EUROPEAN FORUM FOR GEOSTATISTICS

eurostat EUROPEAN COMMISSION

INE Instituto Nacional de Estadística

Using Census 2011 geodata of Spain

European Forum of Geostatistics Conference

23-25 of October 2013, Sofia (Bulgaria)

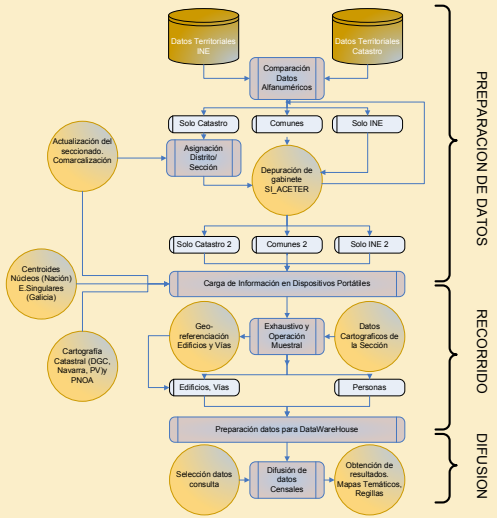
Ignacio Duque (IT Directorate)

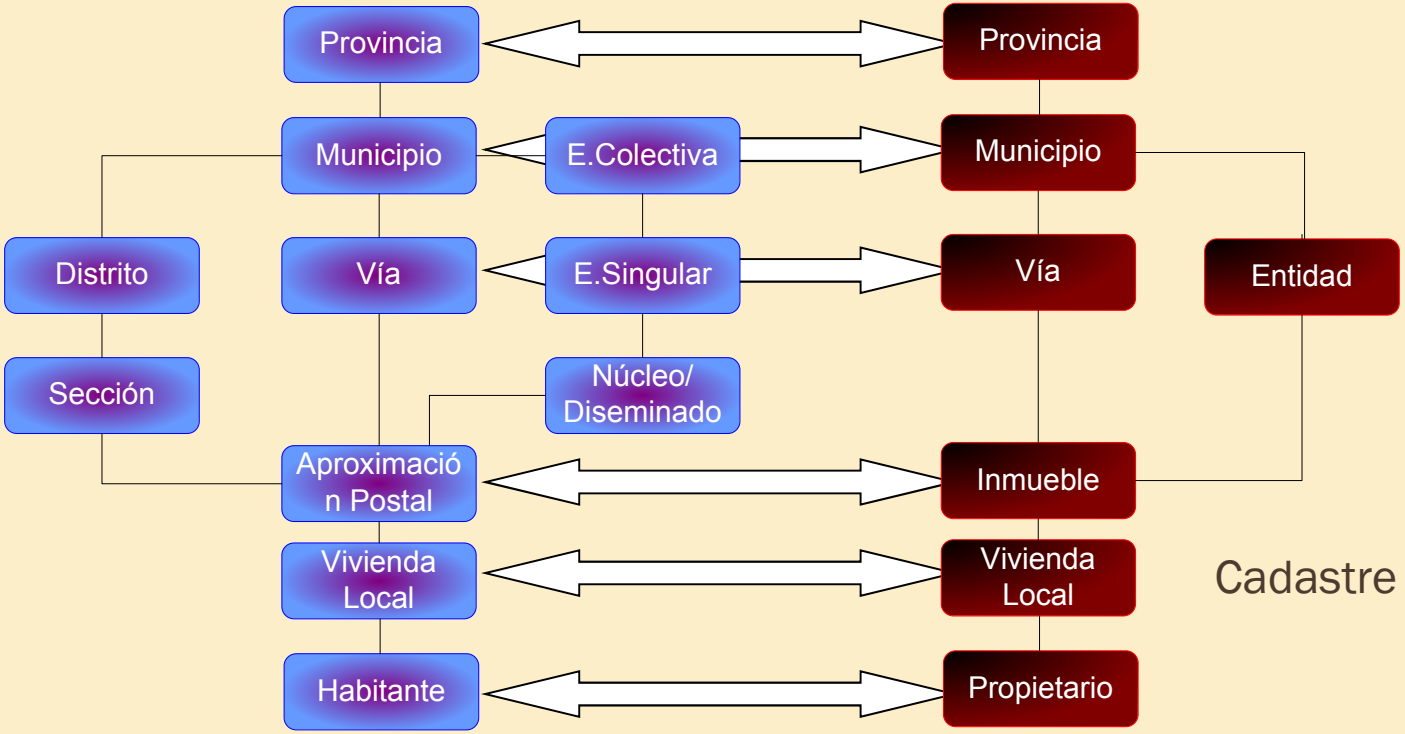
- Introduction.
- General overview of Census 2011 geographic approach in Spain.
- User's tailored tool for Census dissemination: general and geographical queries.
- Census address and spatial data infrastructure for statistics.

1. Census 2011 geographic approach (I)

Geospatial strategy of 2011 Census based on Registers.

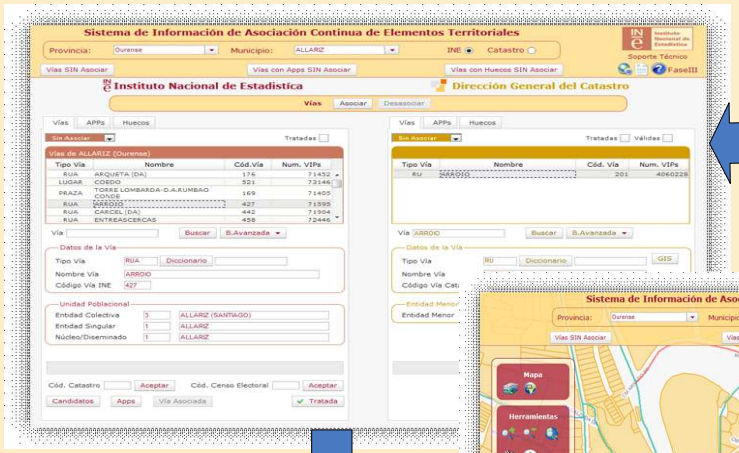
A) Defining the bridges between cadastral and statistical concepts





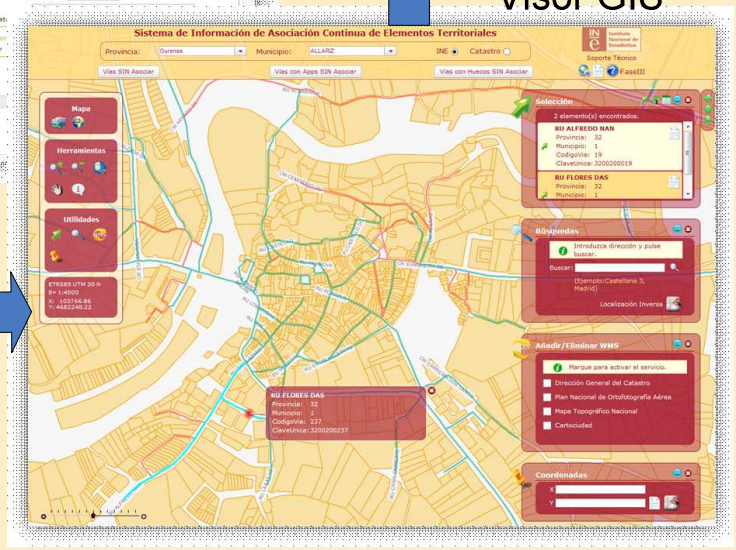
1. Census 2011 geographic approach (II)

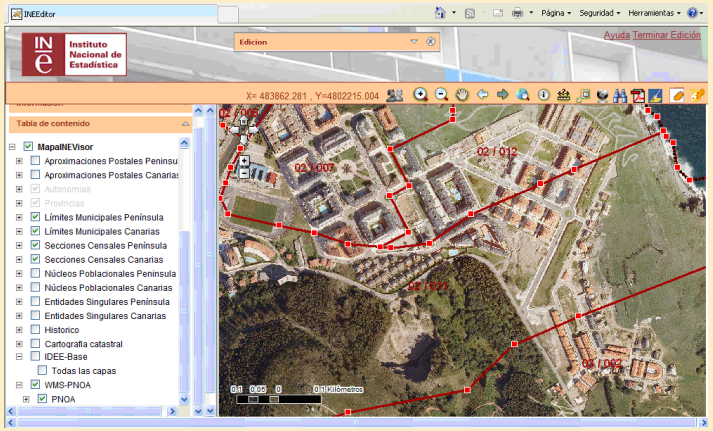
B) Matching streets and addresses:



Datos Alfanuméricos

Visor GIS





1. Census 2011 geographic approach (III)

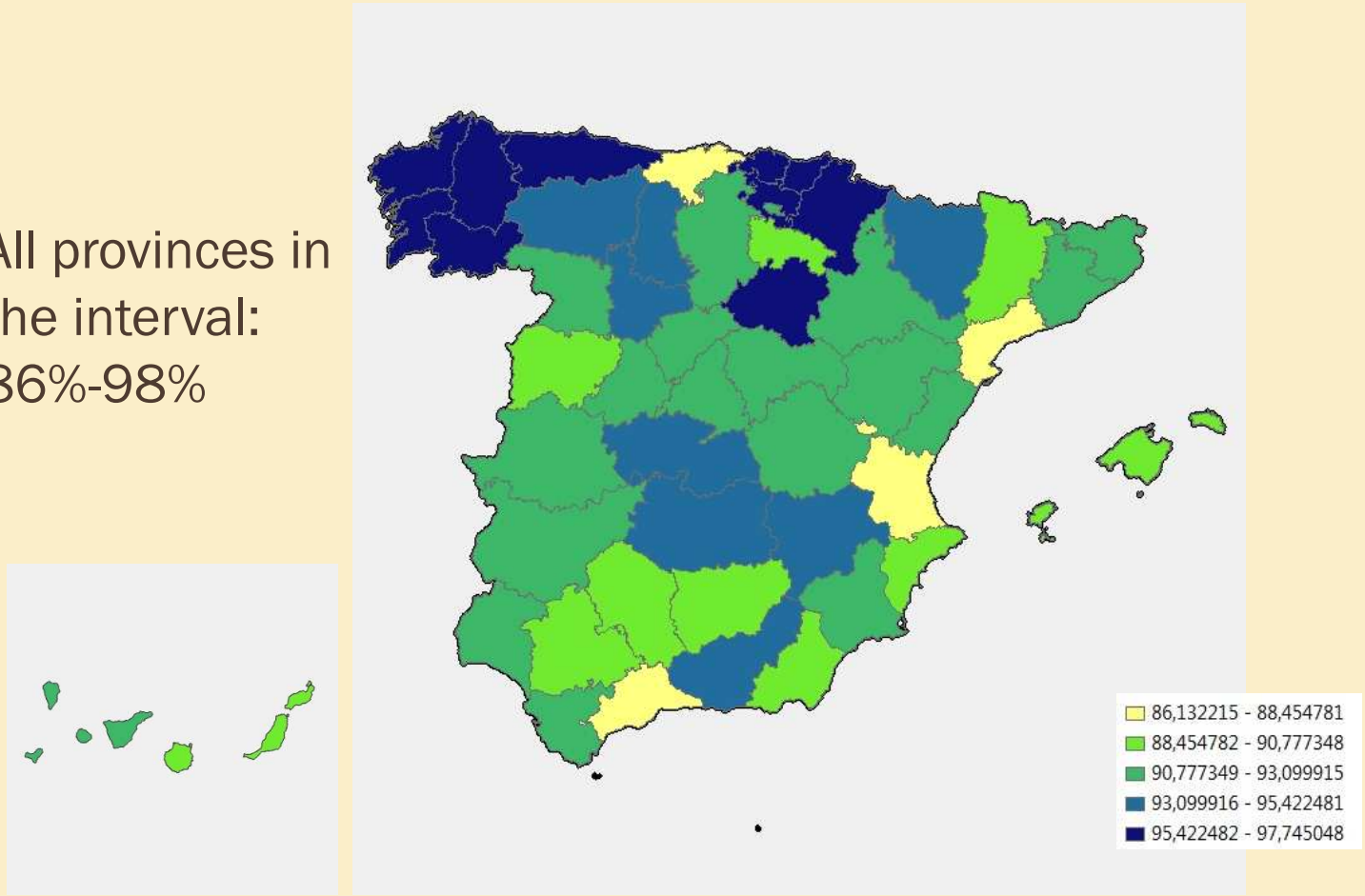
C) Using portable devices for checking or capturing coordinates:

1. Census 2011 geographic approach (III)

Final results, considering active addresses:

91% of Census addresses with coordinates

All provinces in the interval: 86%-98%



1. Census 2011 geographic approach (IV)

Methodological challenges about addresses with Census information and without x-y coordinates:

1. Use other located address with or without census data.
2. Use street axis and interpolation methods.
3. Use different polygon-unit of this address (mainly enumeration districts) and select a point inside.
4. Mixed methods

Key point: the rate of address without coordinates inside each municipality.

1. Census 2011 geographic approach (V)

Methodological challenges about Census based on registers and sample form.

1. All data are weighted.
2. First and elemental data for all population came from exhaustive data.
3. Long-form detailed census data came from sample data.
4. Complementary statistics from collective dwellings.

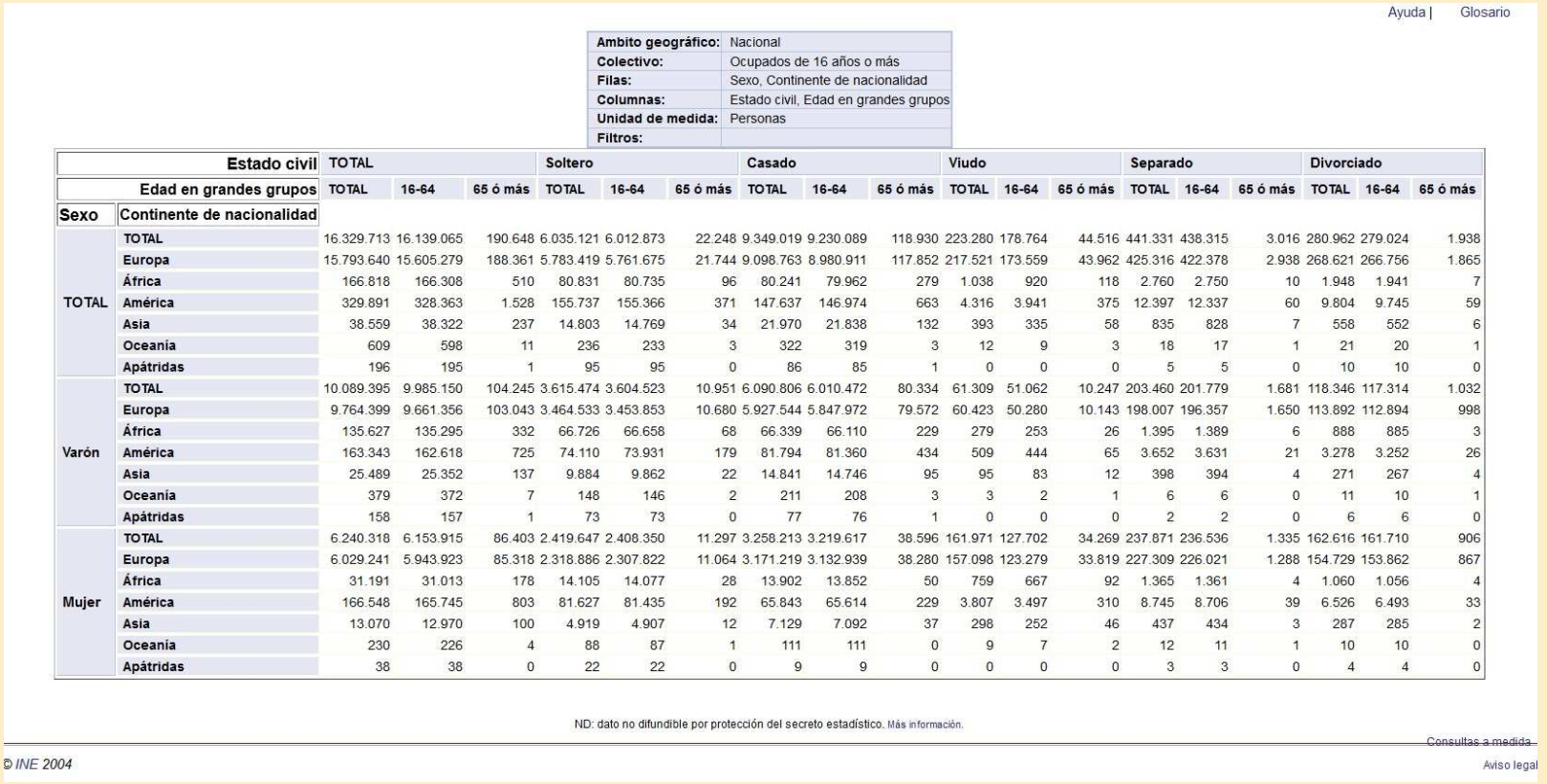
2. Census 2011 dissemination tool (I)

The previous experience:
2001 Census.

User selects:

- Thematic objective: persons, dwellings, households, family units.
- Geographical universe.
- Row variables.
- Column variables.

2. Census 2011 dissemination tool (II)



Ayuda | Glosario

Ambito geográfico:	Nacional
Colectivo:	Ocupados de 16 años o más
Filas:	Sexo, Continente de nacionalidad
Columnas:	Estado civil, Edad en grandes grupos
Unidad de medida:	Personas
Filtros:	

Estado civil		TOTAL			Soltero			Casado			Viudo			Separado			Divorciado		
Edad en grandes grupos		TOTAL	16-64	65 ó más	TOTAL	16-64	65 ó más	TOTAL	16-64	65 ó más	TOTAL	16-64	65 ó más	TOTAL	16-64	65 ó más	TOTAL	16-64	65 ó más
Sexo	Continente de nacionalidad																		
TOTAL	TOTAL	16.329.713	16.139.065	190.648	6.035.121	6.012.873	22.248	9.349.019	9.230.089	118.930	223.280	178.764	44.516	441.331	438.315	3.016	280.962	279.024	1.938
	Europa	15.793.640	15.605.279	188.361	5.783.419	5.761.675	21.744	9.098.763	8.980.911	117.852	217.521	173.559	43.962	425.316	422.378	2.938	268.621	266.756	1.865
	África	166.818	166.308	510	80.831	80.735	96	80.241	79.962	279	1.038	920	118	2.760	2.750	10	1.948	1.941	7
	América	329.891	328.363	1.528	155.737	155.366	371	147.637	146.974	663	4.316	3.941	375	12.397	12.337	60	9.804	9.745	59
	Asia	38.559	38.322	237	14.803	14.769	34	21.970	21.838	132	393	335	58	835	828	7	558	552	6
	Oceanía	609	598	11	236	233	3	322	319	3	12	9	3	18	17	1	21	20	1
	Apátridas	196	195	1	95	95	0	86	85	1	0	0	0	5	5	0	10	10	0
Varón	TOTAL	10.089.395	9.985.150	104.245	3.615.474	3.604.523	10.951	6.090.806	6.010.472	80.334	61.309	51.062	10.247	203.460	201.779	1.681	118.346	117.314	1.032
	Europa	9.764.399	9.661.356	103.043	3.464.533	3.453.853	10.680	5.927.544	5.847.972	79.572	60.423	50.280	10.143	198.007	196.357	1.650	113.892	112.894	998
	África	135.627	135.295	332	66.726	66.658	68	66.339	66.110	229	279	253	26	1.395	1.389	6	888	885	3
	América	163.343	162.618	725	74.110	73.931	179	81.794	81.360	434	509	444	65	3.652	3.631	21	3.278	3.252	26
	Asia	25.489	25.352	137	9.884	9.862	22	14.841	14.746	95	95	83	12	398	394	4	271	267	4
	Oceanía	379	372	7	148	146	2	211	208	3	3	2	1	6	6	0	11	10	1
	Apátridas	158	157	1	73	73	0	77	76	1	0	0	0	2	2	0	6	6	0
Mujer	TOTAL	6.240.318	6.153.915	86.403	2.419.647	2.408.350	11.297	3.258.213	3.219.617	38.596	161.971	127.702	34.269	237.871	236.536	1.335	162.616	161.710	906
	Europa	6.029.241	5.943.923	85.318	2.318.886	2.307.822	11.064	3.171.219	3.132.939	38.280	157.098	123.279	33.819	227.309	226.021	1.288	154.729	153.862	867
	África	31.191	31.013	178	14.105	14.077	28	13.902	13.852	50	759	667	92	1.365	1.361	4	1.060	1.056	4
	América	166.548	165.745	803	81.627	81.435	192	65.843	65.614	229	3.807	3.497	310	8.745	8.706	39	6.526	6.493	33
	Asia	13.070	12.970	100	4.919	4.907	12	7.129	7.092	37	298	252	46	437	434	3	287	285	2
	Oceanía	230	226	4	88	87	1	111	111	0	9	7	2	12	11	1	10	10	0
	Apátridas	38	38	0	22	22	0	9	9	0	0	0	0	3	3	0	4	4	0

ND: dato no difundible por protección del secreto estadístico. Más información.

Consultas a medida

© INE 2004

Aviso legal

Background architecture:

- Data warehouse tools.
- SAS procedures.
- Confidentiality filter.
- Intensive previous work for generating OLAP cubes

2. Census 2011 dissemination tool (III)

General overview of DW project:

Available data-bases:

- Buildings.
- Family dwellings.
- Households.
- People.
- Family groups and couples.

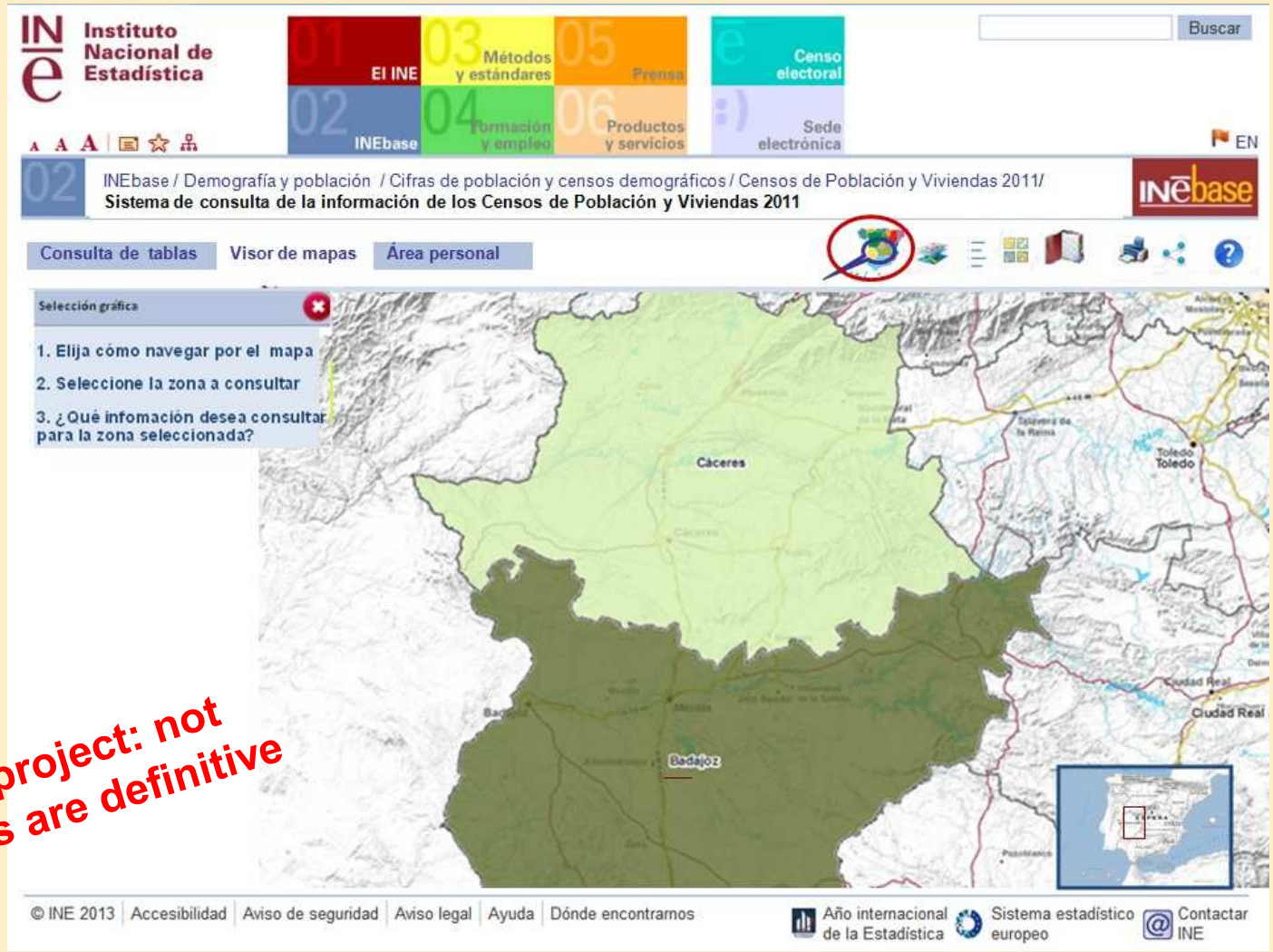
All queries are from sample files (except Buildings)

Query types:

- Classical administrative units.
- Geographical.
- Pre-defined thematic maps.

2. Census 2011 dissemination tool (IV)

Web portal home draft:



Running project: not all details are definitive

2. Census 2011 dissemination tool (V)

Treatment for the few addresses without coordinates:

- Using coordinates of the same numbers, with different street access.
- Using coordinates the same street and closets numbers.
- Using inner centroids of the enumeration district.

2. Census 2011 dissemination tool (VI)

Key issue for confidentiality and geographical queries:

how to avoid small number and the *subtracting behavior*?

Technical strategy:

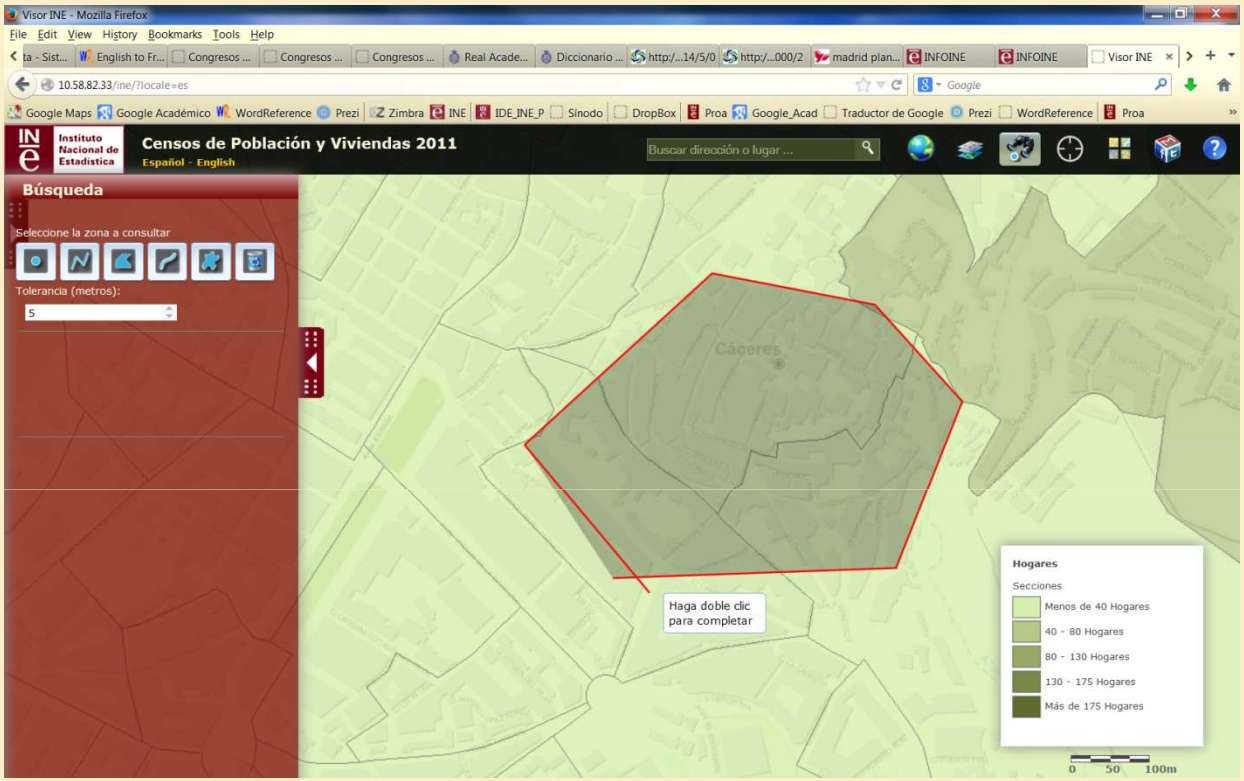
1. Small geographical units
2. Sample units / tabulated cells.
3. Tables and Variables.
4. Final results.

Confidentiality and sample size issues are treated together

1. Small geographical unit.

- Queries are not based in direct address information.
- All queries are based in aggregated units: grid cell or an addition of grid cells.
- Based on Eurostat-EFGS grid of 1km.
- Cells are divided and added for attained en equilibrium:
  - At least 20 sample units (about 220 people).
  - Producing small cells in highly populated areas.

2. Census 2011 dissemination tool (VIII)



The user design a free polygon, but in fact the query look for all the cells partially contained in the defined polygon.

2. The relationship between sample units and tabulated cells.

Sample units = M

- Family dwellings.
- Households.
- People.
- Family groups and couples.

Size limit

$(M / D) > x$

Cells of the table requested = D

3. Variables and table control.

Sample units	Table query	Maximum variables
Low	Only pre-defined	
Intermediate	Pre-defined or free.	4
High	Free	6

→ *Special treatment for sensitive variables*

4. Treatment of the final results.

All the final results are rounded at 0 or 5

DW architecture and works:

- Data warehouse implementation details.
  - Background OLAP cubes.
  - Predefined most frequent queries.
- SAS procedures.
- Previous process for feeding the Data warehouse.

2. Census 2011 dissemination tool (XII)

Thematic maps for a selected Census indicators:

- Total population.
- Age structure.
- Foreign people.
- Educational attainment.
- Employment status.
- Dwelling data: age, surface, etc.
- Household size.
- Family types.

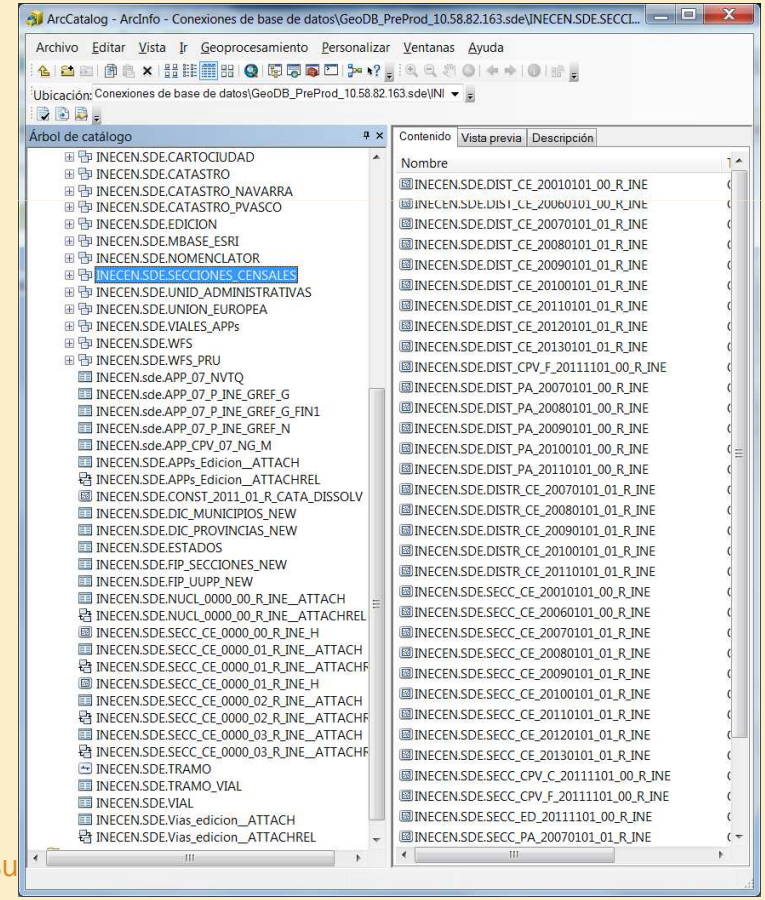
3. Census address and SDI for statistics (I)

Aims of Spatial Data Infrastructure general definition:

A spatial data infrastructure (SDI) is a data infrastructure implementing a framework of ***geographic data, metadata, users and tools*** that are interactively connected in order to use spatial data in an efficient and flexible way.

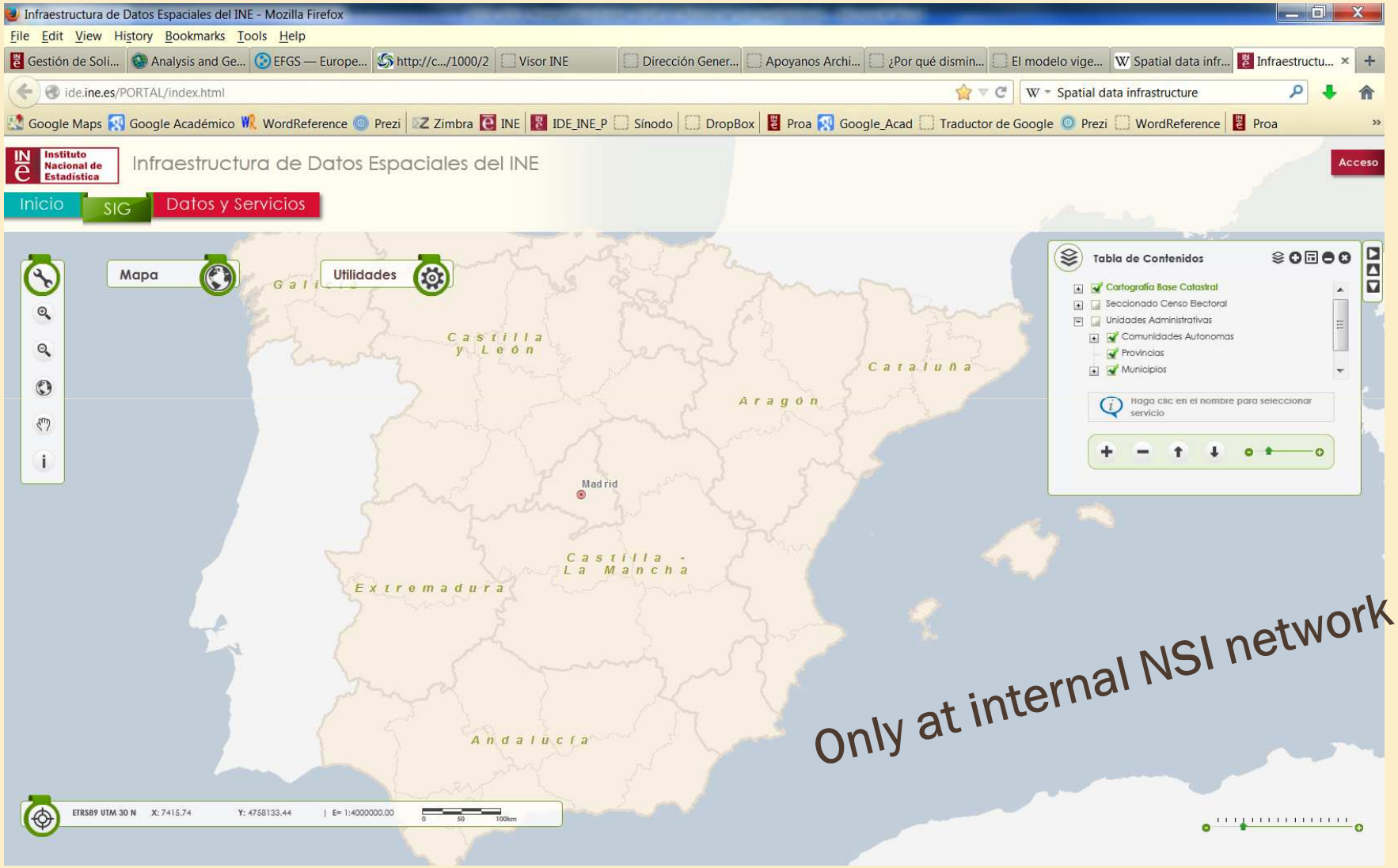
Specific objectives for Statistical Production:

- ***Unique geodatabase with the geometry of:***
  - Statistical units produced by INE.
  - Postal references produced by NSI and others.
  - Cartographic features and images produced by Map and Cadastral Agencies.
- ***Interactive services of statistical units.***

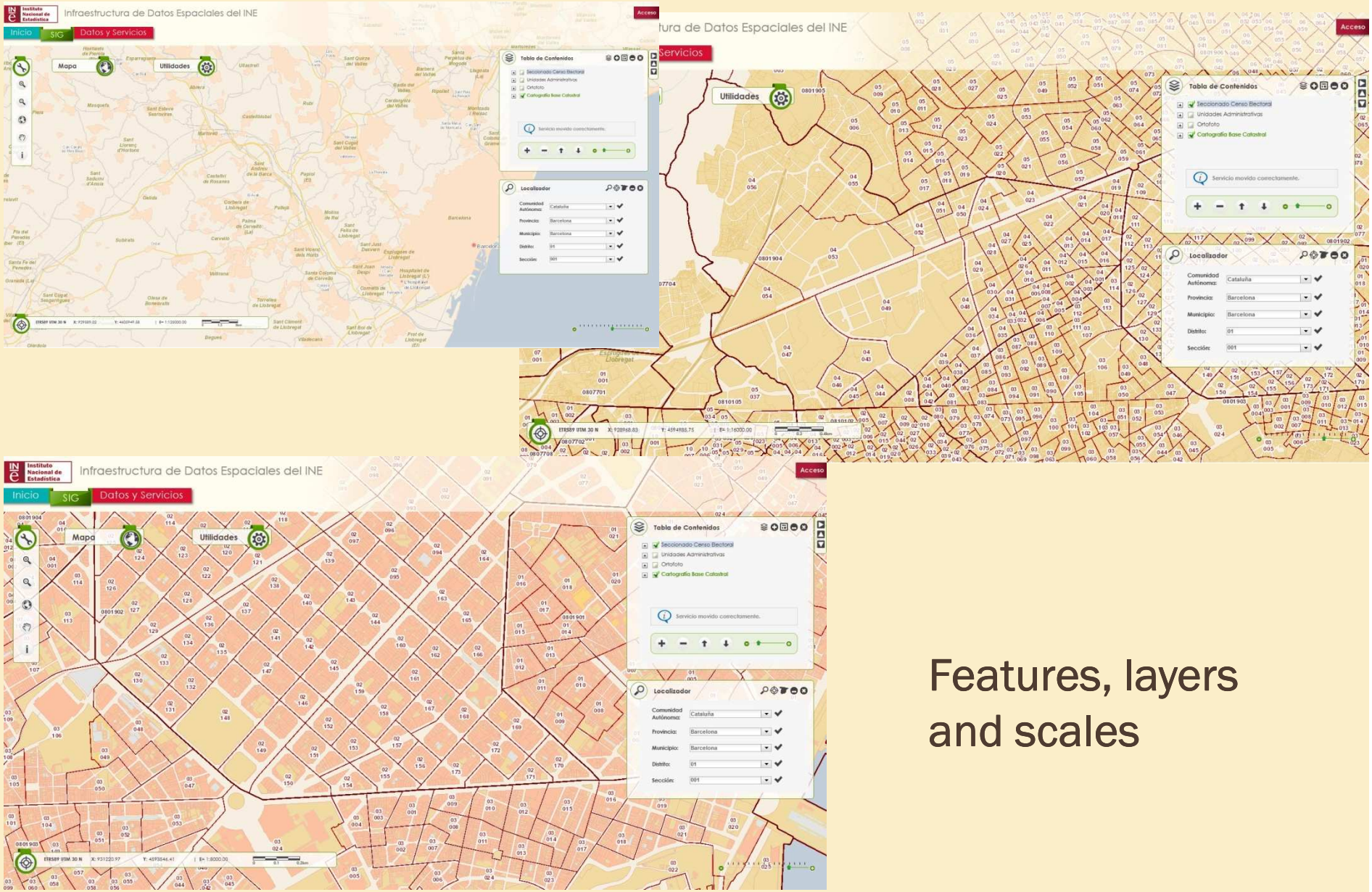


3. Census address and SDI for statistics (II)

The SDI viewer

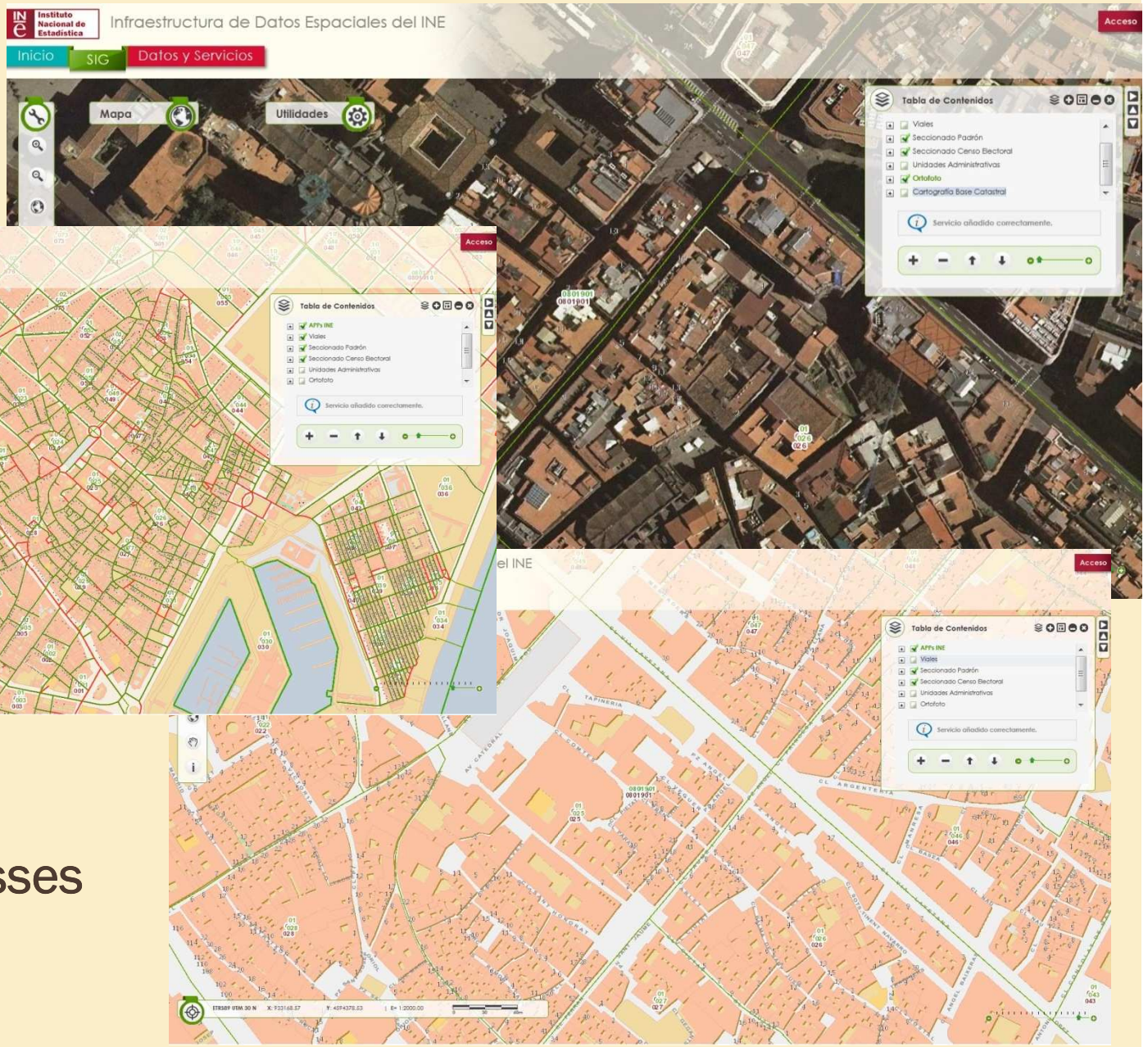


3. Census address and SDI for statistics (III)



Features, layers and scales

3. Census address and SDI for statistics (IV)

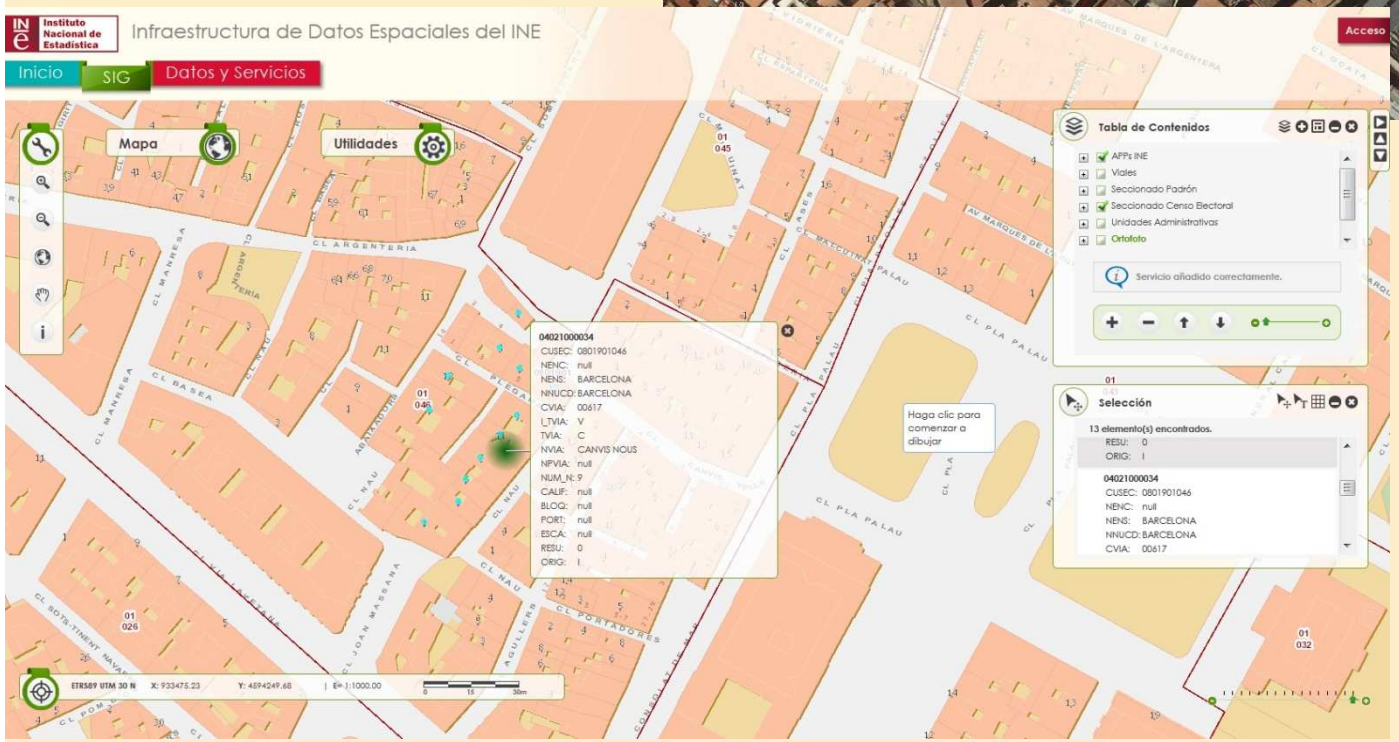


Census addresses

3. Census address and SDI for statistics (V)

Cadastral parcels

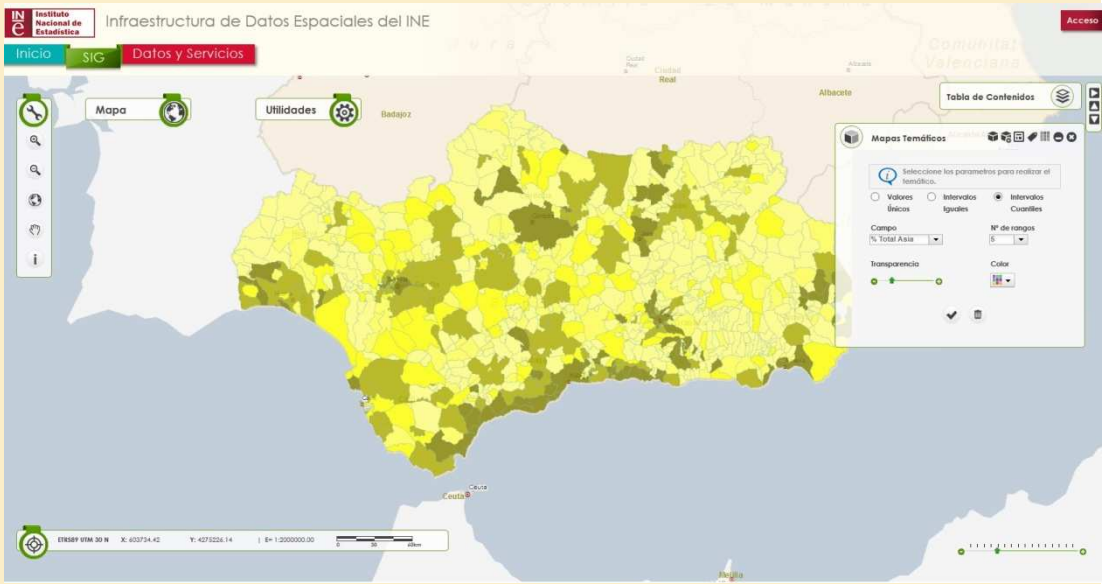
Look and Find statistical, administrative and postal information

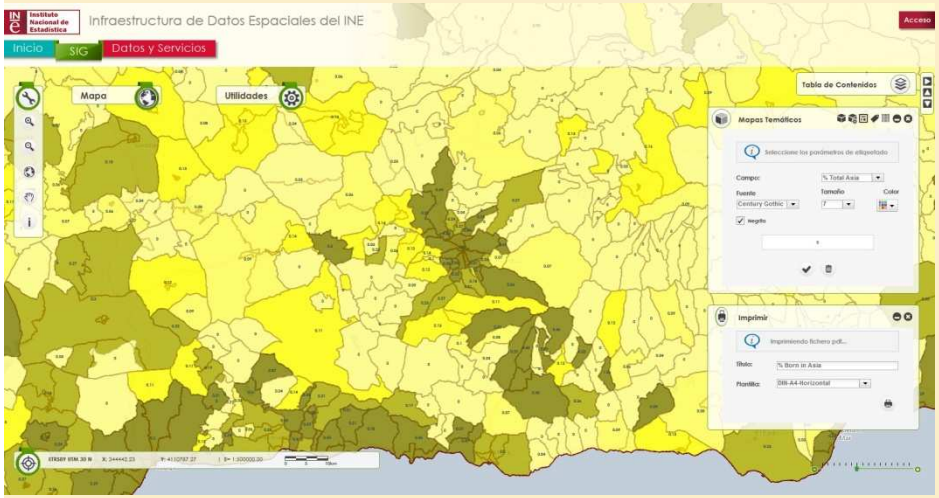


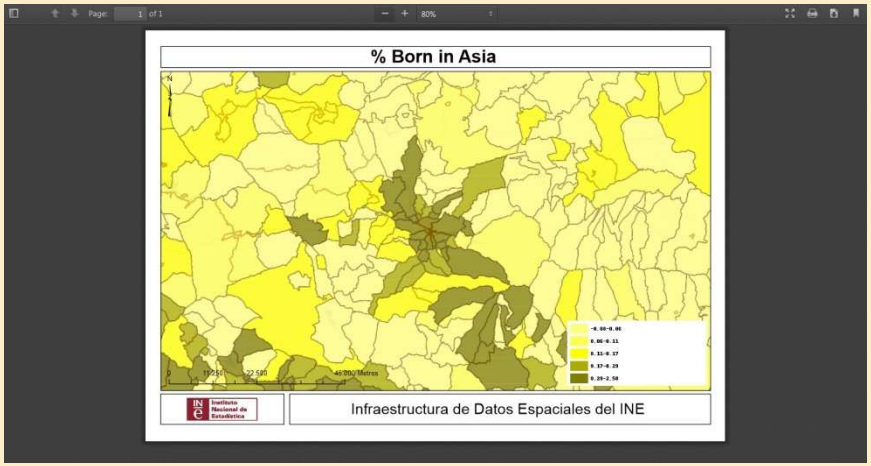
Buildings addresses

3. Census address and SDI for statistics (VI)

Thematic maps

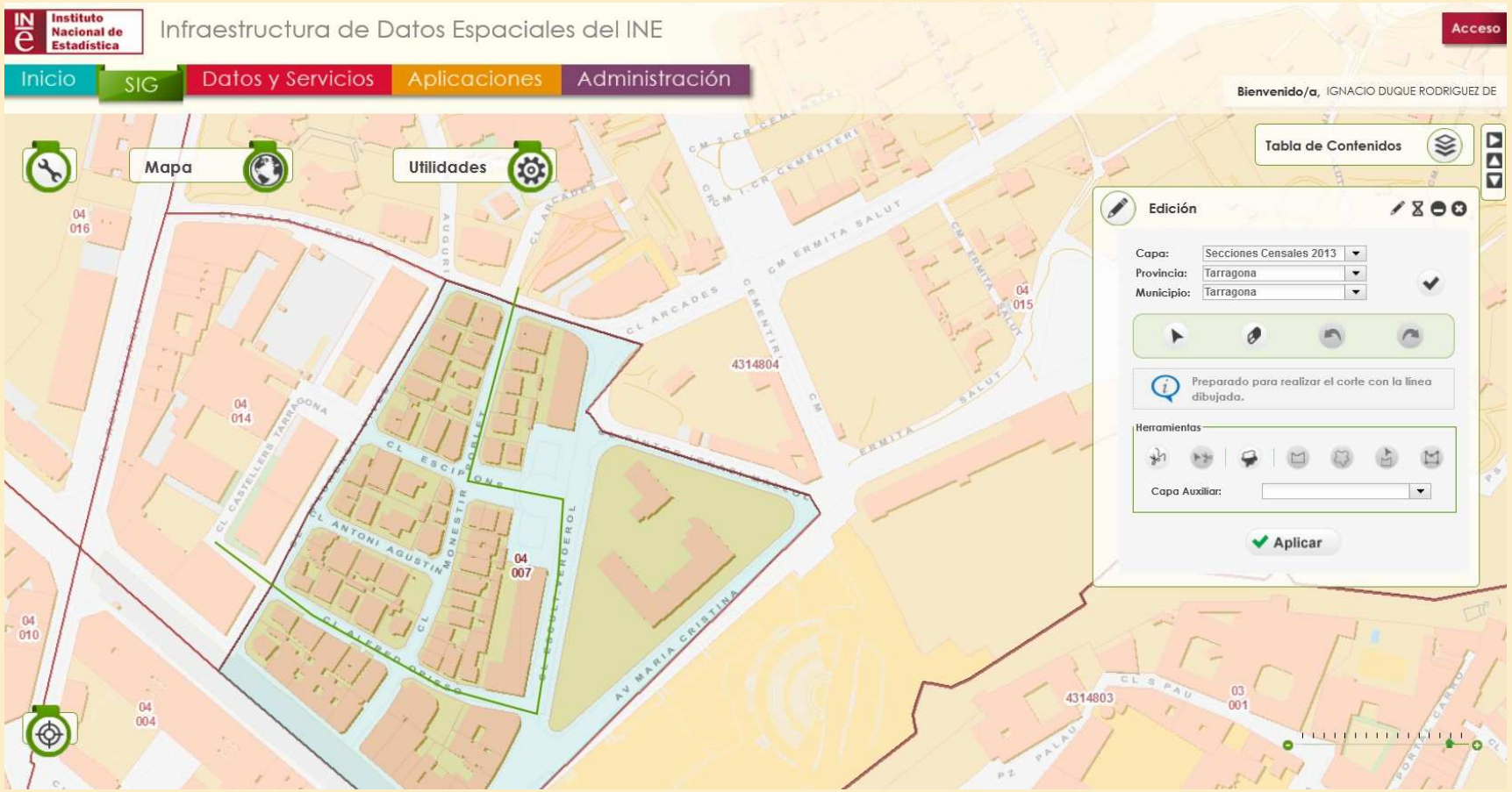






3. Census address and SDI for statistics (VII)

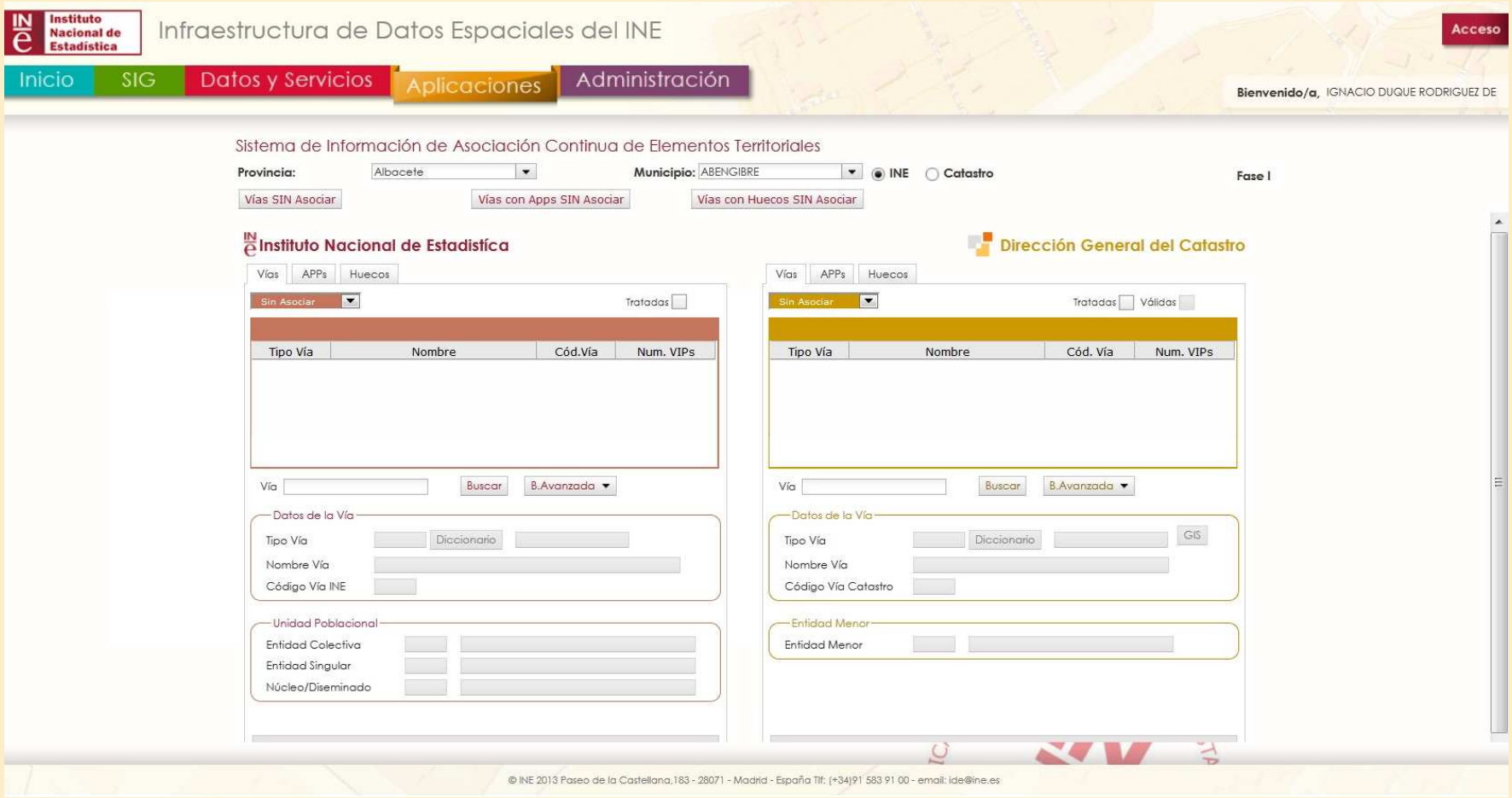
Web edition of statistical units updating process



From 52 offices, videoconference training meetings

3. Census address and SDI for statistics (VIII)

Other web applications developed … or planed



ignacio.duque.rodriguez@ine.es