

10TH EXPERT GROUP MEETING ON

Statistical Data and Metadata eXchange

JANUARY 25-28, 2021

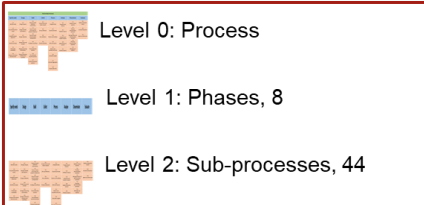
Geospatial Support in **SDMX**

28 01, 2021

Juan Muñoz (INEGI)
Hannes Reuter (Eurostat)

Overarching Processes: Use the Registry and Metadata Reports to manage metadata

Specify needs	Design	Build	Collect	Process	Analyse	Disseminate	Evaluate
1.1 Identify needs Review of existent data and metadata flows	2.1 Design outputs Design data and metadata flows	3.1 Reuse or build collection instruments Select/Build DSDs, MSDs and VTL artefacts, and decide on SDMX supported patterns	4.1 Create frame and select sample Review and make use of available SDMX data and metadata flows	5.1 Integrate data Populate databases from SDMX dataflows, update the registry	6.1 Prepare draft outputs Internal use of SDMX datasets	7.1 Update output systems Automated by SDMX, just ensure that it is connected to dissemination systems	8.1 Gather evaluation inputs Use logs from ICT infrastructure (like queries to web services) if available
1.2 Consult and confirm needs Stablish provision agreements	2.2 Design variable descriptions Design Concept Schemes, DSDs and MSDs	3.2 Reuse or build processing and analysis components Check for available tools in the SDMX.ORG website	4.2 Set up collection Run pilots to check SDMX dataflows	5.2 Classify and code Update codelists, and version DSDs and MSDs as needed	6.2 Validate outputs Check gathered information, review metadata reports	7.2 Produce dissemination products Provide flexible data and metadata services using the SDMX Registry and Web Services	8.2 Conduct evaluation Assess the benefits from actual SDMX implementation and plan for improvements
1.3 Establish output objectives Decide dataflows	2.3 Design collection Design the ICT architecture for collection, integration, exchange and dissemination	3.3 Reuse or build dissemination components Check for available tools in the SDMX.ORG website	4.3 Run collection <u>Collect/Exchange the information</u>	5.3 Review and validate Run VTL artefacts to help in the process	6.3 Interpret and explain outputs Detect discrepancies using VTL artefacts, use the SDMX datasets and metadata reports to build stories	7.3 Manage release of dissemination products Use embargo capabilities to automate this task	8.3 Agree an action plan Make the agreements to improve the actual SDMX implementation to get more benefits
1.4 Identify concepts Review cross-domain and global concepts, build a concept scheme	2.4 Design frame and sample Design constraints	3.4 Configure workflows Deploy the ICT infrastructure, build the mappings	4.4 Finalise collection Automated by SDMX, check it with VTL tools	5.4 Edit and impute Run VTL artefacts to help in the process	6.4 Apply disclosure control Check confidentiality using VTL tools, generate new dissemination DSDs/Web services	7.4 Promote dissemination products Promote the use of new enabled SDMX capabilities, like direct connection to data and metadata	
1.5 Check data availability Check data from available data and metadata flows	2.5 Design processing and analysis Design constraints, validation and transformation artefacts	3.5 Test production systems Check data and metadata flows support (ICT infrastructure and SDMX artefacts)		5.5 Derive new variables and units Run VTL artefacts to help in the process	6.5 Finalise outputs Connect SDMX datasets to dissemination systems, use of internationalization features intrinsic to SDMX	7.5 Manage user support Check the use of SDMX enabled services to be ensure alignment with user's needs	
1.6 Prepare and submit business case Design the strategy to make use of SDMX	2.6 Design production systems and workflow Finish the design of the collection and exchange architecture	3.6 Test statistical business process Run pilots to check resulting SDMX messages and reports		5.6 Calculate weights Run VTL artefacts to help in the process			
		3.7 Finalise production systems Finalize mappings and ICT infrastructure		5.7 Calculate aggregates Run VTL artefacts to help in the process			
				5.8 Finalise data files Use ICT infrastructure to move to dissemination			





SDMX and the Statistical Production Process

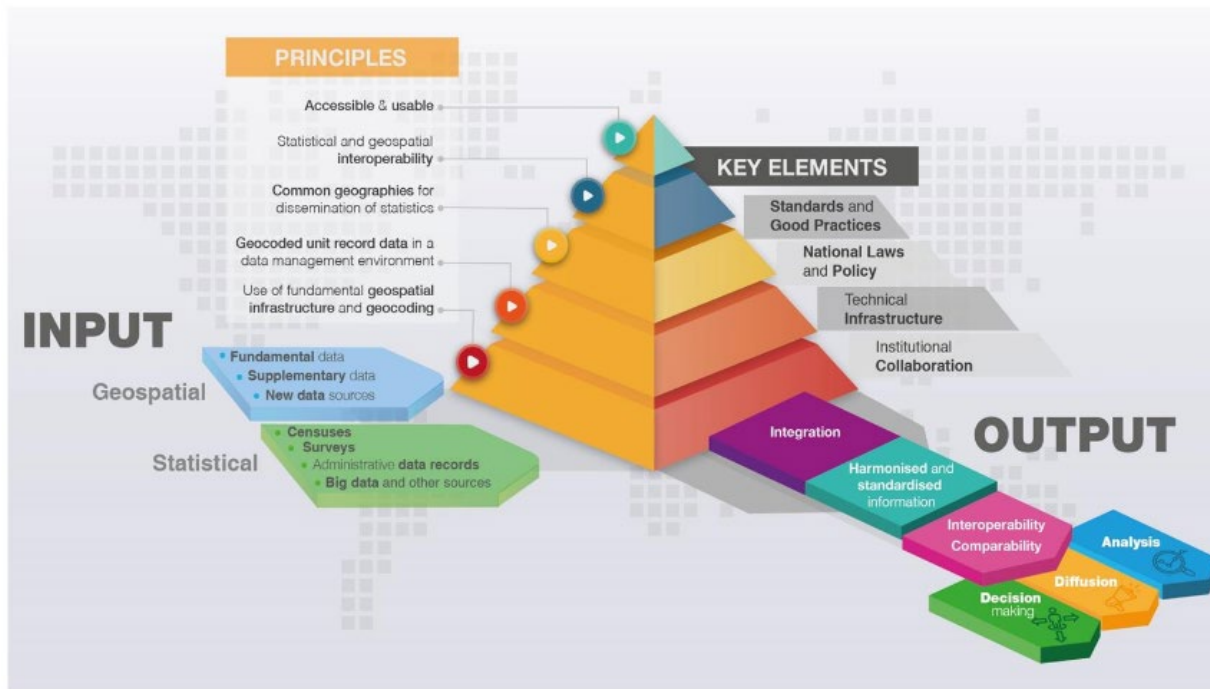
Georeferencing in the Statistical Production Process

- Design of information sets and products
- Define the sample
- Support the field work
- Enhance the Analysis
- New dissemination services



 Change Work
 Ongoing Work

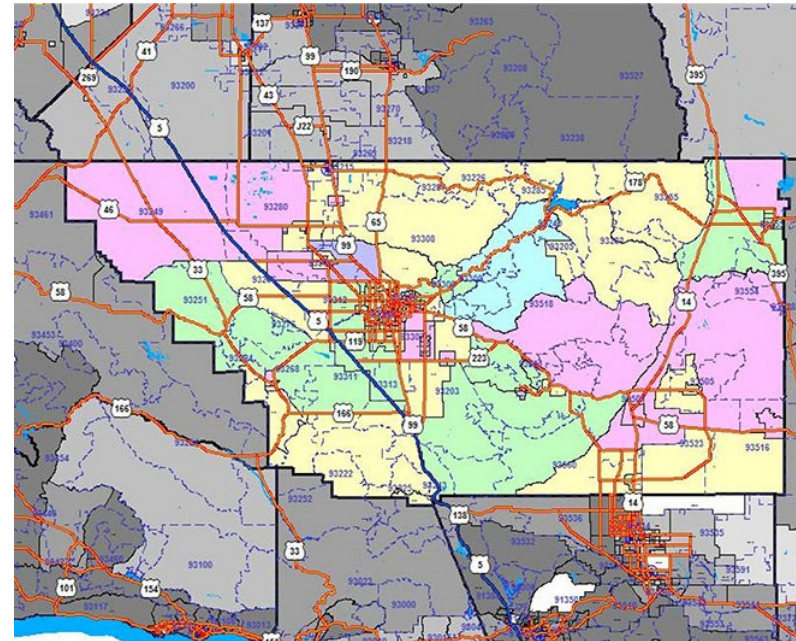
Global Statistical Geospatial Framework



- **Address**, property, building, and location information are accurate and consistent, meeting country-level agreed standards and good practices;
- Geocoding results are as accurate and consistent as possible using common approaches or systems; and,
- Any geocoding issues are consistently managed through application of standardised approaches

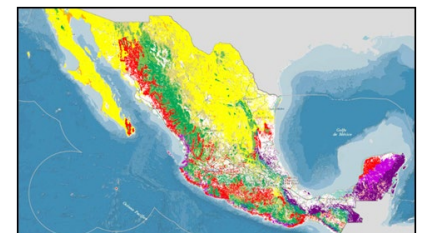
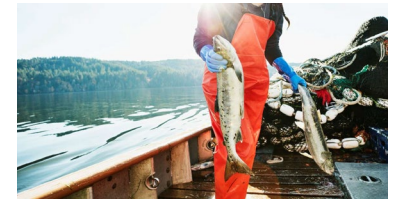
SDMX 3.0 beyond dissemination and reporting

- Statistics and Geographical Information concepts
 - Concept Schemes / Roles
 - Microdata
 - Building data lakes
- Integration of collected information
 - Internal Dataflows
 - Traditional and new sources
- Analysis of Information
 - VTL, SDMX Web Services, Metadata Reports
- Dissemination and Reporting
 - SDMX Web Services/Dataflows
 - Visualization tools



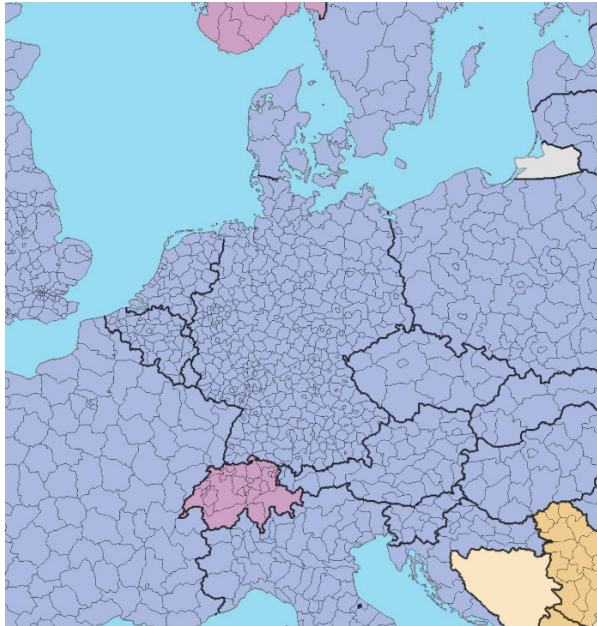
SDMX 3.0 and the Support of New Information Services: Need for Georeferencing

- Adaptive granularity
 - Variable areas
- Geographical features related to new statistical services
 - Examples: Transit over vitalities; Fish production from seas, rivers, lakes, etc.; Hospital coverage; Power Grid Coverage; etc.
- Facts correlation
 - Environmental statistics, relation with demographical and economical statistical information
- Connection Statistics with GIS systems and other georeferenced services

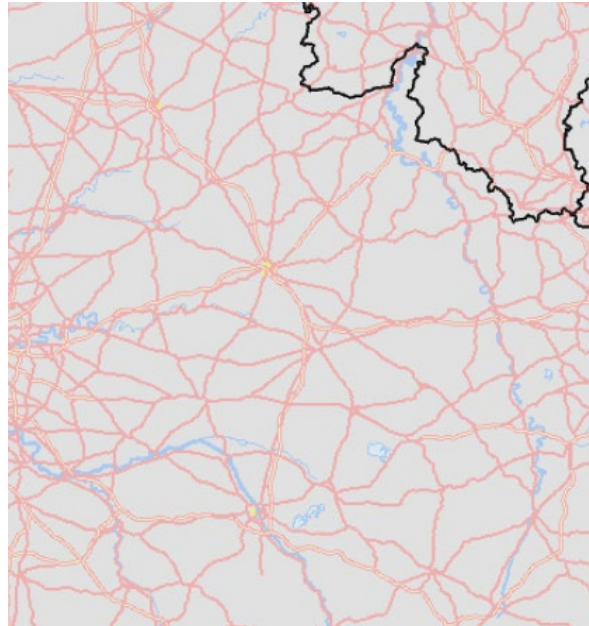


Geospatial Data Types

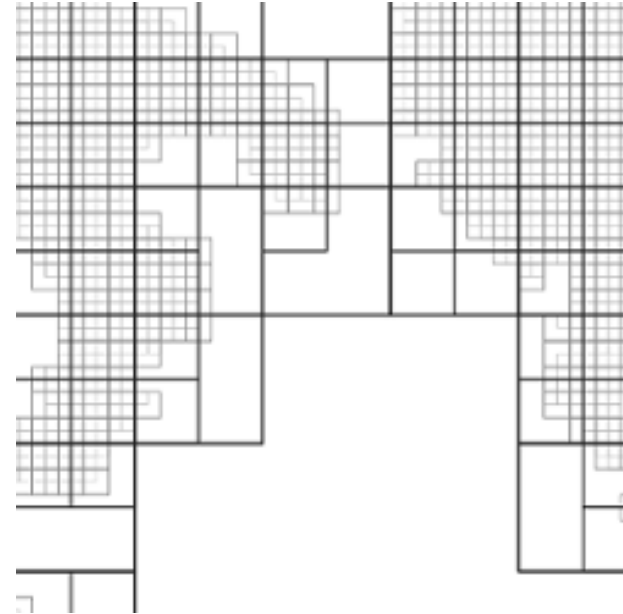
Polygons



Lines/Points



Grids



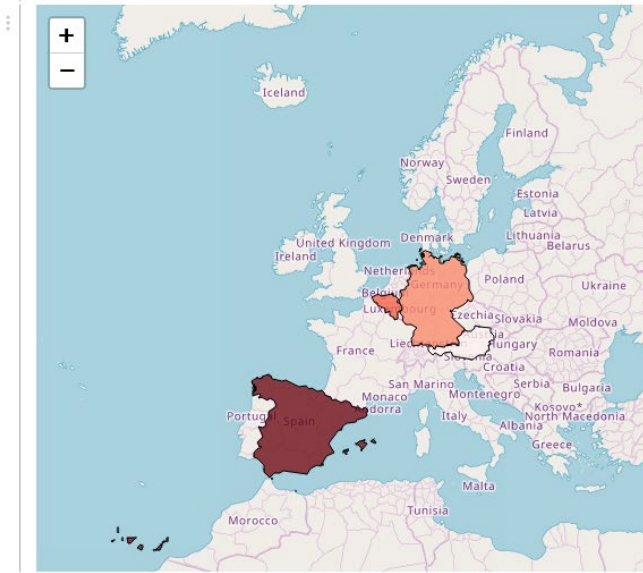
Referencing Geospatial Data

- **Indirect Reference to Geospatial Information**
- **Geospatial Coordinates**
- **A Geographic Code List**

EXAMPLE

SDMX POC

Proof of concept showing the cost of borrowing for corporations, made SDMX data using leafletJS, proj4leaflet, d3.js and xmljs.



Date

2019-09

tiles = ▶ NewClass {_url: "https://gisco-services.ec.europa.eu/maps/

map = ▶ NewClass {options: Object, _handlers: Array(6), _layers: Obj

```
colorScale = d3
  .scaleOrdinal()
  .domain(colorValues)
  .range(["#fef0d9", "#fdcc8a", "#fc8d59", "#d7301f"])
```

```
colorFunction = f(n)
```

```
colorFunction = d3.scaleLinear().domain(colorDomain)
```

Data

```
seriesData = ▶ Array(4) [Object, Object, Object, Object]
```

```
seriesData = jsonData.StructureSpecificTimeSeriesData["message:DataSet"].Series
```

```
jsonData = ▶ Object {declaration: Object, StructureSpecificTimeSeriesData: Object}
```

```
jsonData = JSON.parse(xmljs.xml2json(xmlString, { compact: true, spaces: 4 }))
```

```
xmlString = `<?xml version="1.0" encoding="UTF-8" standalone="yes"?>\r
<StructureSpecificTimeSeriesData xmlns:exr="urn:sdmx:org.sdmx.infomodel.datastructure.DataStructure=ECB:ECB_MIR1(1.0):0"
\r<message:Header>\r
\r<message:ID>Generic</message:ID>\r
\r<message:Test>>false</message:Test>\r
\r<message:Prepared>2010-01-04T16:21:49+01:00</message:Prepared>\r
\r<message:Sender id="ECB"/>\r
\r<message:Structure structureID="MIR1" dimensionAtObservation="TIME_PERIOD" namespace="urn:sdmx:org.sdmx.infomodel.d"
\r<common:Structure>\r
\r<Ref agencyID="ECB" id="ECB_MIR1" version="1.0"/>\r
\r</common:Structure>\r
\r</message:Structure>\r
\r</message:Header>\r
\r<message:DataSet data:structureRef="MIR1" xsi:type="exr:DataSetType" data:dataScope="DataStructure">\r
\r<Series FREQ="M" REF_AREA="AT" BS_REP_SECTOR="B" BS_ITEM="A2I" MATURITY_NOT_IRATE="AM" DATA_TYPE_MIR="R" AMOUNT_CAT_
\r<Obs TIME_PERIOD="2019-09" OBS_VALUE="1.3965" OBS_STATUS="A" OBS_CONF="F"/>\r
\r<Obs TIME_PERIOD="2019-10" OBS_VALUE="1.4161" OBS_STATUS="A" OBS_CONF="F"/>\r
\r<Obs TIME_PERIOD="2019-11" OBS_VALUE="1.3548" OBS_STATUS="A" OBS_CONF="F"/>\r
\r<Obs TIME_PERIOD="2019-12" OBS_VALUE="1.4418" OBS_STATUS="A" OBS_CONF="F"/>\r
\r</Series>\r
Show 21 truncated lines
xmlString = FileAttachment("ecb_mir_ts_CASE1_1 (1)@2.xml").text()
```