

Statistical Viewer

Coping with the Complexity



Overview

- What's DataCove and Ghosting
- EuroStat as a Data Resource
- DataCove Statistical Viewer
- The Was from Here to There
- Conclusions

DataCove e.U.

- **Data Modelling:** OGC & ISO / INSPIRE based data modelling
- **Data Harmonization:** Transforming existing data to align with standardized data models
- **Data Provision:** Making data available via OGC & ISO / INSPIRE compliant web services (WMS, WFS, SOS)
- **Ghosting:** Provision of existing data aligned with INSPIRE
- **Observations:** Special focus on observational data on spatial objects (O&M Data Model, Sensor Web)

DataCove INSPIRE Ghosts

Various datasets online with INSPIRE Formats

- CDDA Protected Sites
- Land Use Data
- Utility & Government Services (Schools)
- Eurostat data for Statistical Units and Population Distribution (more than demography!)
- Eurostat dictionary as GML Dictionary







eurostat

- Statistical Office of the European Union
- Wide range of statistical data, i.e.
Demography, Employment, Education,
Migration, Economy
- Spatial Data available, but NUTS regions or
topology

<http://ec.europa.eu/eurostat/data/database>

Tables by Theme

Data Navigation Tree

- Database by themes
 - Tables by themes
 - General and regional statistics
 - Regional statistics (t_reg)
 - Regional agriculture statistics (t_reg_agr)
 - Regional demographic statistics (t_reg_dem)
 - Regional economic accounts - ESA 95 (t_reg_eco)
 - Regional education statistics (t_reg_educ) **M**
 - Regional science and technology statistics (t_reg_sct)
 - Regional health statistics (t_reg_hlth)
 - Regional tourism statistics (t_reg_tour)
 - Regional transport statistics (t_reg_tran) **M**
 - Regional labour market statistics (t_reg_lmk) **M**
 - Regional digital economy and society (t_reg_isoc)
 -  **ZIP** Households that have internet access at home by NUTS 2 regions (tgs00047) **i**
 -  **ZIP** Households that have broadband access by NUTS 2 regions (tgs00048) **i**
 -  **ZIP** Percentage of households with broadband access in relation to households with internet access, by NUTS 2 regions (tgs00049) **i**
 -  **ZIP** Individuals regularly using the internet by NUTS 2 regions (tgs00050) **i**
 -  **ZIP** Individuals who have never used a computer by NUTS 2 regions (tgs00051) **i**
 -  **ZIP** Individuals who ordered goods or services over the internet for private use in the last year by NUTS 2 regions (tgs00052) **i**
 - Regional environmental statistics (t_reg_env)
 - Regional poverty and social exclusion statistics (t_reg_ilc) **M**
 - Non EU countries (t_noneu)
 - Land cover and land use, landscape (LUCAS) (t_lan) (Important information concerning comparability)

Data Explorer

Key farm variables: area, livestock (LSU), labour force and standard output (SO)

Last update: 29-07-2016

Table Customization [show](#)

TIME + GEO
 + hold: Total number of holdings

GEO	2005	2007	2010	2013
Belgium	51,540	48,010	42,850	37,760
Région de Bruxelles-Capitale	20	20	20	40
Prov. Antwerpen	5,440	5,030	4,330	3,720
Prov. Limburg (BE)	4,470	4,060	3,510	2,980
Prov. Oost-Vlaanderen	9,000	8,380	7,490	6,550
Prov. Vlaams-Brabant	4,340	3,970	3,410	2,950
Prov. West-Vlaanderen	11,160	10,550	9,590	8,680
Prov. Brabant Wallon	1,290	1,230	1,110	1,030
Prov. Hainaut	5,180	4,860	4,470	3,970
Prov. Liège	4,390	4,070	3,660	3,160
Prov. Luxembourg (BE)	3,170	2,980	2,670	2,350
Prov. Namur	3,080	2,860	2,600	2,320
Bulgaria	534,610	493,130	370,490	254,410
Severozapaden	96,710	90,330	51,340	30,430
Severen tsentralen	72,470	63,780	43,320	28,670
Severozitochen	64,420	56,290	43,780	28,030
Yugoiztochen	78,050	66,940	56,980	36,790
Yugozapaden	85,510	82,220	65,560	52,130
Yuzhen tsentralen	137,460	133,570	109,510	78,350
Czech Republic	42,250	39,400	22,860	26,250
Praha	330	280	70	60
Střední Čechy	4,730	4,470	3,030	3,520
Jihozápad	6,740	6,250	4,400	5,050
Severozápad	2,690	2,330	1,750	1,940
Severovýchod	6,620	6,240	3,910	4,330
Jihovýchod	12,710	12,030	5,590	6,340
Střední Morava	5,330	4,900	2,570	3,150
Moravskoslezsko	3,090	2,890	1,550	1,860
Denmark	51,680	44,620	42,100	38,830
Hovedstaden	2,630	2,420	2,290	2,150
Sjælland	8,690	7,850	7,110	6,430
Syddanmark	14,830	12,410	11,730	10,500
Midtjylland	16,030	13,630	12,840	11,990
Nordjylland	9,480	8,310	8,120	7,760
Germany (until 1990 former territories)	389,880	370,480	299,130	285,030
Baden-Württemberg	59,320	56,940	44,510	42,410
Bayern	126,810	120,960	97,870	93,270
Berlin, Bremen, Hamburg	1,340	1,280	:	:
Berlin	:	90	70	60
Brandenburg	6,640	6,670	5,570	5,400
Bremen	:	210	160	160

[Explanatory texts \(meta\)](#)

Key farm variables: area, livestock (LSU), labour force and standard output (SO) by agricultural holding and NUTS 2

Last update: 29-07-2016

Interactive extraction size limit

Current extraction size:

Dimension selection:

AGRAREA GEO INDIC

View

Sorting Sort Ascending Sort Descending

Show Codes Labels

Filtering

Filtering type: Text Numbers

<input type="checkbox"/> Select all	Cod
<input checked="" type="checkbox"/>	HO
<input type="checkbox"/>	AGL
<input type="checkbox"/>	C_L
<input type="checkbox"/>	LFC
<input type="checkbox"/>	SO
<input type="checkbox"/>	TO

Downloads



Full extraction [1 data tables]

On the same sheet

On separate sheets

Flags and footnotes

On the same sheet

On separate sheets

Separator

With thousand separator (format dependent on Excel configuration)

Without thousand separator



Full extraction [1 data tables]

Single file

Multiple files

Flags and footnotes

Cell formatting

1.234,56

1.234.56

1 234.56



Full extraction [1 data tables]

Flags and footnotes

Cell formatting

1.234,56

1.234.56

1 234.56



Full extraction [1 data tables]

Flags and footnotes




Full extraction [1 data tables]

Flags and footnotes

Data Explorer 2

Table ▾
Graph ▾
Map ▾



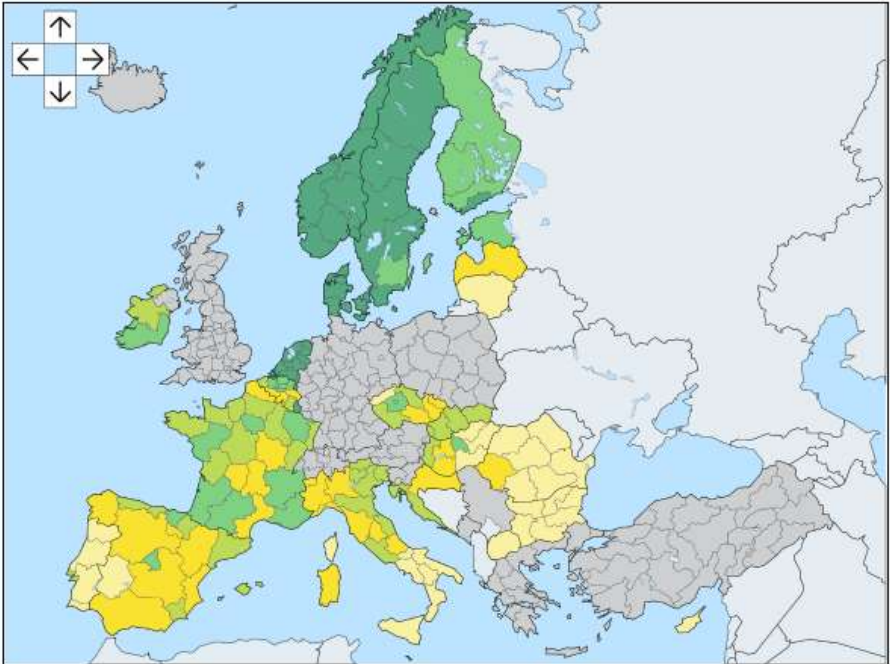
2.4.2-r2159-2016-08-11 (PROD) [Online support](#) [Legal Notice](#)

Households that have internet access at home by NUTS 2 regions

% of households with at least one member aged 16 to 74

The access of households to internet is measured as percentage of households where any ... [more](#)

Code: tgs00047



Click on map to:

Recenter Zoom In/Recenter Zoom Out/Recenter

Types
Data
Classes
Layers
Hide toolbox ▾

Map Type

Coloured Map

Symbol Map by size

Colour with: black

Symbol Map by color

Comparison

No Comparison

Compare with geo Rég. Bruxelles / ▾

Compare with time 2015 ▾

Compare with value

Refresh

© EuroGeographics Association for the administrative boundaries

- 
- What could we do if this data was available via INSPIRE...

At DanubeHack2 we started seeing what we could do with INSPIRE compliant services for:

- `suv:VectorStatisticalUnit`
- `pd:StatisticalDistribution`

Statistical Data Viewer App

Select
Measure:

No measure selected

No measure selected

Life expectancy by age, sex and NUTS 2 region

Population on 1 January by five years age group, sex and NUTS 2 region

Population aged 25-64 by educational attainment level, sex and NUTS 2 regions (%)

Early leavers from education and training by sex and NUTS 2 regions

Hospital beds by NUTS 2 regions

Health personnel by NUTS 2 regions

Average number of usual weekly hours of work in main job by sex, age and NUTS 2 regions (hours)

Asylum and first time asylum applicants - monthly data (rounded)

Asylum and first time asylum applicants - annual aggregated data (rounded)

First instance decisions on asylum applications by type of decision - annual aggregated data



Statistical Data Viewer App

Select Measure: Hospital beds by NUTS 2 regions

Select facility: HBEDI_PSY

- Select unit:
- HBEDI_PSY
 - HBEDT
 - HBEDT_CUR
 - HBEDT_LT
 - HBEDT_OTH
 - HBEDT_REH

Get Dataset (0)



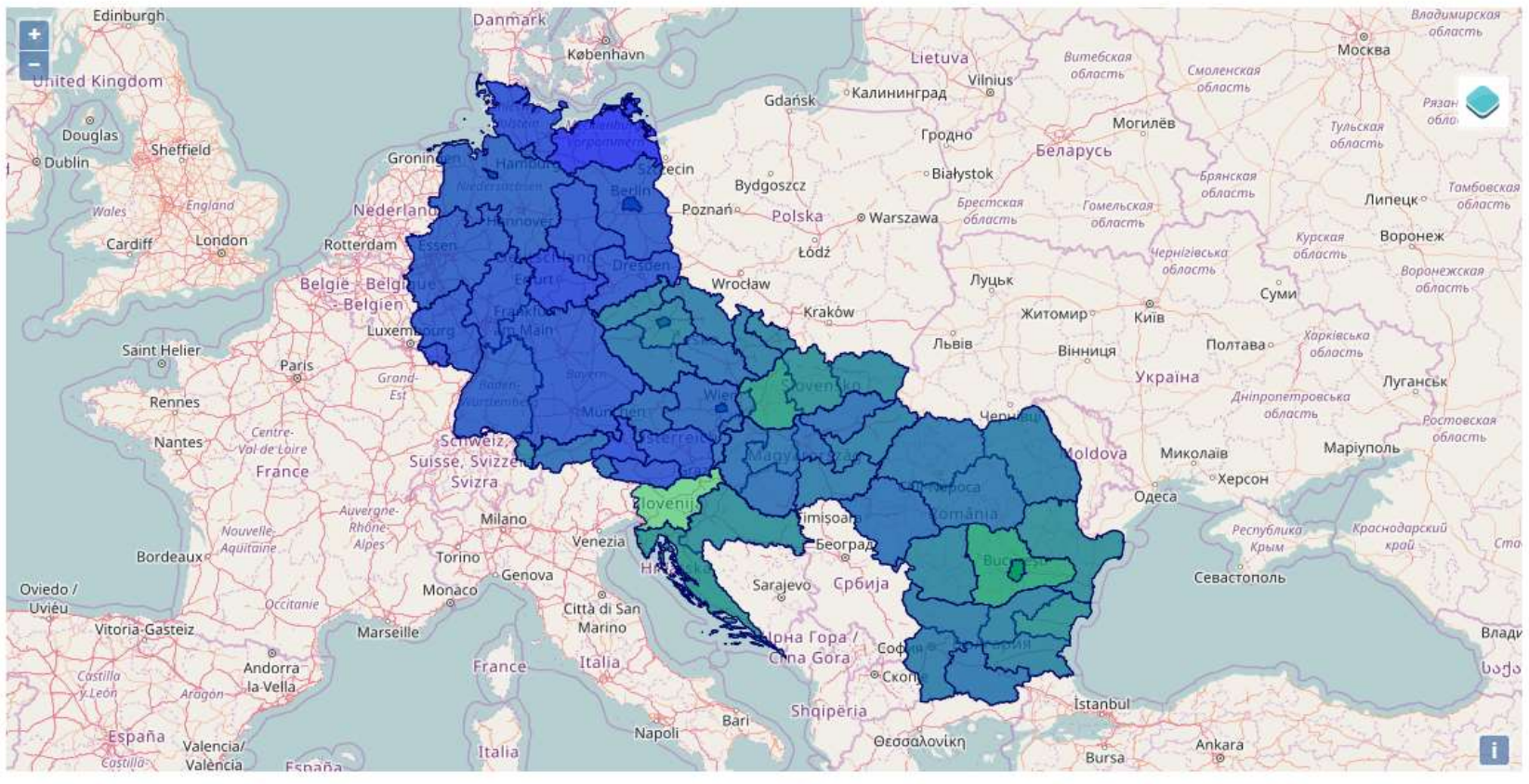
Statistical Data Viewer App

Select Measure: Hospital beds by NUTS 2 regions

Select facility: HBEDT

Select unit: HAB_P

Get Dataset (1)



pd:dimensions	pd:value	Legend
Measure: Hospital beds by NUTS 2 regions (Map layer on)		
http://data.datacove.eu/suv/RO	151.6068	RO
http://data.datacove.eu/suv/RO11	143.2429	RO11
http://data.datacove.eu/suv/RO12	145.0874	RO12
http://data.datacove.eu/suv/RO21	158.7682	RO21
http://data.datacove.eu/suv/RO22	175.6768	RO22
http://data.datacove.eu/suv/RO31	210.722	RO31
http://data.datacove.eu/suv/RO32	101.8395	RO32
http://data.datacove.eu/suv/RO41	166.4265	RO41
http://data.datacove.eu/suv/RO42	137.9413	RO42
http://data.datacove.eu/suv/AT	130.3117	AT
http://data.datacove.eu/suv/AT11	166.8048	AT11
http://data.datacove.eu/suv/AT12	142.5129	AT12
http://data.datacove.eu/suv/AT13	126.8926	AT13
http://data.datacove.eu/suv/AT21	112.5228	AT21
http://data.datacove.eu/suv/AT22	114.7306	AT22
http://data.datacove.eu/suv/AT31	138.7597	AT31
http://data.datacove.eu/suv/AT32	104.2626	AT32
http://data.datacove.eu/suv/AT33	142.9471	AT33
http://data.datacove.eu/suv/AT34	166.636	AT34
http://data.datacove.eu/suv/SK	169.1825	SK
http://data.datacove.eu/suv/SK01	124.0053	SK01
http://data.datacove.eu/suv/SK02	212.0759	SK02

Statistical Data Viewer App

Select Measure: Health personnel by NUTS 2 regions

Select unit: HAB_P

Select isco08: OC221

Calculator

Loaded datasets:

- A. hlth_rs_bdsrg(52)
- B. hlth_rs_bdsrg(52)
- C. hlth_rs_prsrg(69)

Equation: $(a+b)/c$

Eval

Close

Get Dataset (3)

Open calculator



How did we get there?

- Eurostat Data Ghosted using INSPIRE Data Specifications and GeoServer:
 - `suv:VectorStatisticalUnit`
 - `pd:StatisticalDistribution`
- Viewer using OpenLayers
- But – there were still quite a few bits missing!

Missing bits

- Queries - no standardized stores queries in INSPIRE:
 - List of all measures
 - List of classifications with items per measure
 - Features available for a specific measure
- Filters – some `pd:StatisticalDistribution` features large
 - Filter by selected classification items

PHP Filter

- Depending on Statistical Measure selected, features can be very large (the Data, not the Geometry!)
- Statistical Data Provided for different classifications:
 - Male/Female/Total
 - Age Groups
 - Level of Education
 - ...
- Only one value can be visualized, so filter before download!

PHP Filter

- Provide FeatureId & Classification Items (and Service URI if not default). Example:
 - classItem=http://code.datacove.eu/codeList/age/Y_GE25,<http://code.datacove.eu/codeList/sex/M>
 - featureId=AT_lfst_r_lfu3rt_age_sex_2015
- Returns Feature as JSON, only values corresponding to classItem
- http://bolegweb.geof.unizg.hr:2017/danubehack2/pd-viewer/api/PD_Filter2.php

Problems Encountered (GeoServer)

- No Stored Queries on Complex Features
 - Made simple features from same tables
- No Stored Queries on Views (simple feature)
 - Created tables instead of views
- Problem with WMS on MultiSurface
 - WMS from Simple Features
- Crash when querying multiple complex features
 - Query features individually

Architecture

OGC Services

INSPIRE (WFS):

- Statistical Units (SU)
- Population Distribution (PD)

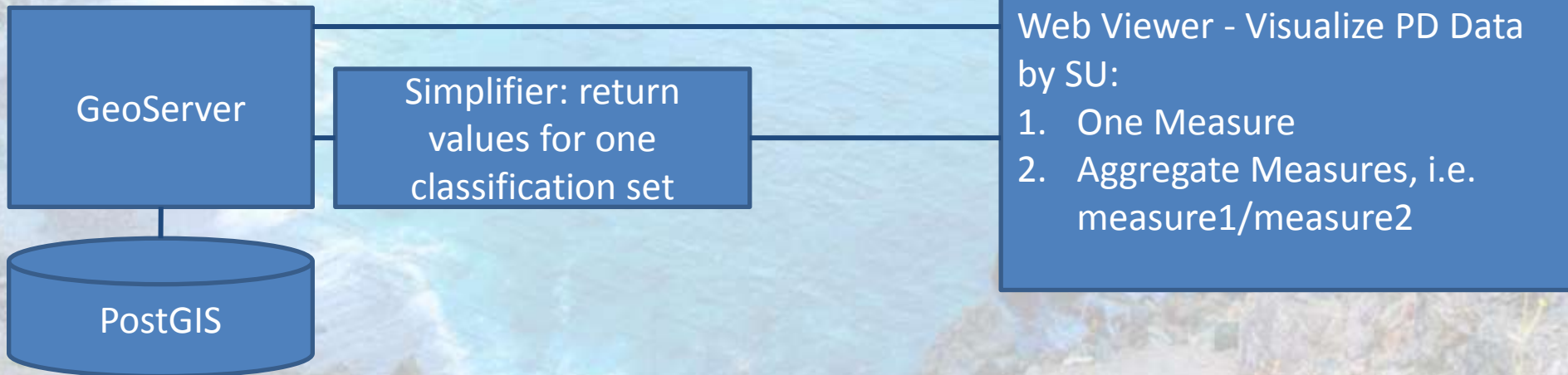
Additions:

Simple Feature SU (for WMS)

Simple Feature: All Measures for PD

Simple Feature + Stored Query: Classifications for Measure PD

Simple Feature + Stored Query: FeatureId for Measure, Country & Year



Insights for INSPIRE

- Fitness for Use checks would be valuable: while PD data logically structured, not ideal for application development
 - include real use cases & usability checks
- Standardized Stored Queries for Data Themes
 - some usage dependent, but many reusable
- PD Structure logical in statistical context, difficult in spatial context
 - usability checks

Conclusions

- More and more data online and available, but not much harmonized (→ essential for reuse!)
- INSPIRE Data Specifications good start, but:
 - Difficult to implement with current technology
 - Standardized Stored Queries missing
 - Fitness for Use should be further analysed
- Important for enabling applications across data sources

Statistical Viewer:

<http://bolegweb.geof.unizg.hr:2017/danubehack2/pd-viewer/>

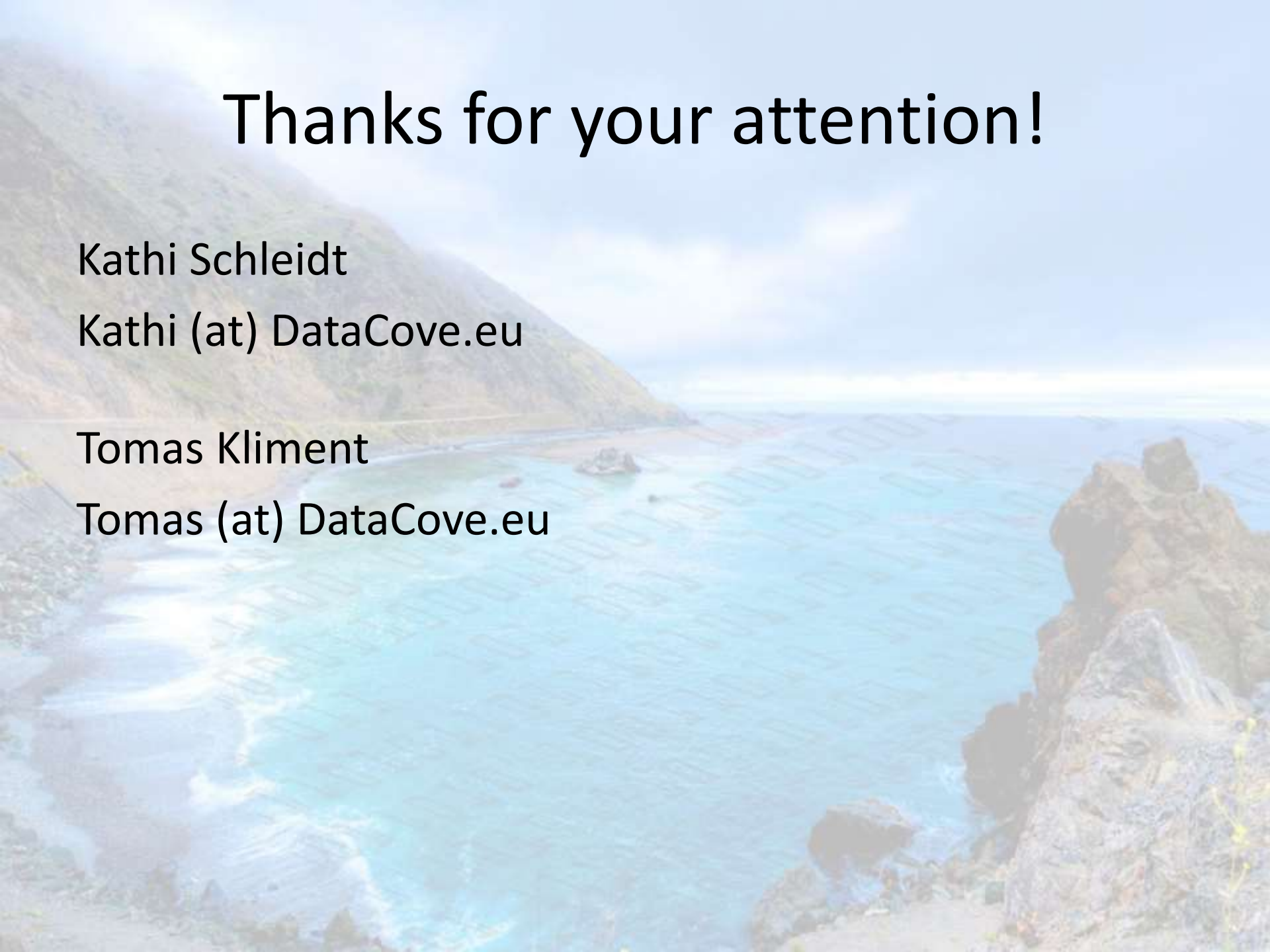
Thanks for your attention!

Kathi Schleidt

Kathi (at) DataCove.eu

Tomas Kliment

Tomas (at) DataCove.eu



DataCove Ghosts

WFS:

<http://data.datacove.eu:8080/geoserver/ows?service=WFS>

Eurostat Dictionaries (example):

<http://www.schleidt.org/Codelists/EuroStat/sex.xml>

Some real National Services

WFS:

<http://luft.umweltbundesamt.at/inspire/wfs?service=WFS&version=2.0.0&request=GetCapabilities>

SOS

<http://luft.umweltbundesamt.at/inspire/sos?service=SOS&version=2.0.0&request=getCapabilities>