

## **International initiatives**

Franck Cotton – Institut National de la Statistique et des Études Économiques

# Contents

**Linked Open Statistics ESS Project**

**Unece Ontology Work**

**Statistical Data on the Web Best Practise**

**Other initiatives**

DDI

XKOS

Controlled vocabularies

DDI 4

SDMX

# Linked Open Statistics ESS Project

## What is it?

“ESSnet”: project financed by Eurostat and conducted by a consortium of national statistical institutes (here: Bulgaria, Ireland, Italy, France)

Also involved an Irish academic consortium (ADAPT / Insight / Derilinx)

In the framework of the DIGICOM “Vision Implementation Project”

Ended in May 2019

## Outcomes

Use cases, methods and tools

Lessons learned and ideas for future work

# Linked Open Statistics ESS Project – Outcomes

Data integration use case:  
“The **DG’s query**”

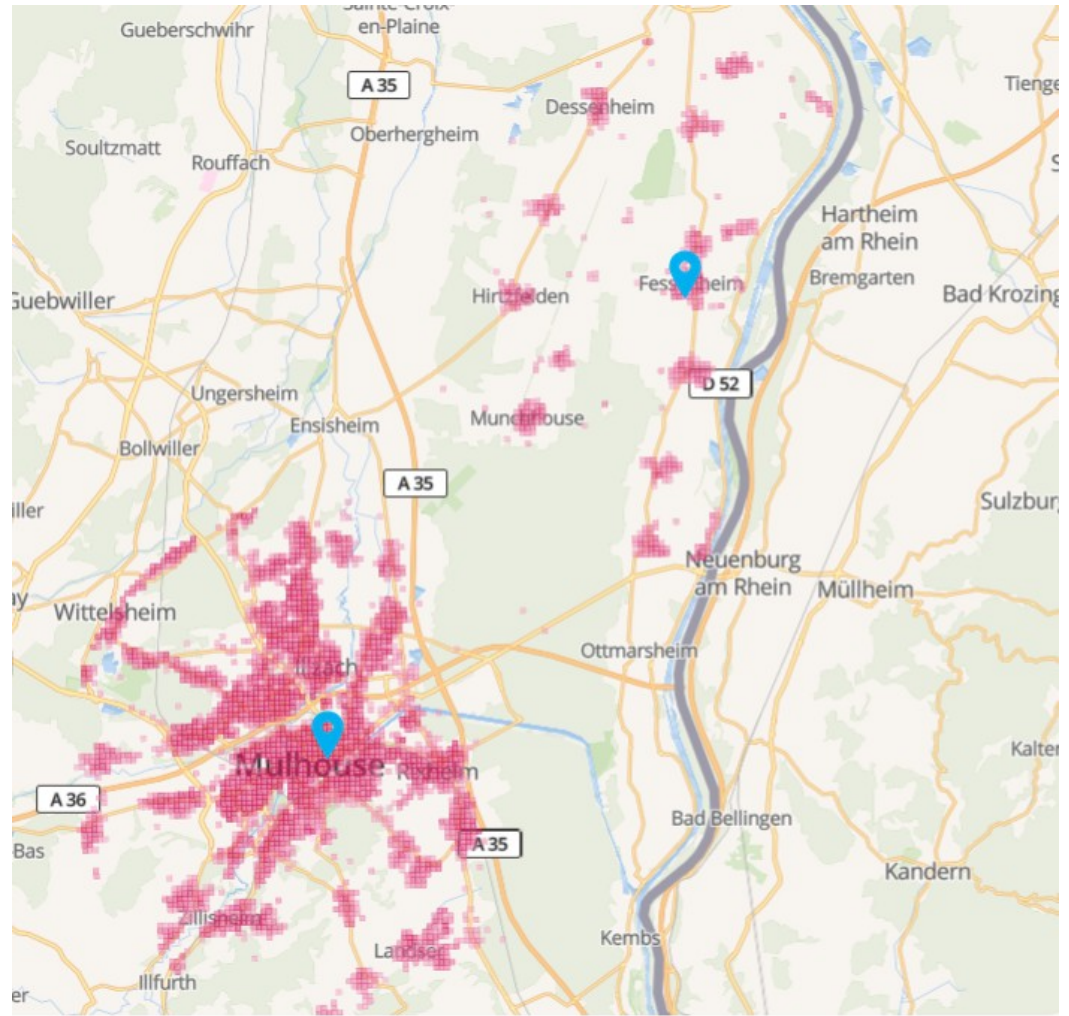
How many people live below a  
given distance of given types  
of establishments?

Here:

Establishment > 500 employees

Activity: NACE 35.11 - Production of  
electricity

Distance: 10 km



# Linked Open Statistics ESS Project – Outcomes

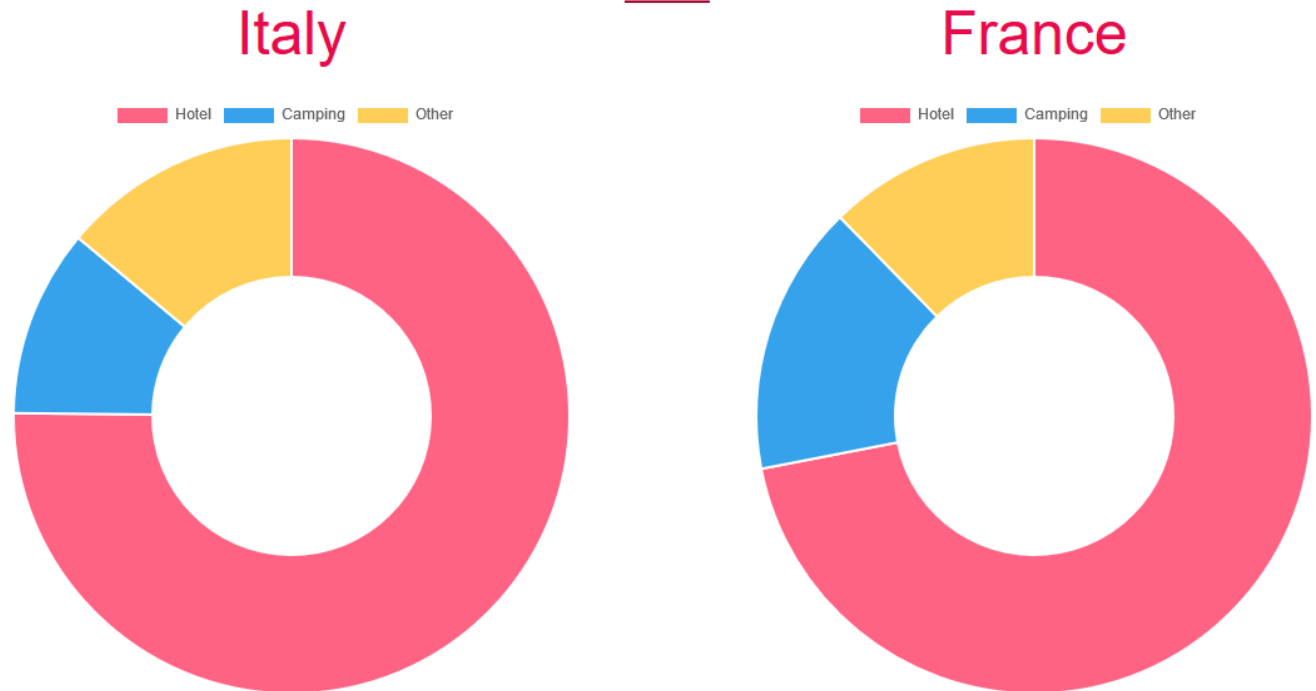
Use case on  
**comparison** between  
states (distributed  
queries)

Here:

Type of accommodation  
used by tourists in 2017,  
in percentages

Type of housing

Year  
2017



# Linked Open Statistics ESS Project – Outcomes

## Prototype of a Linked Open Statistical Data Hub

### Tools to:

transform data to RDF

browse data cubes

The screenshot shows the homepage of the Linked Open Statistical Data Hub. At the top is a dark blue navigation bar with the project logo on the left and links for 'DATASETS', 'ORGANIZATIONS', and 'ABOUT' on the right, along with a search bar. Below the navigation bar is a large heading 'Welcome to Linked Open Statistical Data Hub' followed by a brief description of the hub as a catalog of datasets and resources. A row of four blue buttons with icons and text provides access to various tools: 'Data conversion using Juma', 'Querying our RDF Store using SPARQL or GraphQL', 'Visualization and Analysis using OLAP Browser or Cube Visualiser', and 'Learning Platform'. The main content area displays a grid of dataset cards from various national statistical offices. The first column includes cards for the Central Statistics Office (Ireland) with datasets like 'ILO Unemployment Rate for Ages 15-74 by NUTS region Ireland', 'SILC\_IE\_Hackathon', and 'Irish HC55 Census Data'; and the National statistical institute (Bulgaria) with 'Bulgarian LFS', 'Bulgarian Tourism Data', and 'Bulgarian SILC'. The second column includes cards for Istat (Italy) with 'CIS' and 'ILO Unemployment Rate for Ages 15-74 by NUTS region Italy', and Insee (France) with 'French Tourism Data' and 'ILO Unemployment Rate for Ages 15-74 by NUTS3 regions France Quarterly'.

# Linked Open Statistics ESS Project - Outcomes

## Lessons learned

Technology not really an issue

Lot of work left to have data comparability at European level

Examples

- Measurement of population in the Census

- Seasonal adjustment of results

- Codes for NUTS

Eurostat to publish “key semantic assets”

**DIGICOM final event on 26 and 27 November 2019 in Brussels**

# Unece Ontology Work

## Unece

UN Economic Commission for Europe (membership well beyond Europe)

Supports a “High-level group” for the Modernization of Official Statistics

The HLG oversees various activities led through projects or groups

Activities branded as The logo for ModernStats features the word 'modernstats' in a lowercase, sans-serif font. The letters are black, but the 'o' and 'd' are stylized with horizontal bars in red, orange, and green. Below the word, the text 'by HLG - MOS' is written in a smaller, black, sans-serif font.

## ModernStats standards

A lot of models, frameworks: GSBPM, GSIM, GSDEM, GAMSO, CSPA, CSDA...

Problems: isolated and not always coherent

Work started to formalize a high-level framework: the COOS



# Unece Ontology Work – COOS

## Core Ontology for Official Statistics

### Objectives

- Link ModernStats Models to one another

- Link with external models (SKOS, PROV-O, ORG...)

- Provide machine-actionable version of main features of the models

### Timetable

- 2017: first [proof of concept](#)

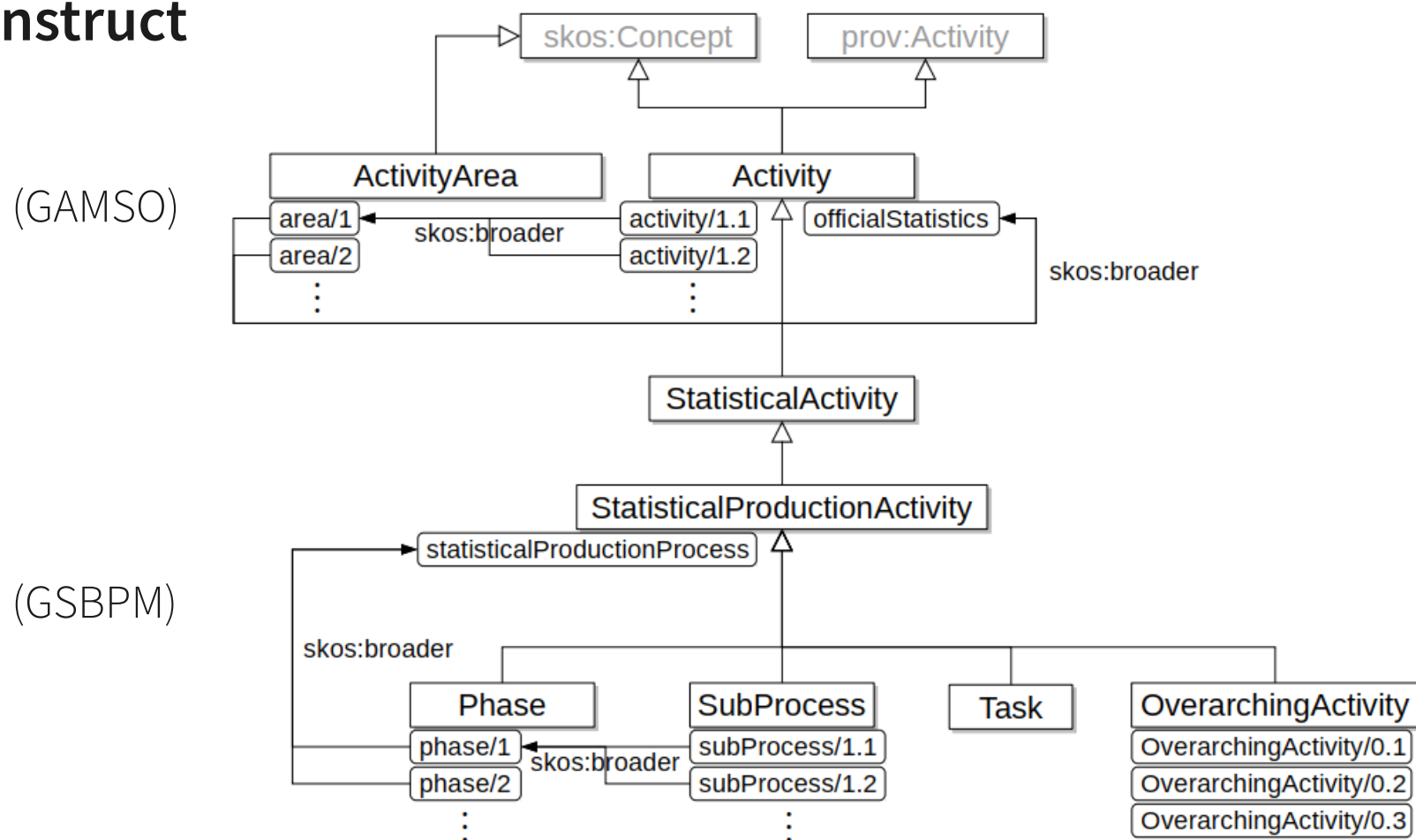
- November 2018: activity launched

- Today: first version of ontology created by working group

- Next HLG-MOS workshop (end of November) to greenlight next steps

# Unece Ontology Work – COOS

## Example of construct



# Unece Ontology Work – COOS

## Next steps

Marginal additions to the ontology

Detailed description of components

Public review

Document on the governance of the COOS

Presentation to November 2020 HLG-MOS workshop for adoption

## Contributions welcome!

<https://github.com/linked-statistics/COOS>

# Statistical Data on the Web Best Practise

## What is it?

Document good/best practices for sharing statistical data via the web

Inspired by similar efforts for data on the web and spatial data on the web

Group effort through W3C Semantic Statistics Community Group and W3C/OGC Spatial Data on the Web Interest Group

Coordinated by Bill Roberts (Swirrl) and Franck Cotton (INSEE)

# Statistical Data on the Web Best Practise

## Challenges

Data discovery

Understanding methodology

Classification schemes and aggregation

Annotations and 'data markers'

Versioning

Use of data in common tools, such as R

# Statistical Data on the Web Best Practise

## Challenges

Agree on terminology

List issues with RDF Data Cube

Identify connections between existing standards

Take a generic approach – not narrowly focused on RDF but also considering other formats, RESTful APIs, etc.

# Statistical Data on the Web Best Practise

## Timetable

First call took place on 14 October

Conference calls about once a month

Most work via GitHub and mailing lists

Aim for complete document(s) around October 2020

## Contributions welcome!

# Other Initiatives

## DDI

XKOS

Representation of statistical classifications in RDF

[Officially published](#) as a DDI standard

Work on future version and best practice document

Controlled vocabularies

DDI4

## SDMX



# **Thank you for your attention**

**Any questions?**