si no sabes a qué grupo pertenece trabajas a priori
→ **Clustering**

Si sabes el grupo a que pertenece lo que vas a entrenar a un modelo para que clasifique con nuevos datos
→ **Clasificación** (redes neuronales)



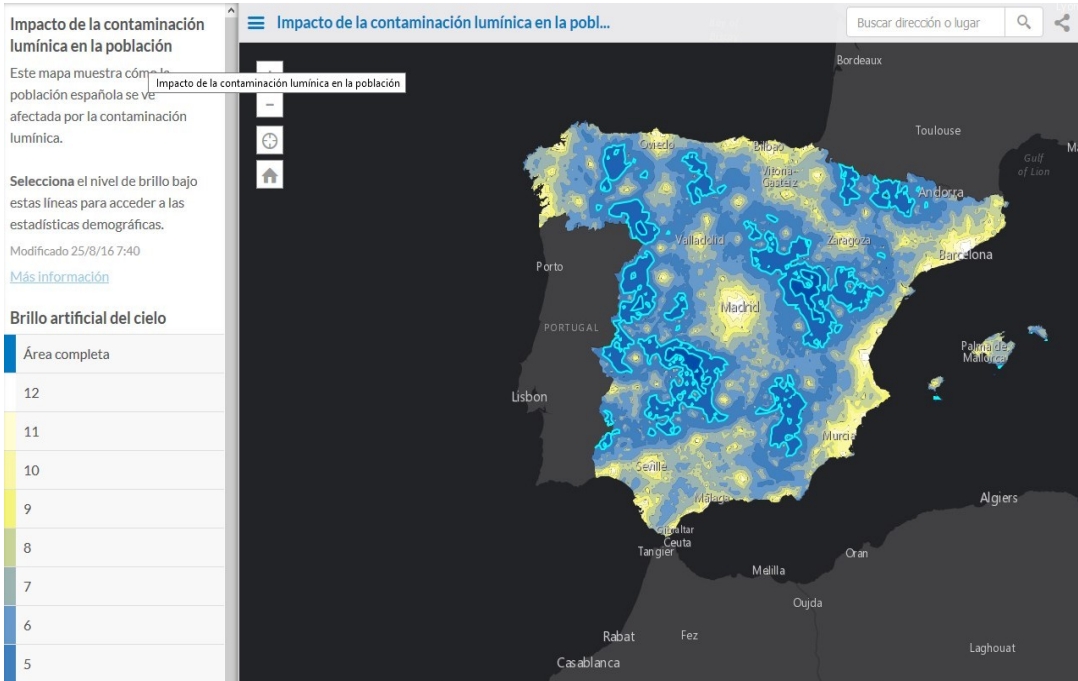
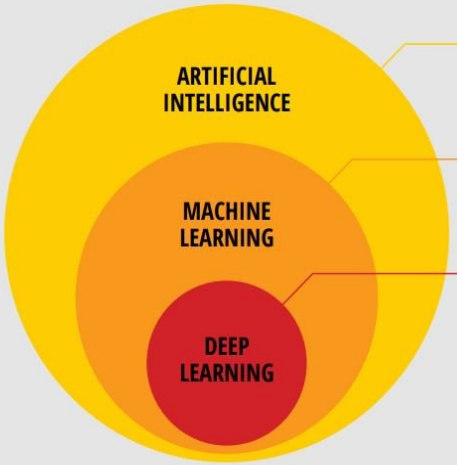


Figure 1 — Artificial intelligence (AI), machine learning and deep learning[7]



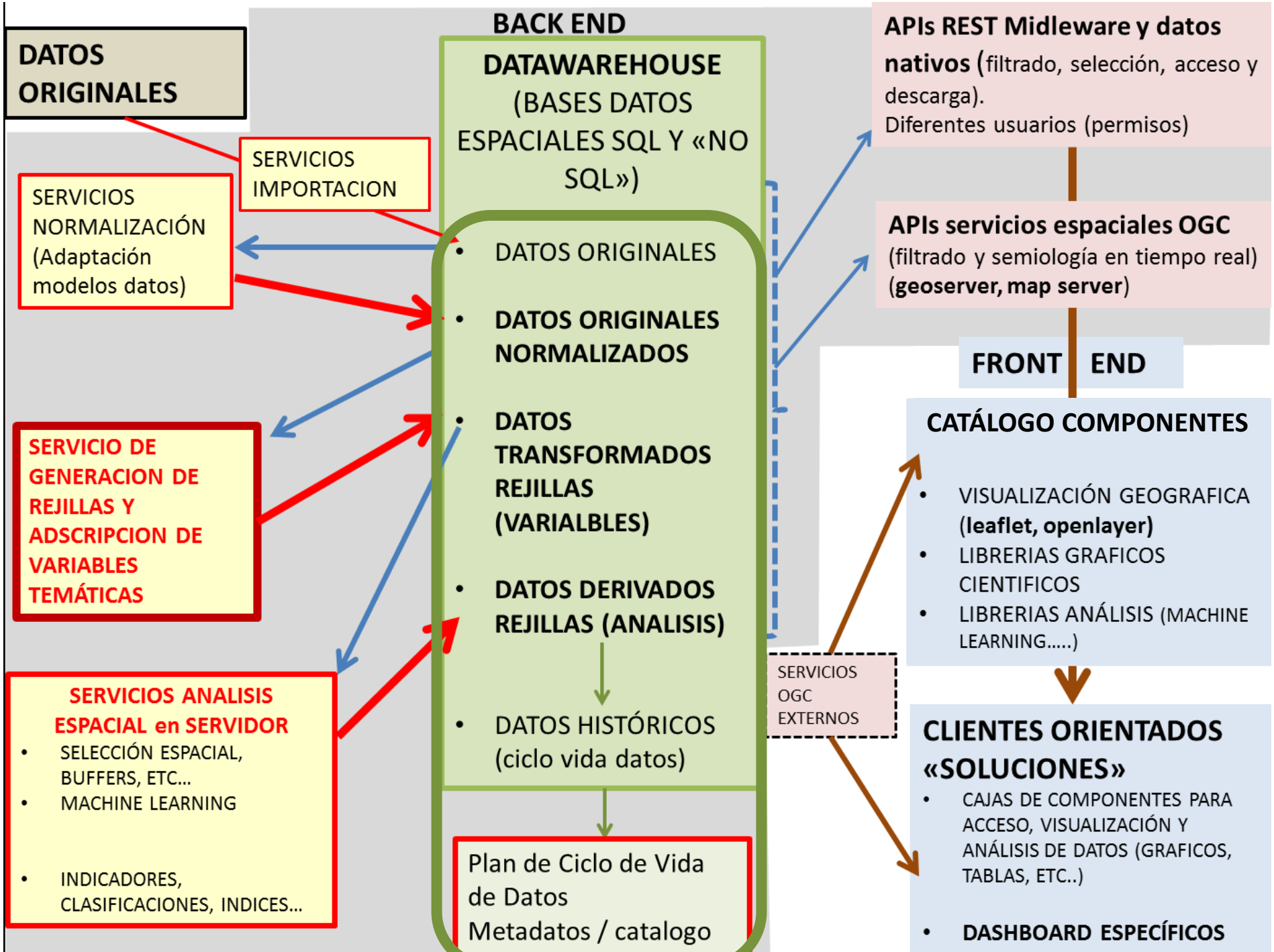
ARTIFICIAL INTELLIGENCE is the study of devices that perceive their environment and define a course of action that will maximize its chance of achieving a given goal.[8]

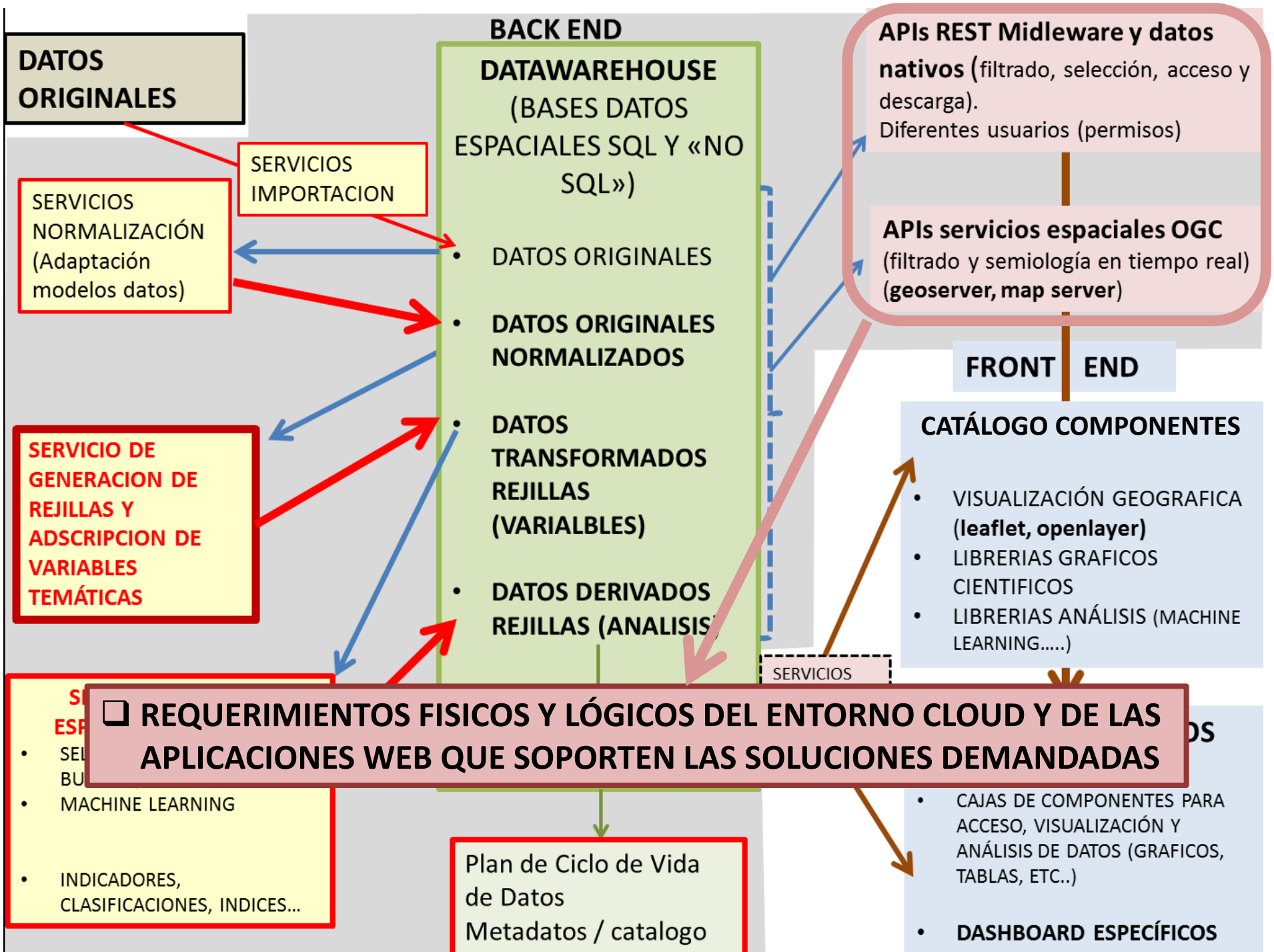
MACHINE LEARNING is a subset of artificial intelligence, in which machines learn how to to complete a certain task without being explicitly programmed to do so.

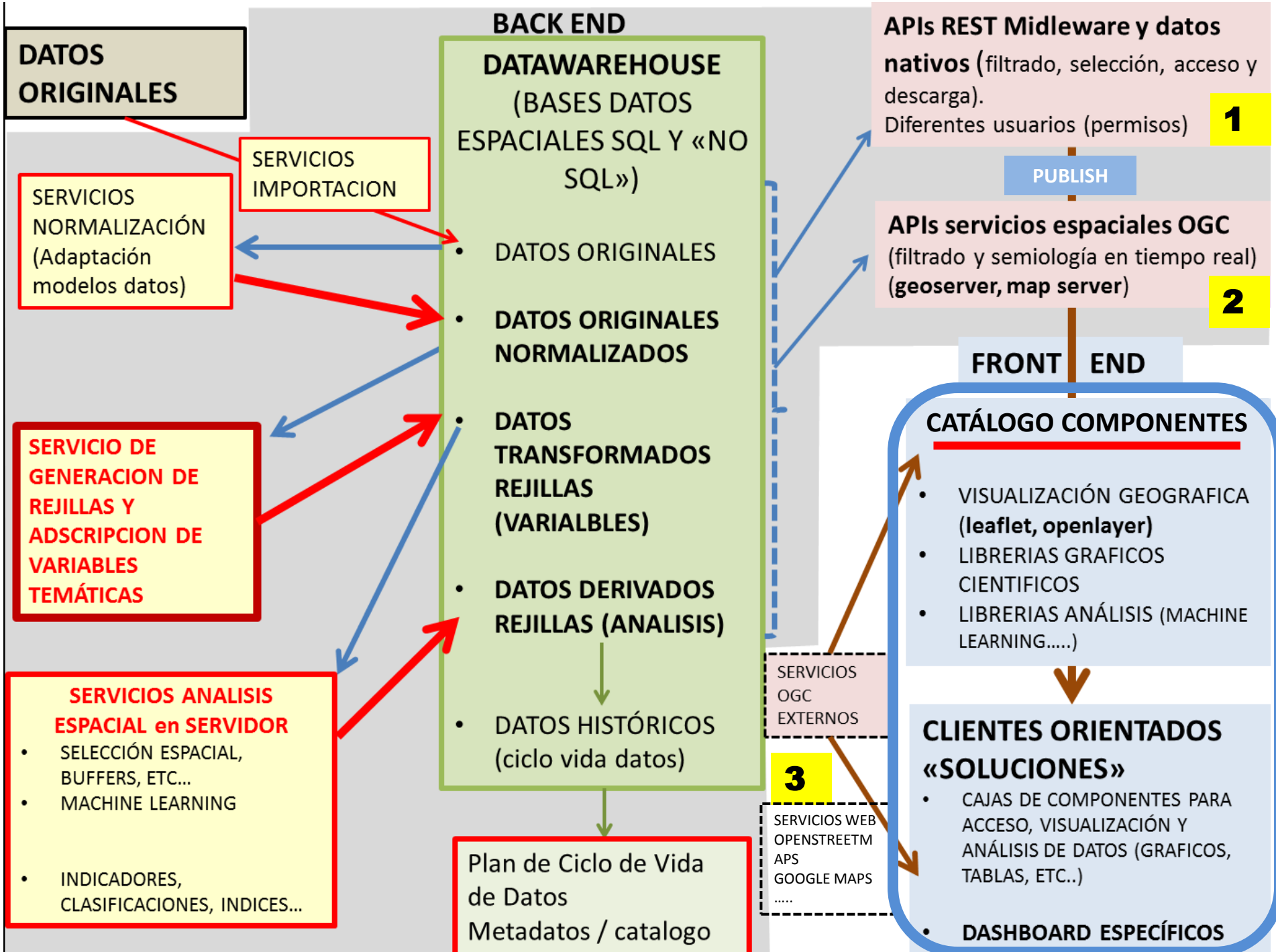
DEEP LEARNING is a subset of machine learning in which the tasks are broken down and distributed onto machine learning algorithms that are organised in consecutive layers. Each layer builds up on the output from the previous layer. Together the layers constitute an artificial neural network that mimics the distributed approach to problem-solving carried out by neurons in a human brain.

¿Tensorflow? ¿Caffe? ¿Torch?

TensorFlow + Keras


BACK END
DATOS ORIGINALES
SERVICIOS IMPORTACION
SERVICIOS NORMALIZACIÓN (Adaptación modelos datos)
DATAWAREHOUSE (BASES DATOS ESPACIALES SQL Y «NO SQL»)
DATOS ORIGINALES
DATOS ORIGINALES NORMALIZADOS
DATOS TRANSFORMADOS REJILLAS (VARIALBLES)
DATOS DERIVADOS REJILLAS (ANALISIS)
DATOS HISTÓRICOS (ciclo vida datos)
Plan de Ciclo de Vida de Datos Metadatos / catalogo
SERVICIO DE GENERACION DE REJILLAS Y ADSCRIPCION DE VARIABLES TEMÁTICAS
SERVICIOS ANALISIS ESPACIAL en SERVIDOR
SELECCIÓN ESPACIAL, BUFFERS, ETC…
MACHINE LEARNING
INDICADORES, CLASIFICACIONES, INDICES…
APIs REST Midleware y datos nativos (filtrado, selección, acceso y descarga).
Diferentes usuarios (permisos)
APIs servicios espaciales OGC (filtrado y semiología en tiempo real) (geoserver, map server)
FRONT END
CATÁLOGO COMPONENTES
VISUALIZACIÓN GEOGRAFICA (leaflet, openlayer)
LIBRERIAS GRAFICOS CIENTIFICOS
LIBRERIAS ANÁLISIS (MACHINE LEARNING…..)
SERVICIOS OGC EXTERNOS
CLIENTES ORIENTADOS «SOLUCIONES»
CAJAS DE COMPONENTES PARA ACCESO, VISUALIZACIÓN Y ANÁLISIS DE DATOS (GRAFICOS, TABLAS, ETC..)
DASHBOARD ESPECÍFICOS


DATOS ORIGINALES
BACK END
DATAWAREHOUSE (BASES DATOS ESPACIALES SQL Y «NO SQL»)
SERVICIOS IMPORTACION
SERVICIOS NORMALIZACIÓN (Adaptación modelos datos)
DATOS ORIGINALES
DATOS ORIGINALES NORMALIZADOS
DATOS TRANSFORMADOS REJILLAS (VARIALBLES)
DATOS DERIVADOS REJILLAS (ANALISIS)
SERVICIO DE GENERACION DE REJILLAS Y ADSCRIPCION DE VARIABLES TEMÁTICAS
APIs REST Midleware y datos nativos (filtrado, selección, acceso y descarga).
Diferentes usuarios (permisos)
APIs servicios espaciales OGC (filtrado y semiología en tiempo real) (geoserver, map server)
FRONT END
CATÁLOGO COMPONENTES
VISUALIZACIÓN GEOGRAFICA (leaflet, openlayer)
LIBRERIAS GRAFICOS CIENTIFICOS
LIBRERIAS ANÁLISIS (MACHINE LEARNING.....)
SERVICIOS
MACHINE LEARNING
INDICADORES, CLASIFICACIONES, INDICES...
Plan de Ciclo de Vida de Datos
Metadatos / catalogo
CAJAS DE COMPONENTES PARA ACCESO, VISUALIZACIÓN Y ANÁLISIS DE DATOS (GRAFICOS, TABLAS, ETC..)
DASHBOARD ESPECÍFICOS
❑ REQUERIMIENTOS FISICOS Y LÓGICOS DEL ENTORNO CLOUD Y DE LAS APLICACIONES WEB QUE SOPORTEN LAS SOLUCIONES DEMANDADAS


BACK END
DATOS ORIGINALES
SERVICIOS IMPORTACION
SERVICIOS NORMALIZACIÓN (Adaptación modelos datos)
DATAWAREHOUSE (BASES DATOS ESPACIALES SQL Y «NO SQL»)
DATOS ORIGINALES
DATOS ORIGINALES NORMALIZADOS
DATOS TRANSFORMADOS REJILLAS (VARIALBLES)
DATOS DERIVADOS REJILLAS (ANALISIS)
DATOS HISTÓRICOS (ciclo vida datos)
SERVICIO DE GENERACION DE REJILLAS Y ADSCRIPCION DE VARIABLES TEMÁTICAS
SERVICIOS ANALISIS ESPACIAL en SERVIDOR
SELECCIÓN ESPACIAL, BUFFERS, ETC…
MACHINE LEARNING
INDICADORES, CLASIFICACIONES, INDICES…
Plan de Ciclo de Vida de Datos
Metadatos / catalogo
APIs REST Midleware y datos nativos (filtrado, selección, acceso y descarga).
Diferentes usuarios (permisos)
1
PUBLISH
APIs servicios espaciales OGC (filtrado y semiología en tiempo real) (geoserver, map server)
2
FRONT END
CATÁLOGO COMPONENTES
VISUALIZACIÓN GEOGRAFICA (leaflet, openlayer)
LIBRERIAS GRAFICOS CIENTIFICOS
LIBRERIAS ANÁLISIS (MACHINE LEARNING…..)
CLIENTES ORIENTADOS «SOLUCIONES»
CAJAS DE COMPONENTES PARA ACCESO, VISUALIZACIÓN Y ANÁLISIS DE DATOS (GRAFICOS, TABLAS, ETC..)
DASHBOARD ESPECÍFICOS
SERVICIOS OGC EXTERNOS
3
SERVICIOS WEB OPENSTREETMAPS GOOGLE MAPS …..

CATÁLOGO DE COMPONENTES DE CLIENTES WEB

Los componentes de un cliente básico y generalista se podría desarrollar sobre las tecnologías **HTML5** y **JavaScript** y tendría los siguientes **componentes y funcionalidades básicos** :

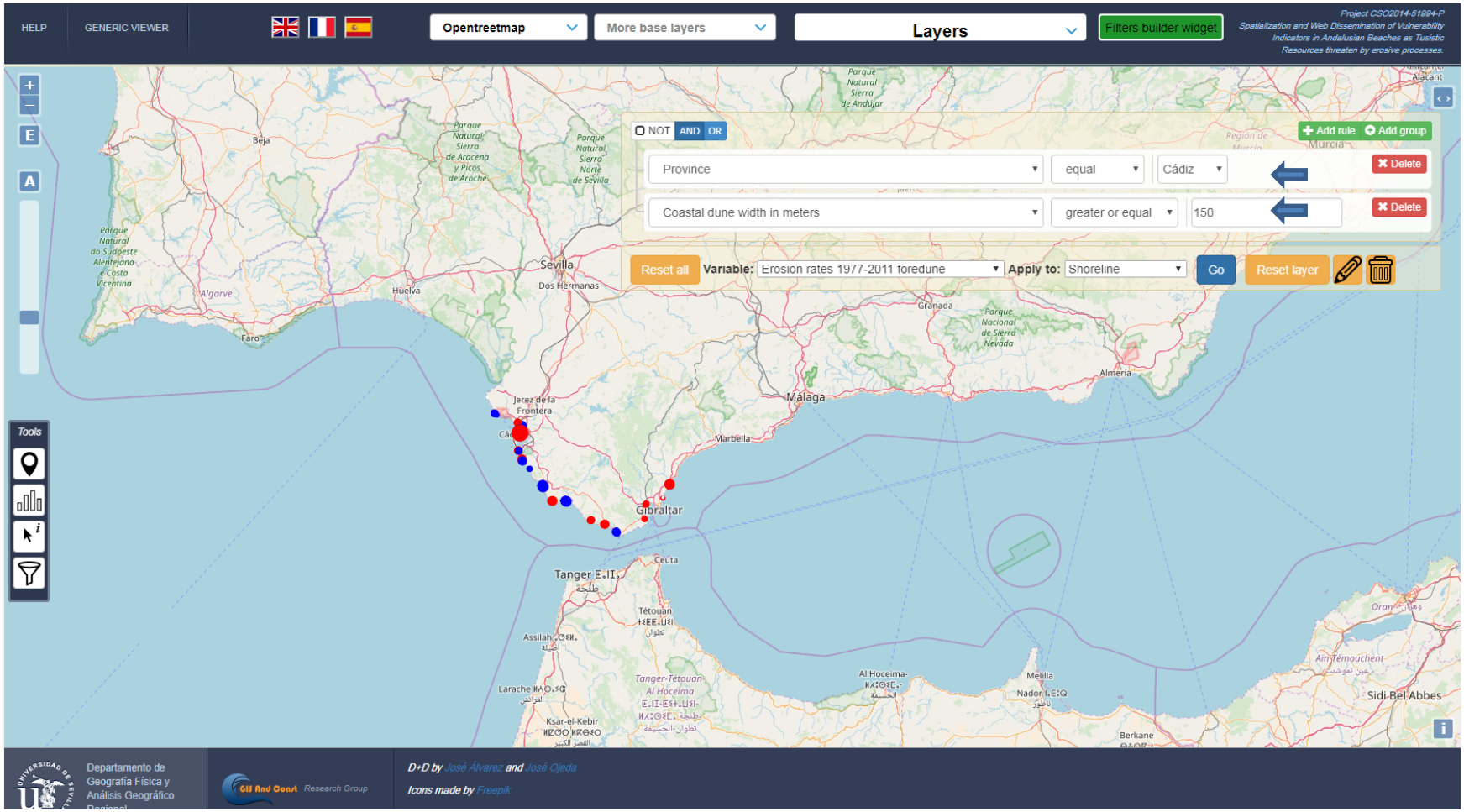
Componentes **para la visualización cartográfica: Geovisor web**. Para el desarrollo de este geovisor web se podrían utilizar librerías como **Leaflet** u **Openlayer** que ofrecen **las funcionalidades genéricas** de un visor. Este cliente debería visualizar:

- Los **servicios de mapas o imágenes de grandes servidores cartográficos ((OpenStretMap, Google..),** por lo se impondrá el sistema de referencia de coordenadas "Google Mercator" (EPSG:3857) para la visualización cartográfica.
- Tanto a los **servicios OGC externos como los generados en el sistema** a **partir de los datos adscritos a las rejillas multiescalares**. En el caso que los servicios sean publicados en otros SRC diferentes a EPSG:3857, podría utilizar la librería **Proj4JS** para la transformación geométrica de coordenadas. Debería igualmente ser capaz de acceder a **Servicios OGC filtrados** (CQL filter)

CATÁLOGO DE COMPONENTES DE CLIENTES WEB

Los componentes de un cliente básico y generalista se desarrollaría sobre las tecnologías **HTML5** y **JavaScript** y tendría los siguientes **componentes y funcionalidades**:

Filtrado ECQL (visualización a demanda filtrando la componente espacial y temática)

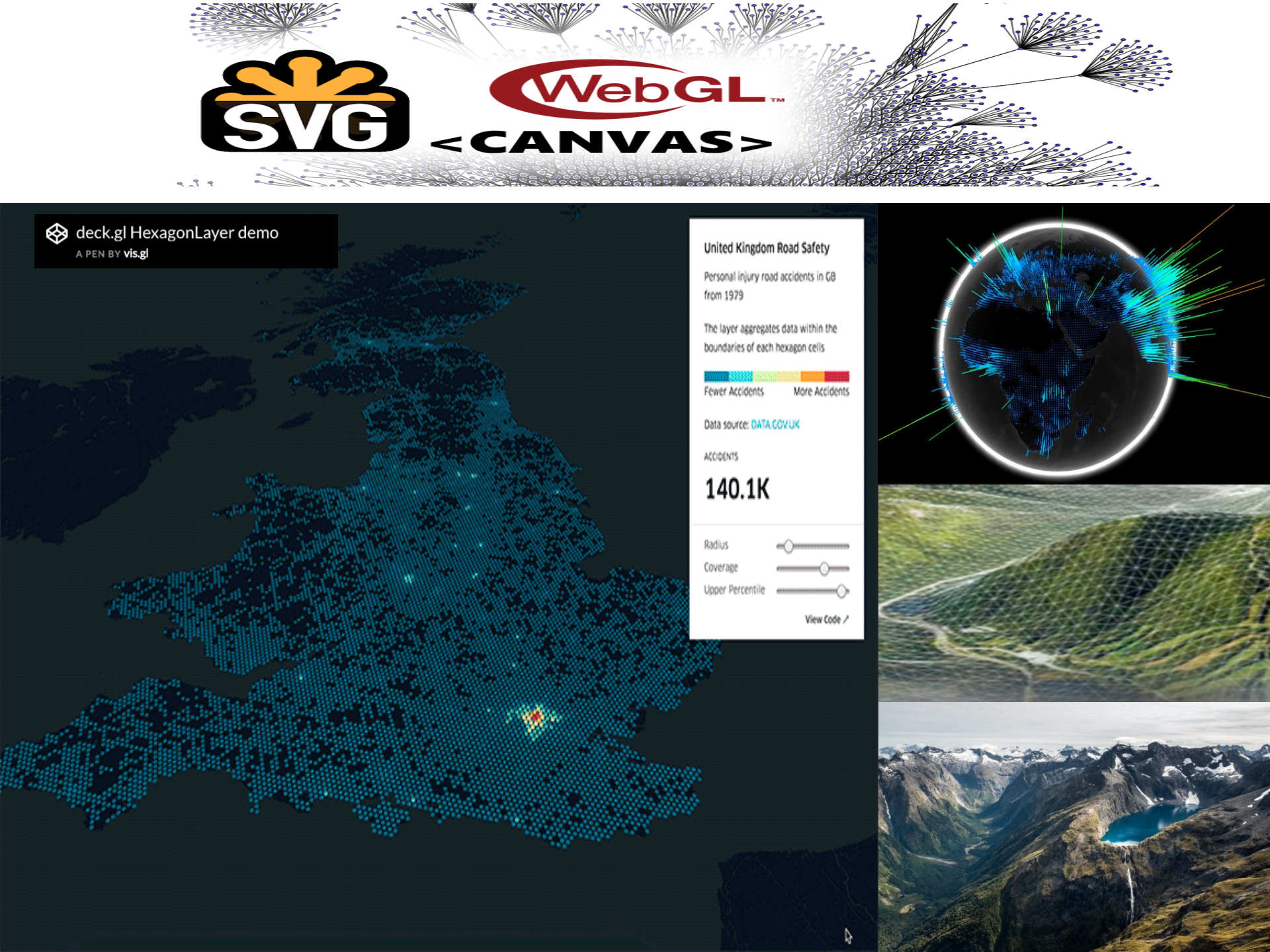


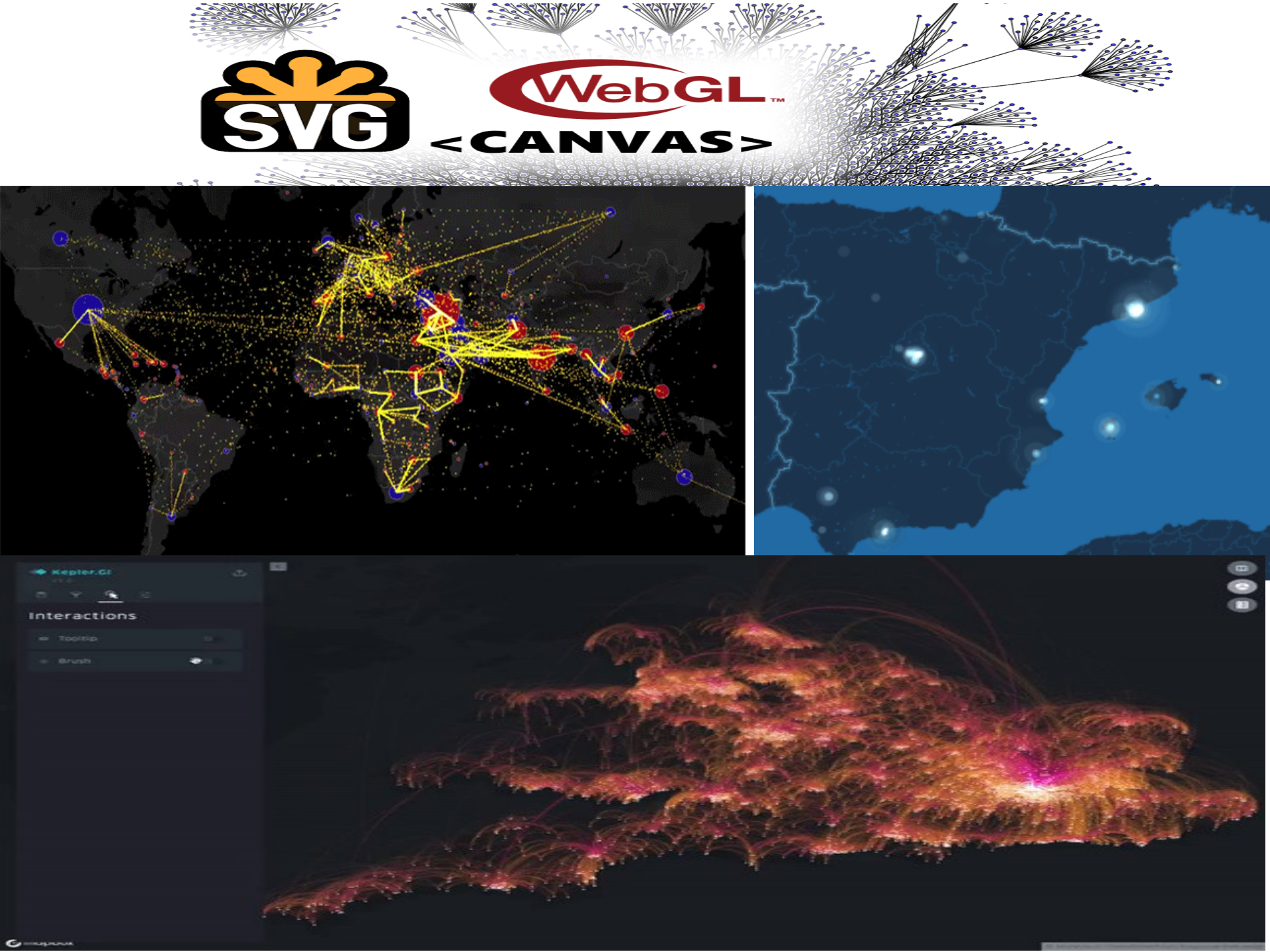
CATÁLOGO DE COMPONENTES DE CLIENTES WEB

Los componentes de un cliente básico y generalista se desarrollaría sobre las tecnologías **HTML5** y **JavaScript** y tendría los siguientes **componentes y funcionalidades básicos** :

Componentes **para la visualización cartografica: Geovisor web**. Para el desarrollo de este geovisor web se podrían utilizar librerías como **Leaflet** u **Openlayer** que ofrecen **las funcionalidades genéricas** de un visor. Este cliente debería visualizar:

- Los servicios de mapas o imágenes de grandes servidores cartográficos **((OpenStretMap, Google..),** por lo se impondrá el sistema de referencia de coordenadas "Google Mercator" (EPSG:3857) para la visualización cartográfica.
- Tanto a los **servicios OGC externos como los generados en el sistema** a **partir de los datos adscritos a las rejillas multiescalares**. En el caso que los servicios sean publicados en otros SRC diferentes a EPSG:3857, podría utilizar la librería **Proj4JS** para la transformación geométrica de coordenadas. Deberia igualmente ser capaz de acceder a **Servicios OGC filtrados** (CQL filter)
- Deberia tener acceso a los **servicios de los datos nativos de un número determinado de las variables (a elegir por el usuarios) de la rejilla multiescalar**, aunque **solo al nivel de escala que exija su visualización** (acceso a través de una API especifica). Sera necesario desarrollar el **software específico** en el cliente para que reciba los datos a través de esta API y los despliegue sobre Leaflet u Openlayer. Este componente de software conlleva una **importante carga innovadora** en relación a los clientes ligeros de visualización cartográfica clásicos donde la renderización o el tratamiento semiológico se realiza en a la parte servidora, ya que, tras recibir los datos de las variables seleccionadas al nivel escalar dependiente de la escala de visualización, **deberá proporcionar diferentes funcionalidades tanto para el tratamiento semiológico de los datos como para su filtrado temático o espacial en tiempo real en el cliente en tiempo real.** Para su desarrollo se podrían emplear librerías de visualización avanzada de datos (**D3.js**) que generan información visual en formato SVG o bien desarrollarse **WebGL**


SVG
WebGL
<CANVAS>
deck.gl HexagonLayer demo
A PEN BY vis.gl
United Kingdom Road Safety
Personal injury road accidents in GB from 1979
The layer aggregates data within the boundaries of each hexagon cells
Fewer Accidents
More Accidents
Data source: DATA.GOV.UK
ACCIDENTS
140.1K
Radius
Coverage
Upper Percentile
View Code


SVG
WebGL
<CANVAS>
Kepler.GL
Interactions
Tooltip
Brush

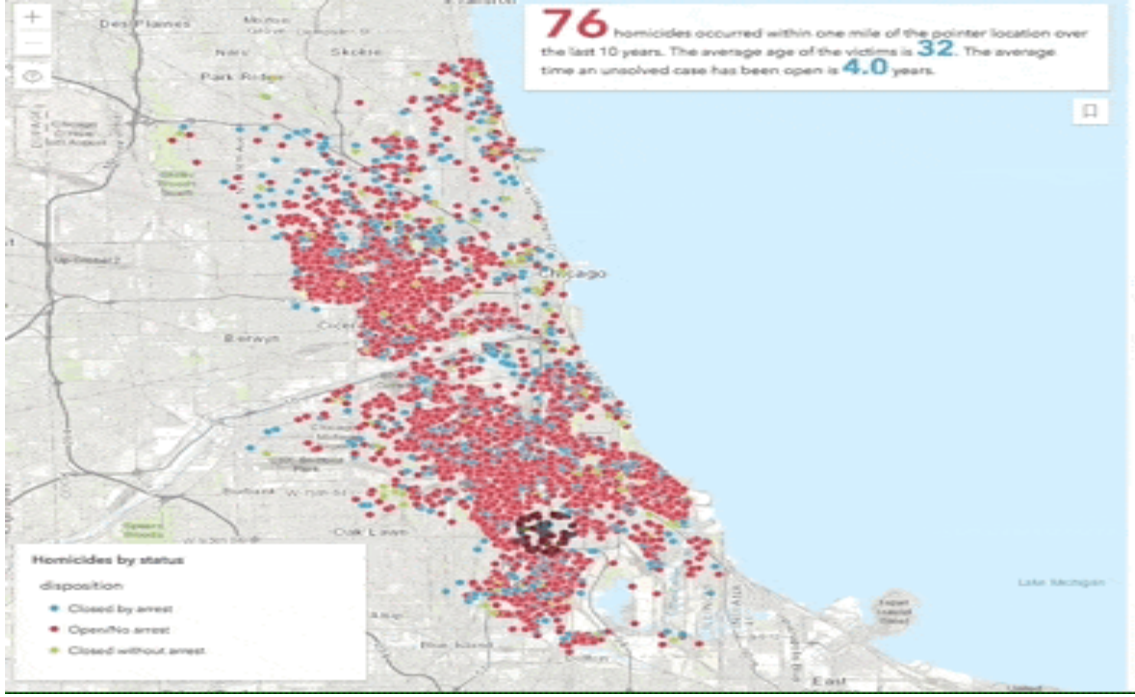
CATÁLOGO DE COMPONENTES DE CLIENTES WEB

Los componentes de un cliente básico y generalista se desarrollaría sobre las tecnologías **HTML5** y **JavaScript** y tendría los siguientes **componentes y funcionalidades básicos**:

Componentes **para la visualización cartografica: Geovisor web**. Para el desarrollo de este geovisor web se podrían utilizar librerías como **Leaflet** u **Openlayer** que ofrecen **las funcionalidades genéricas** de un visor. Este cliente debería visualizar:

- Los servicios de mapas o imágenes de grandes servidores cartográficos **((OpenStretMap, Google..),** por lo se impondrá el sistema de referencia de coordenadas "Google Mercator" (EPSG:3857) para la visualización cartográfica.
- Tanto a los **servicios OGC externos como los generados en el sistema** a **partir de los datos adscritos a las rejillas multiescalares**. En el caso que los servicios sean publicados en otros SRC diferentes a EPSG:3857, podría utilizar la librería **Proj4JS** para la transformación geométrica de coordenadas. Deberia igualmente ser capaz de acceder a **Servicios OGC filtrados** (CQL filter)
- Deberia tener acceso a los **servicios de los datos nativos de un número determinado de las variables (a elegir por el usuarios) de la rejilla multiescalar**, aunque **solo al nivel de escala que exija su visualización** (acceso a través de una API especifica -ver epígrafe 4.5-). Sera necesario desarrollar el **software específico** en el cliente para que reciba los datos a través de esta API y los despliegue sobre Leaflet u Openlayer. Este componente de software conlleva una **importante carga innovadora** en relación a los clientes ligeros de visualización cartográfica clásicos donde la renderización o el tratamiento semiológico se realiza en a la parte servidora, ya que, tras recibir los datos de las variables seleccionadas al nivel escalar dependiente de la escala de visualización, **deberá proporcionar diferentes funcionalidades tanto para el tratamiento semiológico de los datos como para su filtrado temático o espacial en tiempo real en el cliente en tiempo real.** Para su desarrollo se podrían emplear librerías de visualización avanzada de datos (**D3.js**) que generan información visual en formato SVG o bien desarrollarse **WebGL**

Componentes para **visualización gráfica de datos**. Dado que serán los datos nativos de las variables (no su representación cartográfica) los que llegarán al cliente, este deberá ofrecer **funcionalidades para la representación gráfica** de los datos (graficos de barras, donut, etc..) que llegan al cliente (o que se filtran en el mismo tras su llegada en el cliente) en tiempo real**.** Existen **diferentes librerías** que se podrían utilizar **(D3.js, Highcharts**, etc…)


76 homicides occurred within one mile of the pointer location over the last 10 years. The average age of the victims is 32. The average time an unsolved case has been open is 4.0 years.
Homicides by status
disposition
Closed by arrest
Open/No arrest
Closed without arrest
Des Plaines
Skokie
Park Ridge
Chicago
Berwyn
Oak Lawn
Lake Michigan

Homicides by year
Solved by 2017
Remains unsolved
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
Homicides by age
65+
45-64
31-44
18-30
0-18
Not sure
Status of the case
Gender of the victim
Race of the victim
Closed by arrest
Closed without arrest
Open/No arrest
Male (solved)
Male (unsolved)
Female (solved)
Female (unsolved)
Asian (solved)
Asian (unsolved)
Black (solved)
Black (unsolved)
Hispanic (solved)
Hispanic (unsolved)
White (solved)
White (unsolved)

Kepler.Gl
Layers
Add Data
Point
Basic
Latitude
Longitude
Color
Magnitude
Radius
NEVADA
UTAH
ARIZONA
LAS VEGAS
SACRAMENTO
FRESNO
SANTA BARBARA
LOS ANGELES
SAN DIEGO
PHOENIX
FLAGSTAFF
SALT LAKE CITY

CATÁLOGO DE COMPONENTES DE CLIENTES WEB

Los componentes de un cliente básico y generalista se desarrollaría sobre las tecnologías **HTML5** y **JavaScript** y tendría los siguientes **componentes y funcionalidades básicos**:

Componentes **para la visualización cartografica: Geovisor web**. Para el desarrollo de este geovisor web se podrían utilizar librerías como **Leaflet** u **Openlayer** que ofrecen **las funcionalidades genéricas** de un visor. Este cliente debería visualizar:

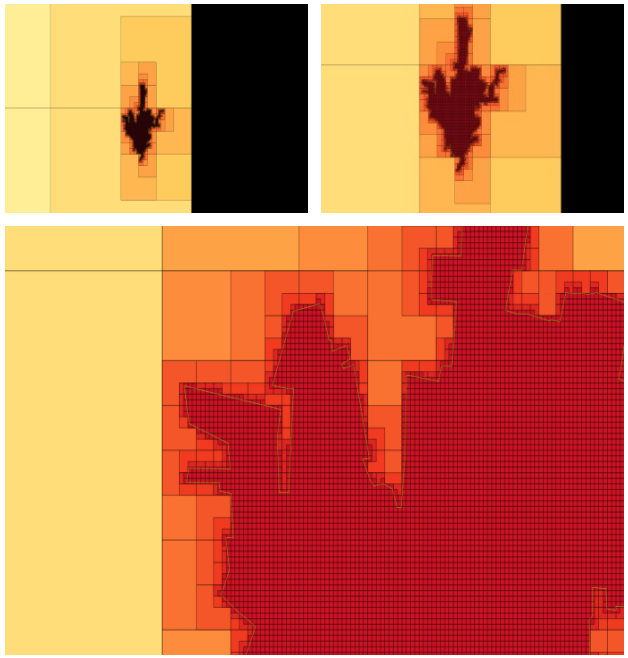
- Los servicios de mapas o imágenes de grandes servidores cartográficos **((OpenStretMap, Google..),** por lo se impondrá el sistema de referencia de coordenadas "Google Mercator" (EPSG:3857) para la visualización cartográfica.
- Tanto a los **servicios OGC externos como los generados en el sistema** a **partir de los datos adscritos a las rejillas multiescalares**. En el caso que los servicios sean publicados en otros SRC diferentes a EPSG:3857, podría utilizar la librería **Proj4JS** para la transformación geométrica de coordenadas. Deberia igualmente ser capaz de acceder a **Servicios OGC filtrados** (CQL filter)
- Deberia tener acceso a los **servicios de los datos nativos de un número determinado de las variables (a elegir por el usuarios) de la rejilla multiescalar**, aunque solo al nivel de escala que exija su visualización (acceso a través de una API especifica -ver epígrafe 4.5-). Sera necesario desarrollar el **software específico** en el cliente para que reciba los datos a través de esta API y los despliegue sobre Leaflet u Openlayer. Este componente de software conlleva una **importante carga innovadora** en relación a los clientes ligeros de visualización cartográfica clásicos donde la renderización o el tratamiento semiológico se realiza en a la parte servidora, ya que, tras recibir los datos de las variables seleccionadas al nivel escalar dependiente de la escala de visualización, **deberá proporcionar diferentes funcionalidades tanto para el tratamiento semiológico de los datos como para su filtrado temático o espacial en tiempo real en el cliente en tiempo real.** Para su desarrollo se podrían emplear librerías de visualización avanzada de datos (**D3.js**) que generan información visual en formato SVG o bien desarrollarse **WebGL**

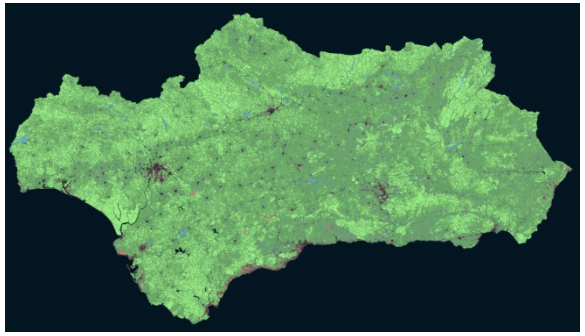
Componentes para **visualización gráfica de datos**. Dado que serán los datos nativos de las variables (no su representación cartográfica) los que llegarán al cliente, este deberá ofrecer **funcionalidades para la representación gráfica** de los datos (graficos de barras, donut, etc..) que llegan al cliente (o que se filtran en el mismo tras su llegada en el cliente) en tiempo real**.** Existen **diferentes librerías** que se podrían utilizar **(D3.js, Highcharts**, etc…)

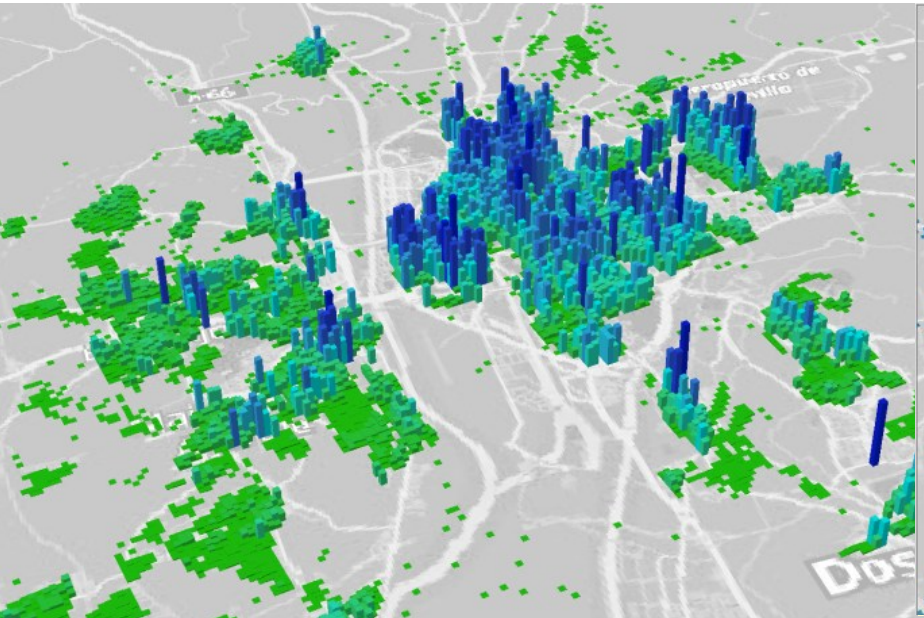
Componentes para el **análisis:** Dado que serán los datos nativos de las variables (no su representación cartográfica) los que llegarán al cliente, se abre la posibilidad para incorporar al mismo nuevas funcionalidades básicas:

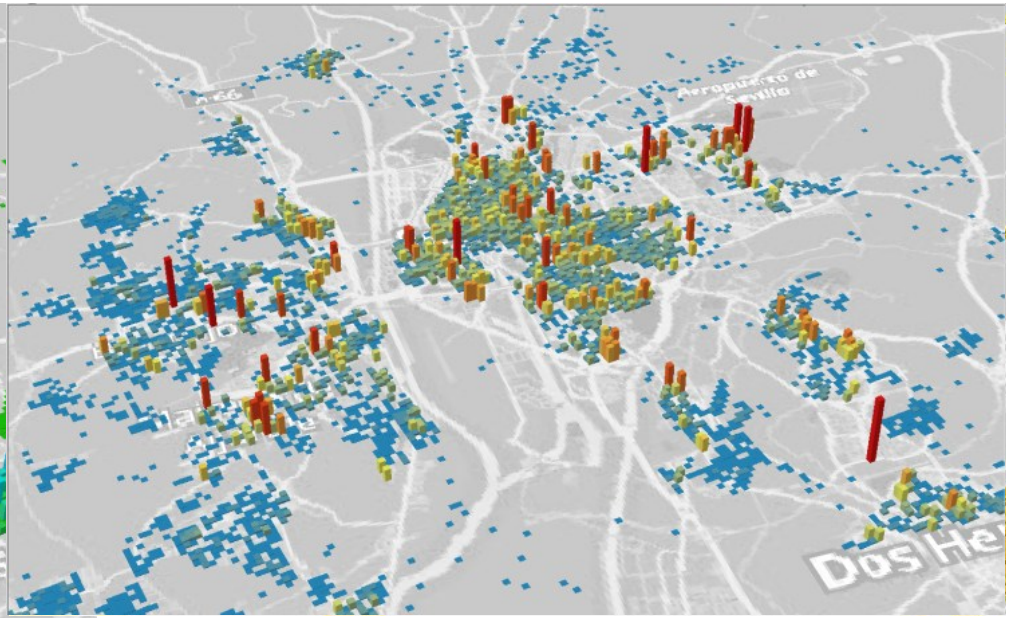
- **Calculadora de indicadores**: es de gran interés que el cliente contara con un nuevo componente para el cálculo de indicadores en tiempo real, calculados a través de funciones matemáticas básicas, pero **solo con los datos de las variables multitemáticas seleccionadas y para el nivel de rejilla que se esté visualizando.**
- **Machine Learning:** El cliente podrá llevar a cabo sencillas tareas de Machine Learning (las técnicas de clusterización K, o clasificadores), **solo con los datos de las variables multitemáticas seleccionadas y para el nivel de rejilla que se esté visualizando.**. Aunque JavaScript no es por ahora un lenguaje de referencia en el campo de Machine Learning, ya están comenzando a aparecer desarrollos orientados en esa dirección **(MLIS, TensorFlowJs…**).

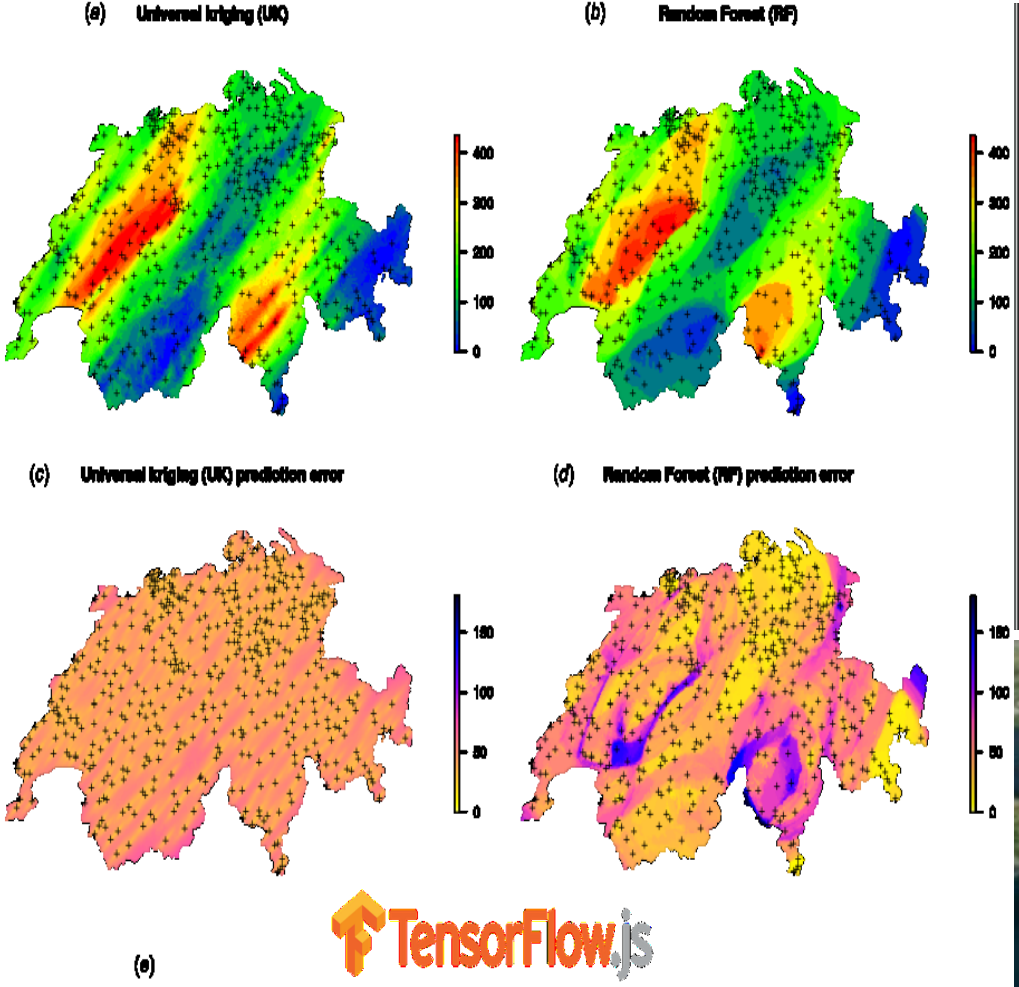
INTEGRACIÓN ESPACIAL DE VARIABLES SOCIO-AMBIENTALES EN ESTRUCTURAS DE REJILLAS MULTIESCALARES PARA SU DIFUSIÓN WEB

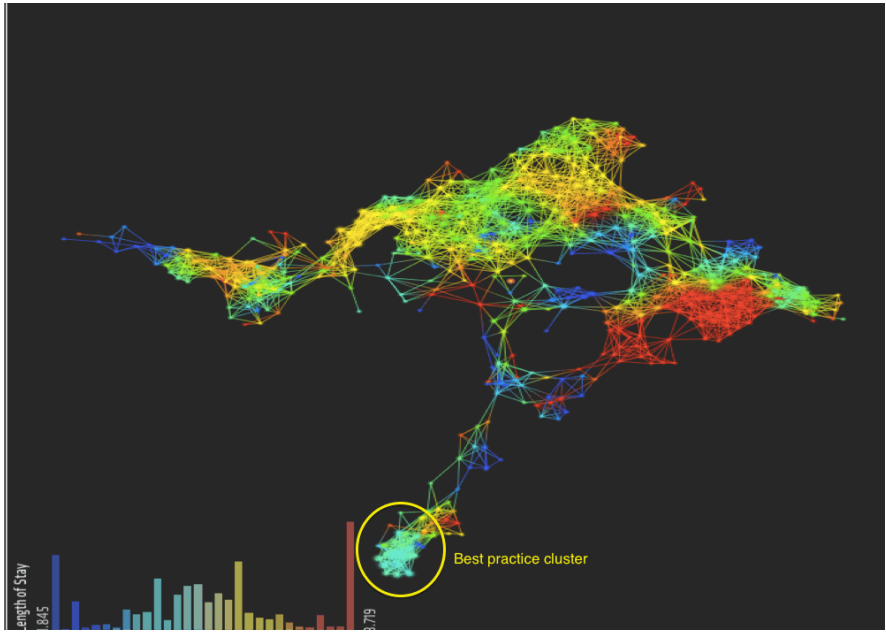


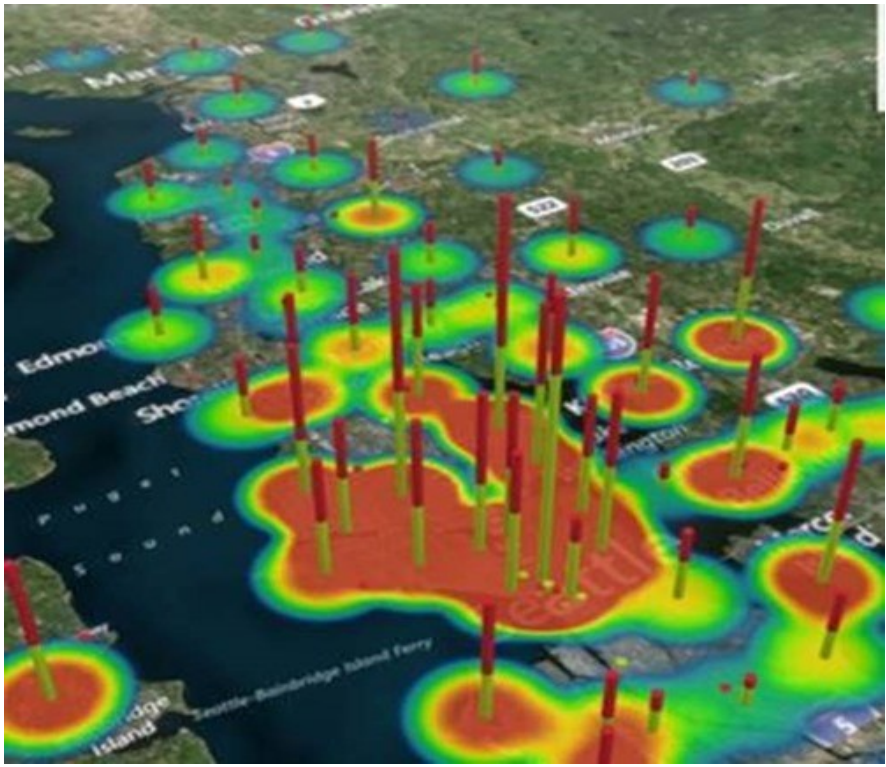


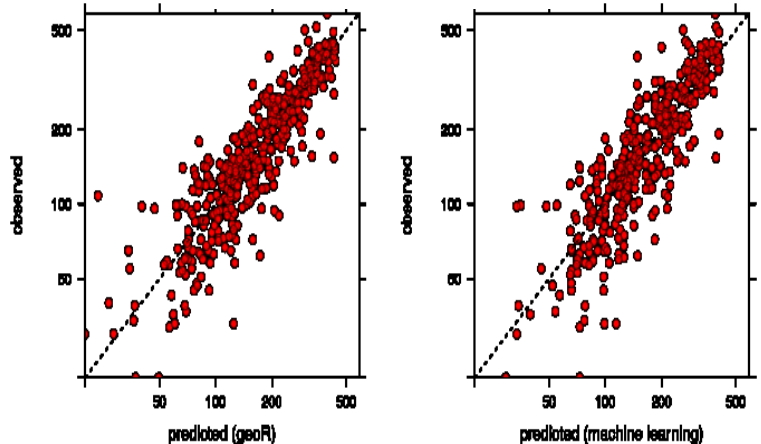


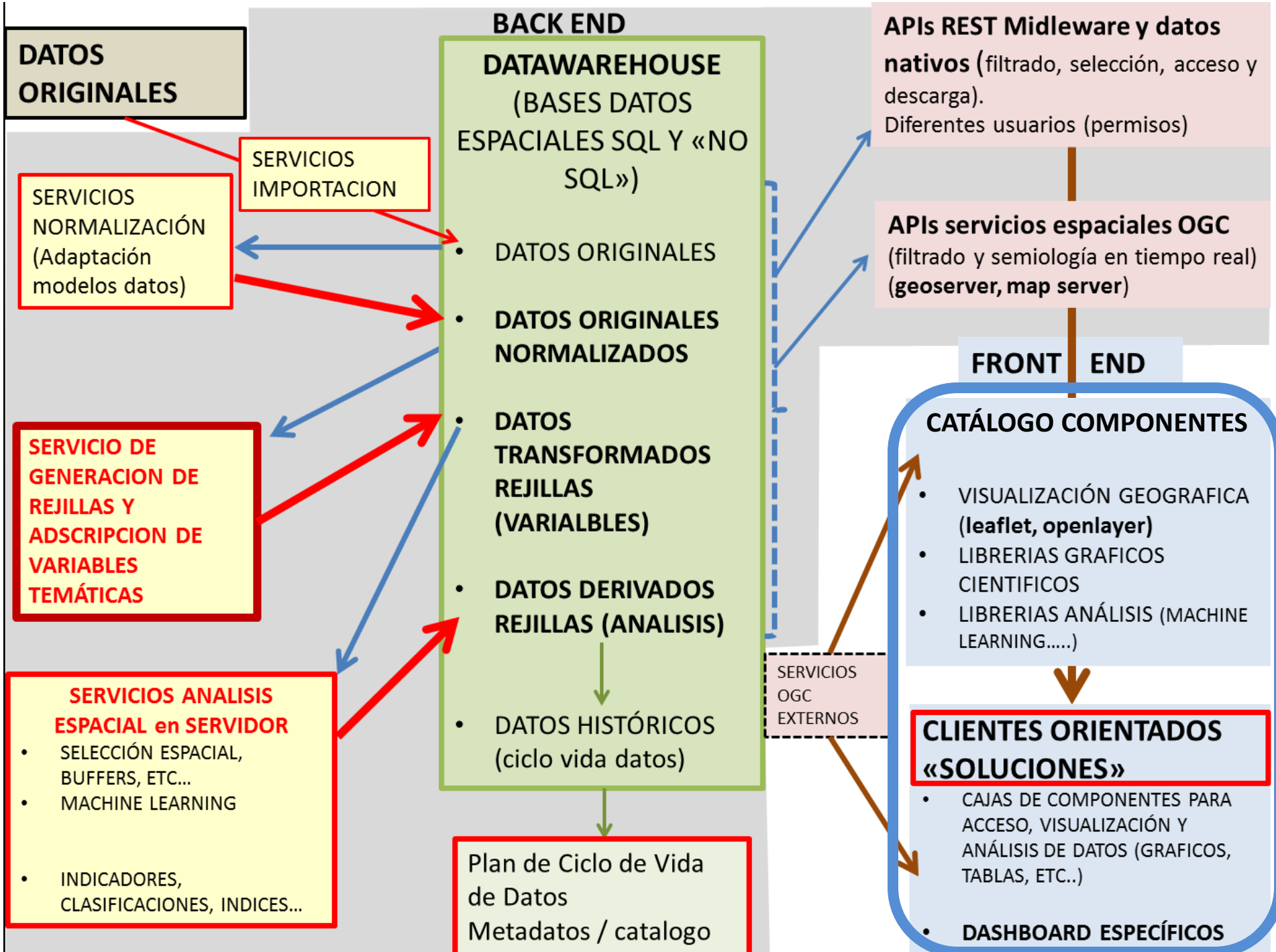



(a) Universal kriging (UK)
(b) Random Forest (RF)
400
300
200
100
0
(c) Universal kriging (UK) prediction error
(d) Random Forest (RF) prediction error
150
100
50
0


Best practice cluster

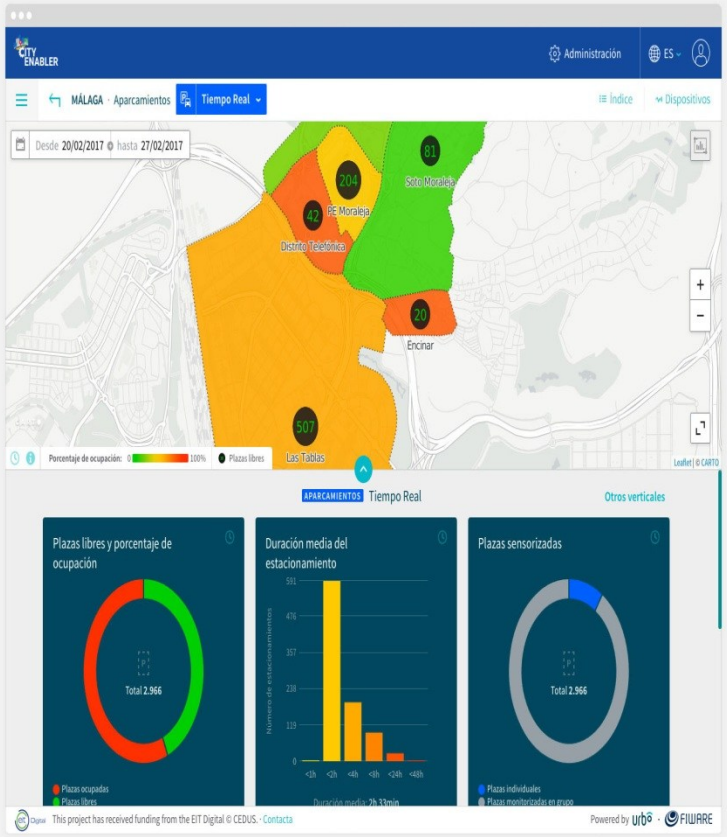


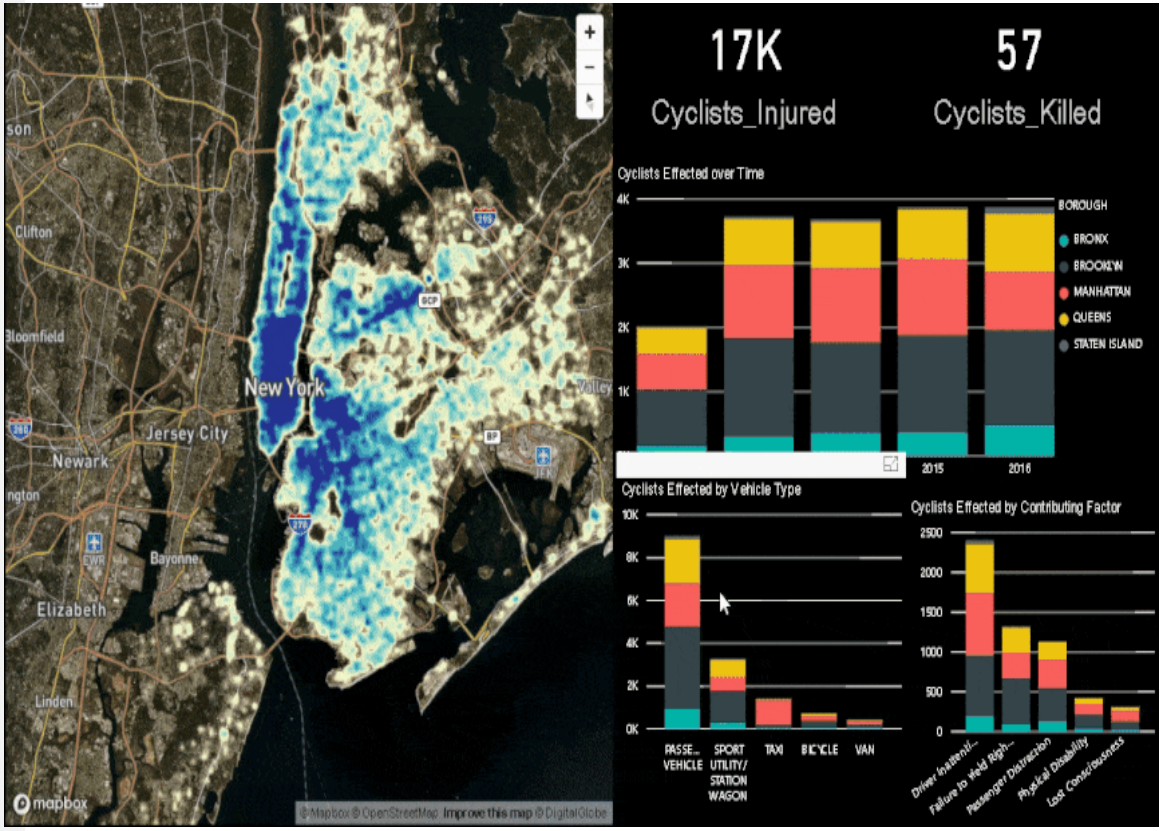

TensorFlow.js
(e)
observed
predicted (geoR)
predicted (machine learning)
50
100
200
500


BACK END
DATOS ORIGINALES
SERVICIOS IMPORTACION
SERVICIOS NORMALIZACIÓN (Adaptación modelos datos)
DATAWAREHOUSE (BASES DATOS ESPACIALES SQL Y «NO SQL»)
DATOS ORIGINALES
DATOS ORIGINALES NORMALIZADOS
DATOS TRANSFORMADOS REJILLAS (VARIALBLES)
DATOS DERIVADOS REJILLAS (ANALISIS)
DATOS HISTÓRICOS (ciclo vida datos)
SERVICIO DE GENERACION DE REJILLAS Y ADSCRIPCION DE VARIABLES TEMÁTICAS
SERVICIOS ANALISIS ESPACIAL en SERVIDOR
SELECCIÓN ESPACIAL, BUFFERS, ETC…
MACHINE LEARNING
INDICADORES, CLASIFICACIONES, INDICES…
Plan de Ciclo de Vida de Datos
Metadatos / catalogo
APIs REST Midleware y datos nativos (filtrado, selección, acceso y descarga).
Diferentes usuarios (permisos)
APIs servicios espaciales OGC (filtrado y semiología en tiempo real) (geoserver, map server)
FRONT END
CATÁLOGO COMPONENTES
VISUALIZACIÓN GEOGRAFICA (leaflet, openlayer)
LIBRERIAS GRAFICOS CIENTIFICOS
LIBRERIAS ANÁLISIS (MACHINE LEARNING…..)
SERVICIOS OGC EXTERNOS
CLIENTES ORIENTADOS «SOLUCIONES»
CAJAS DE COMPONENTES PARA ACCESO, VISUALIZACIÓN Y ANÁLISIS DE DATOS (GRAFICOS, TABLAS, ETC..)
DASHBOARD ESPECÍFICOS

CLIENTES ORIENTADOS A **SOLUCIONES: DASHBOARD**

Con los componentes antes descritos del cliente generalista es posible **la personalización y generación de una amplia gama de <u>clientes específicos claramente orientados a las potenciales "soluciones"</u>** que demanden los usuarios (técnicos administración, científicos , ciudadanos…). Para ello se debe desarrollar una **<u>aplicación/servicio</u>** que permita **seleccionar de la caja de componentes del cliente generalista** aquellos que sean de interés para una solución específica y una sencilla personalización gráfica. En cierta medida funcionaría como una especie de **<u>constructor de Dashboard</u> (visualización cartográfica+graficos+ etc).**





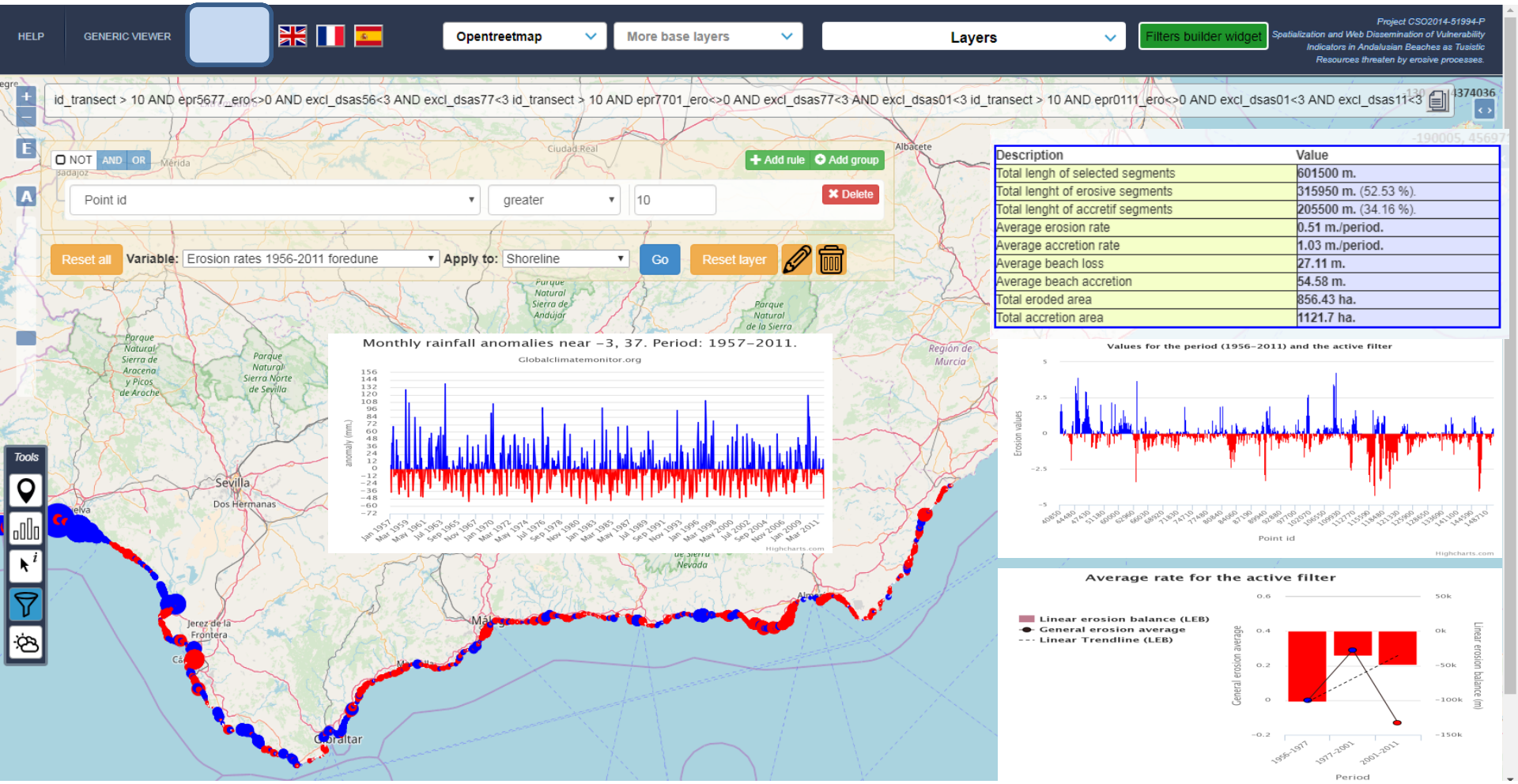
CLIENTES ORIENTADOS A **SOLUCIONES: DASHBOARD**

Con los componentes antes descritos del cliente generalista es posible **la personalización y generación de una amplia gama de clientes específicos claramente orientados a las potenciales "soluciones"** que demanden los usuarios (técnicos administración, científicos , ciudadanos…). Para ello se debe desarrollar una **aplicación/servicio** que permita seleccionar de la caja de componentes del cliente generalista aquellos que sean de interés para una solución específica y una sencilla personalización gráfica. En cierta medida funcionaría como una especie de **constructor de Dashboard (visualización cartográfica+graficos+ etc).**

En los casos de la demanda de **soluciones más complejas** que exijan algún desarrollo nuevo podrían diseñarse **nuevos componentes a partir de los microservicios ya existentes en la parte servidora.** Por ejemplo, supongamos que para una aplicación concreta se necesita poder calcular un indicador nuevo sin las restricciones del cliente genérico (solo trabaja con los datos que le llegan del nivel escalar que se corresponde con el espacio de visualización). Es decir, se necesita poder lanzar una petición de datos o calculo analítico para una extensión superficial concreta (municipio o comarca) al máximo nivel de resolución espacial de la rejilla para esa variable o variables. En este caso, podría habilitarse un **nuevo componente** (siempre que el usuario dispusiera de los correspondientes permisos y licencias) que enviase esa petición al sistema y resolver la petición con las capacidades de los microservicios desarrollados en la parte servidora.

petición de **datos o calculo analítico** para una extensión superficial concreta (municipio o comarca) **al máximo nivel de resolución espacial de la rejilla para esa variable o variables**

Se envia la petición al **sistema** y se **resuelve con las capacidades de los microservicios desarrollados en la parte servidora**



Description	Value
Total lengh of selected segments	601500 m.
Total lenght of erosive segments	315950 m. (52.53 %).
Total lenght of accretif segments	205500 m. (34.16 %).
Average erosion rate	0.51 m./period.
Average accretion rate	1.03 m./period.
Average beach loss	27.11 m.
Average beach accretion	54.58 m.
Total eroded area	856.43 ha.
Total accretion area	1121.7 ha.

CLIENTES ORIENTADOS A **SOLUCIONES: DASHBOARD**

Con los componentes antes descritos del cliente generalista es posible **la personalización y generación de una amplia gama de clientes específicos claramente orientados a las potenciales "soluciones"** que demanden los usuarios (técnicos administración, científicos , ciudadanos…). Para ello se debe desarrollar una **aplicación/servicio** que permita seleccionar de la caja de componentes del cliente generalista aquellos que sean de interés para una solución específica y una sencilla personalización gráfica. En cierta medida funcionaría como una especie de **constructor de Dashboard (visualización cartográfica+graficos+ etc).**

En los casos de la demanda de **soluciones más complejas** que exijan algún desarrollo nuevo podrían diseñarse **nuevos componentes a partir de los microservicios ya existentes en la parte servidora.** Por ejemplo, supongamos que para una aplicación concreta se necesita poder calcular un indicador nuevo sin las restricciones del cliente genérico (solo trabaja con los datos que le llegan del nivel escalar que se corresponde con el espacio de visualización). Es decir, se necesita poder lanzar una petición de datos o calculo analítico para una extensión superficial concreta (municipio o comarca) al máximo nivel de resolución espacial de la rejilla para esa variable o variables. En este caso, podría habilitarse un **nuevo componente** (siempre que el usuario dispusiera de los correspondientes permisos y licencias) que enviase esa petición al sistema y resolver la petición con las capacidades de los microservicios desarrollados en la parte servidora.

Cada uno de los nuevos componentes se quedaría almacenado en el **catálogo de componentes** de cliente y podría ser reutilizado por otros usuarios siempre que se dispusiera de los permisos y licencias para ello. De esta forma, el catálogo de componentes se iría enriqueciendo en la medida que un mayor número de usuarios utilicen el sistema.

Cloud-IA “""E-infraestructura de Información Ambiental para el Desarrollo de Soluciones Basadas en la Integración Espacial Normalizada de Variables"

OBJETO: Creación de una e-infraestructura de información ambiental integrada espacialmente, con gran capacidad de almacenamiento, gestión, análisis y difusión web en un entorno cloud, que permita la generación de soluciones a medida para gestores, investigadores y ciudadanía no expertos en TIGs, en ámbitos relacionados con la gestión ambiental y territorial.

- **REQUERIMIENTOS TEMÁTICOS DE LOS DATOS Y FUNCIONALIDADES DE LA APLICACIONES A DESARROLLAR**

- **REQUERIMIENTOS FISICOS Y LÓGICOS DEL ENTORNO CLOUD Y DE LAS APLICACIONES WEB QUE SOPORTEN LAS SOLUCIONES DEMANDADAS**