

Input for the dissemination of the Demographic Yearbook

Andrea Fernández Conde
andrea.fernandez@inegi.org.mx

November 11, 2020

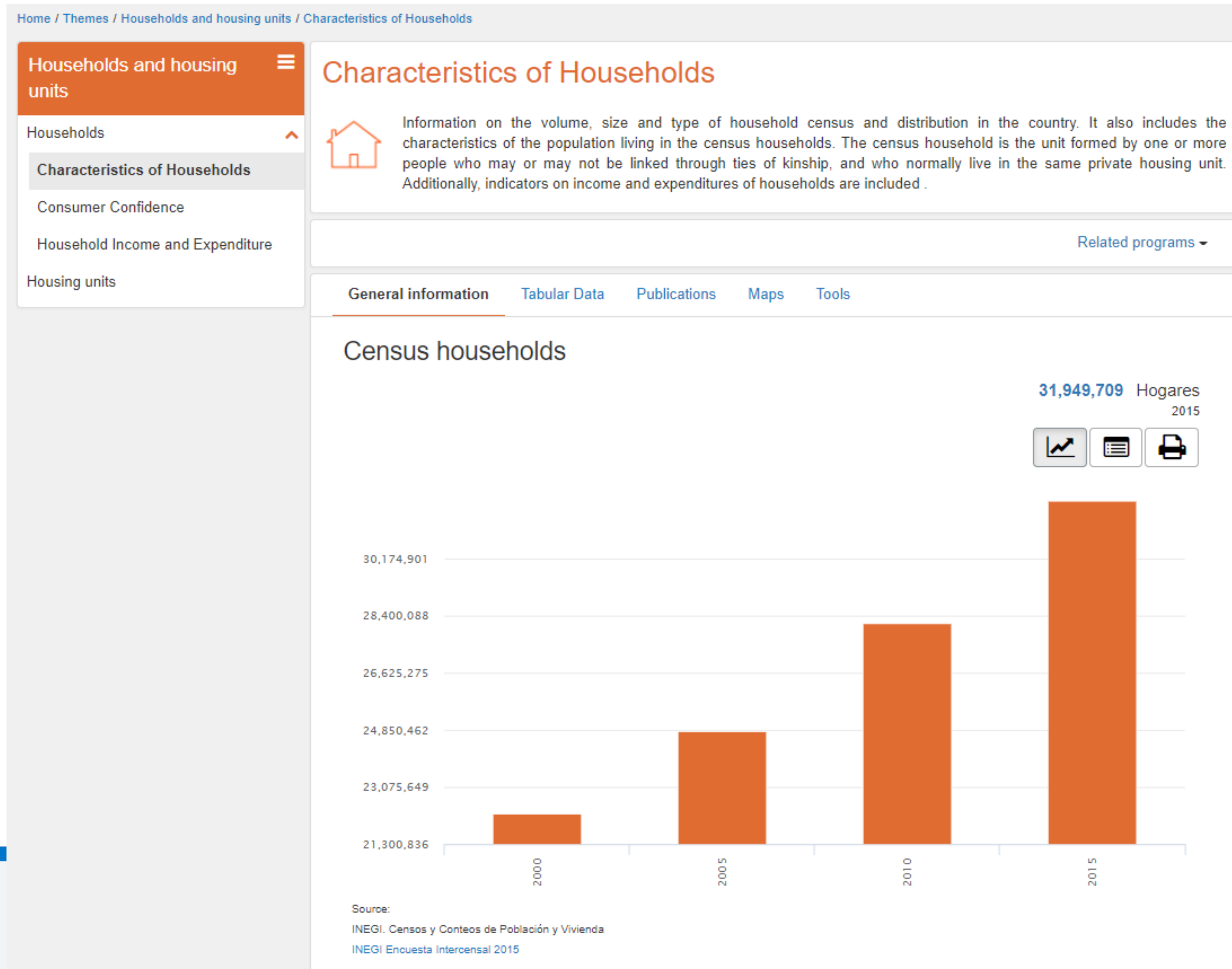
What tools in disseminating demographic data and metadata does your NSO use?

Dissemination

- Our vision is that metadata and data go together.
- We are trying more and more to disseminate data always alongside of its metadata.
- Our main tools for dissemination are:
 - By topic
 - By statistical programme
 - Specialized tools:
 - Indicators bank: search of indicators by topic with functionalities for visualizing and downloading data in a variety of formats.
 - Mapping tool: functionality to create cartograms with enriched layers and the final data downloadable to the user.

Topic search

- Each topic contains a selected series of indicators in a graph or tabular format.
- Reference to the source is always available as a footnote to a specific graph.
- The user has access to the data in tabular format within the graph (default) and as a series of tables.



Topic general structure

- We provide “Interactive tabular data” of the main indicators shown in each topic.
- Referrals to basic metadata is included.

Characteristics of Households



Information on the volume, size and type of household census and distribution in the country. It also includes the characteristics of the population living in the census households. The census household is the unit formed by one or more people who may or may not be linked through ties of kinship, and who normally live in the same private housing unit. Additionally, indicators on income and expenditures of households are included.

[Related programs](#) ▾

[General information](#)

[Tabular Data](#)

[Publications](#)

[Maps](#)

[Tools](#)

Search in title












Results found: **9**

– Title

Period

– Interactive tabular data

 Households by state according to head of household	2000 - 2015
 Population in households by state according to head of household	2000 - 2015
 Average size of households by state according to head of household	2000 - 2015
 Households by state according to type	2014 - 2017
 Nuclear family households by state according to head	2014 - 2017
 Extended family households by state according to head	2014 - 2017
 Single-person non-family households by state according to head	2014 - 2017
 Single-parent households by state according to the head	2014 - 2017
 Two-parent households by state according to head	2014 - 2017

Interactive tabular data

- Interactive tabular data is useful because users can select the desired option and generate a file in a downloadable open format for their use.
- The tool used to generate this interactive tabular data is PC-Axis (a tool made by other NSOs).

Hogares por entidad federativa según jefatura del hogar, 2000 a 2015

Selecciona las variables Ver el tabulado

Entidad federativa ↓ ↑

Buscar...

Estados Unidos Mexicanos
Aguascalientes
Baja California
Baja California Sur
Campeche
Coahuila de Zaragoza
Colima

33 de 33 seleccionados

Periodo ↓ ↑

Buscar...

2000
2005
2010
2015

4 de 4 seleccionados

Seleccionar

Jefatura del hogar ↓ ↑

Buscar...

Total
Jefatura masculina
Jefatura femenina

3 de 3 seleccionados

Interactive tabular data

- Basic metadata is included in the footnote.

Hogares por entidad federativa según jefatura del hogar, 2000 a 2015

[Selecciona las variables](#)

[Ver el tabulado](#)



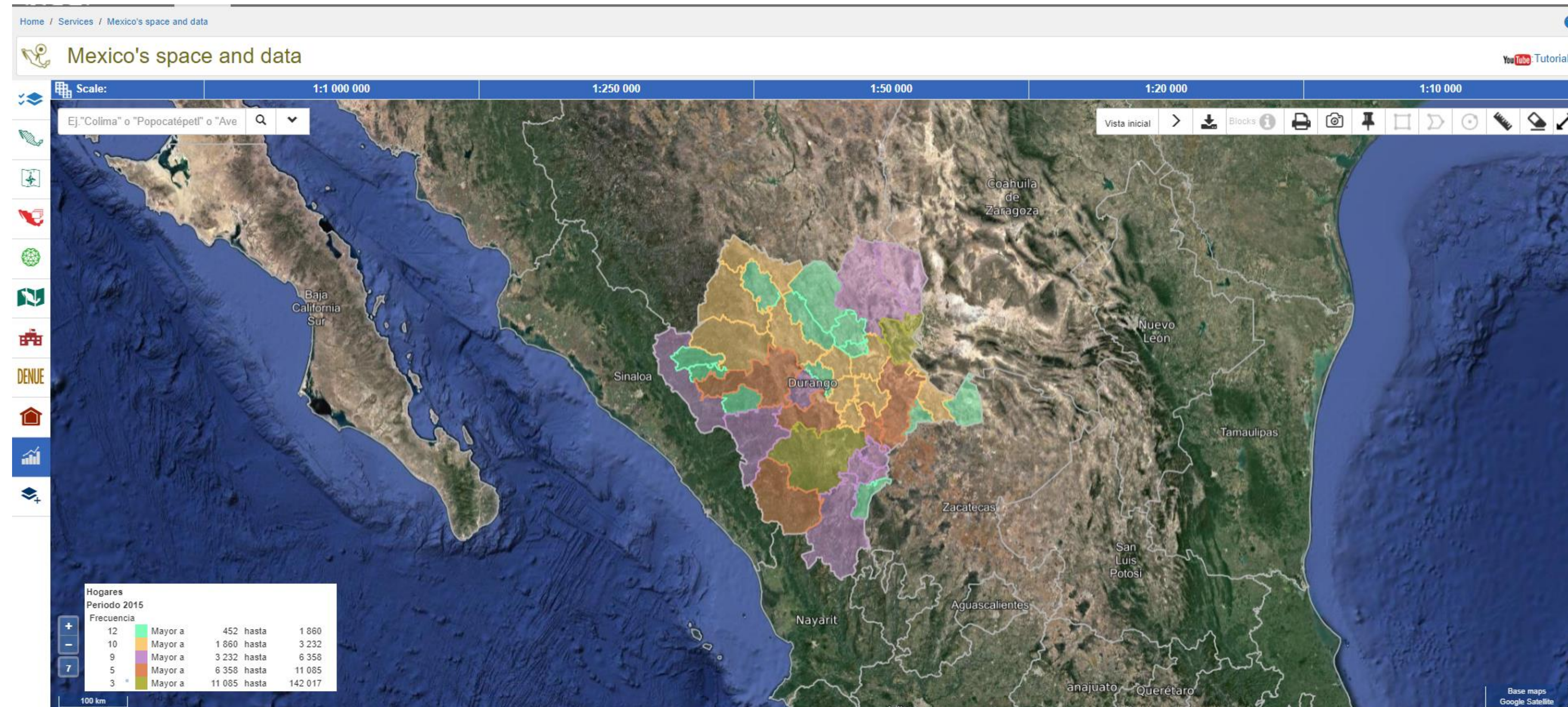
	2000			2005			2010			2015		
	Total	Jefatura masculina	Jefatura femenina	Total	Jefatura masculina	Jefatura femenina	Total	Jefatura masculina	Jefatura femenina	Total	Jefatura masculina	Jefatura femenina
Estados Unidos Mexicanos	22 268 916	17 671 681	4 597 235	24 803 625	19 085 966	5 717 659	28 159 373	21 243 167	6 916 206	31 949 709	22 683 498	9 266 211
Aguascalientes	208 167	169 445	38 722	248 905	196 553	52 352	289 575	224 643	64 932	334 589	247 011	87 578
Baja California	568 090	446 585	121 505	697 250	528 337	168 913	858 676	635 594	223 082	967 863	666 287	301 576
Baja California Sur	107 009	87 717	19 292	132 233	105 309	26 924	175 046	134 525	40 521	209 834	151 763	58 071
Campeche	160 492	131 947	28 545	186 134	147 761	38 373	211 632	162 574	49 058	244 471	172 986	71 485
Coahuila de Zaragoza	552 024	457 712	94 312	627 101	508 345	118 756	715 158	565 393	149 765	809 275	614 713	194 562
Colima	128 295	100 179	28 116	147 092	110 477	36 615	177 848	131 791	46 057	205 243	145 182	60 061
Chiapas	808 149	674 349	133 800	924 967	749 336	175 631	1 072 560	856 153	216 407	1 239 007	938 446	300 561
Chihuahua	744 159	590 665	153 494	822 586	632 614	189 972	910 647	688 119	222 528	1 033 658	724 088	309 570
Ciudad de México	2 180 243	1 618 625	561 618	2 292 069	1 630 794	661 275	2 388 534	1 638 790	749 744	2 601 323	1 672 203	929 120
Durango	329 552	262 277	67 275	360 308	278 890	81 418	398 471	301 613	96 858	455 989	324 994	130 995
Guanajuato	990 119	791 437	198 682	1 105 564	851 435	254 129	1 266 772	964 206	302 566	1 443 035	1 039 266	403 769
Guerrero	674 177	508 497	165 680	713 788	526 136	187 652	805 230	588 351	216 879	895 157	602 071	293 086
Hidalgo	503 151	398 384	104 767	562 857	434 755	128 102	662 651	504 119	158 532	757 252	538 386	218 866
Jalisco	1 441 069	1 141 537	299 532	1 598 029	1 219 866	378 163	1 802 424	1 359 424	443 000	2 059 987	1 480 280	579 707
México	2 848 992	2 319 180	529 812	3 221 617	2 545 534	676 083	3 689 053	2 841 143	847 910	4 168 206	3 009 938	1 158 268
Michoacán de	887 958	697 630	190 328	937 373	714 661	222 712	1 066 630	809 569	257 061	1 191 884	859 451	332 433

Notas:

Para 2000, la información está referida al 14 de febrero.
 Para 2005, la información está referida al 17 de octubre.
 Para 2010, la información corresponde a hogares censales y está referida al 12 de junio.
 Para 2015, la información corresponde a hogares censales y está referida al 15 de marzo.

Mapping

- Maps are also available (wherever possible) for the same indicators shown by topic.
- The user can select additional layers and can download a shapefile for all layers selected.



Tools

- The user is linked by topic to general purpose applications that provide the user with more detailed functionality and Access to data.
- One such tool is our “indicators bank”.

Characteristics of Households



Information on the volume, size and type of household census and distribution in the country. It also includes the characteristics of the population living in the census households. The census household is the unit formed by one or more people who may or may not be linked through ties of kinship, and who normally live in the same private housing unit. Additionally, indicators on income and expenditures of households are included .

[Related programs](#) ▼

[General information](#)

[Tabular Data](#)

[Publications](#)

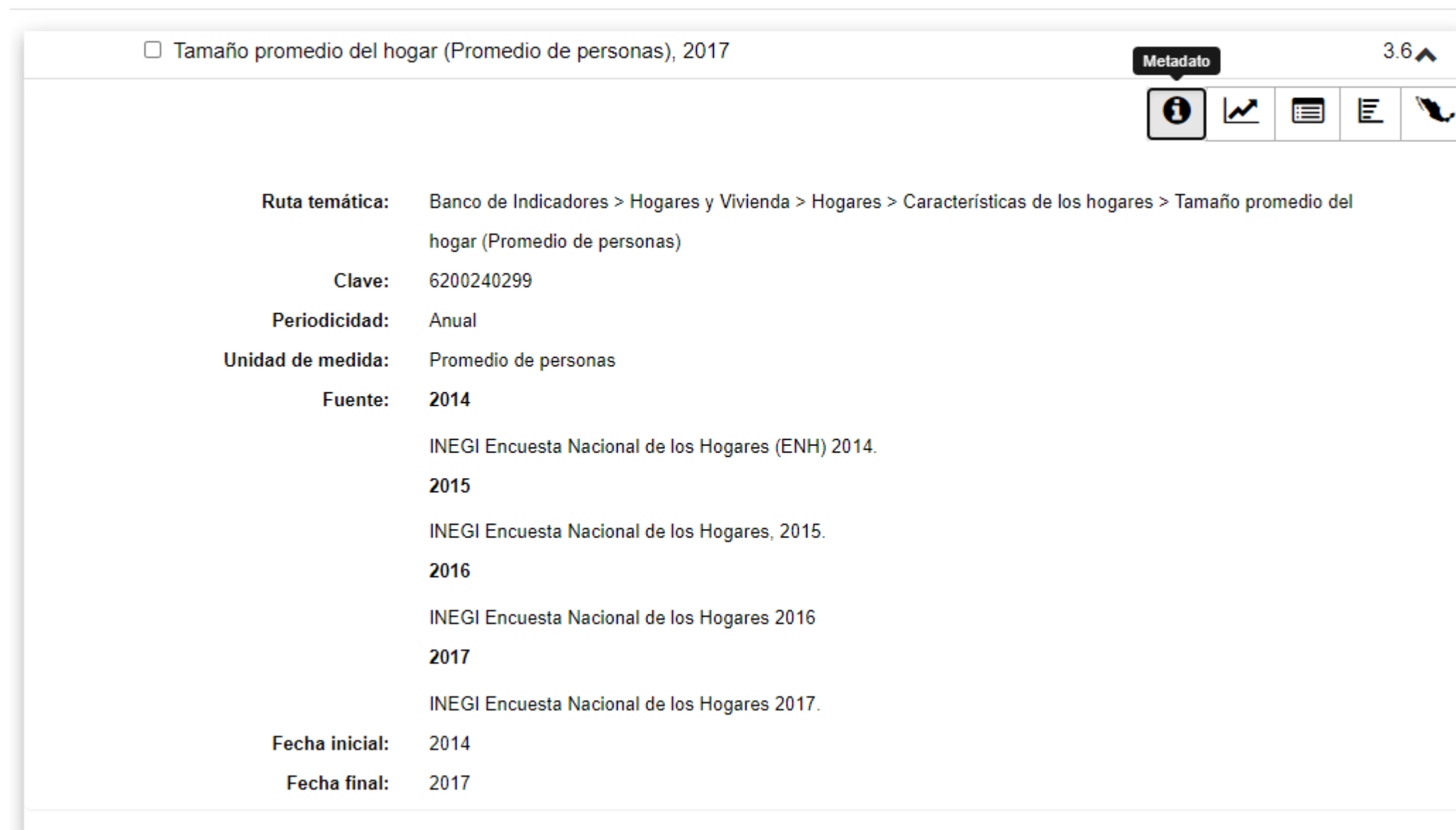
[Maps](#)

[Tools](#)

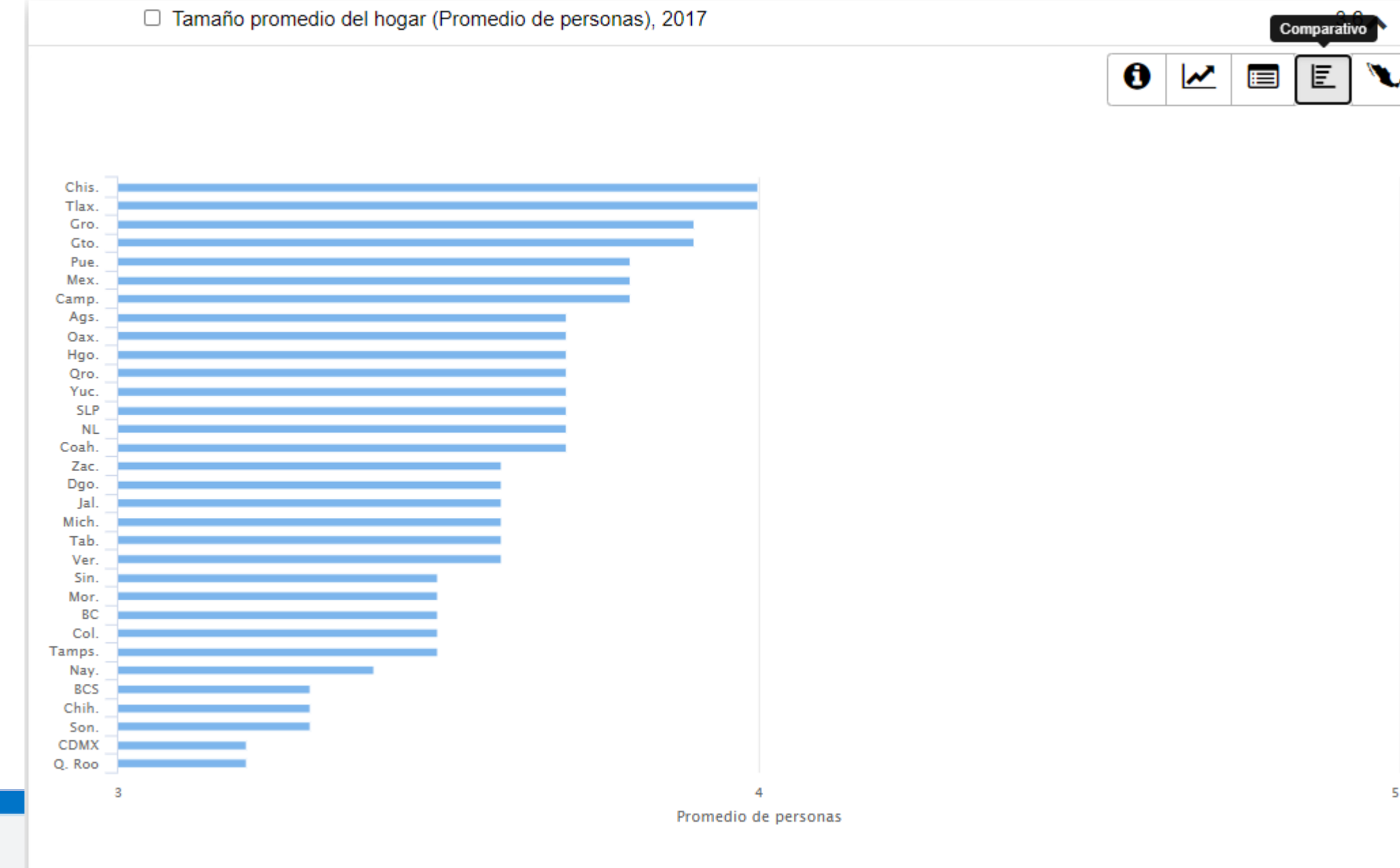
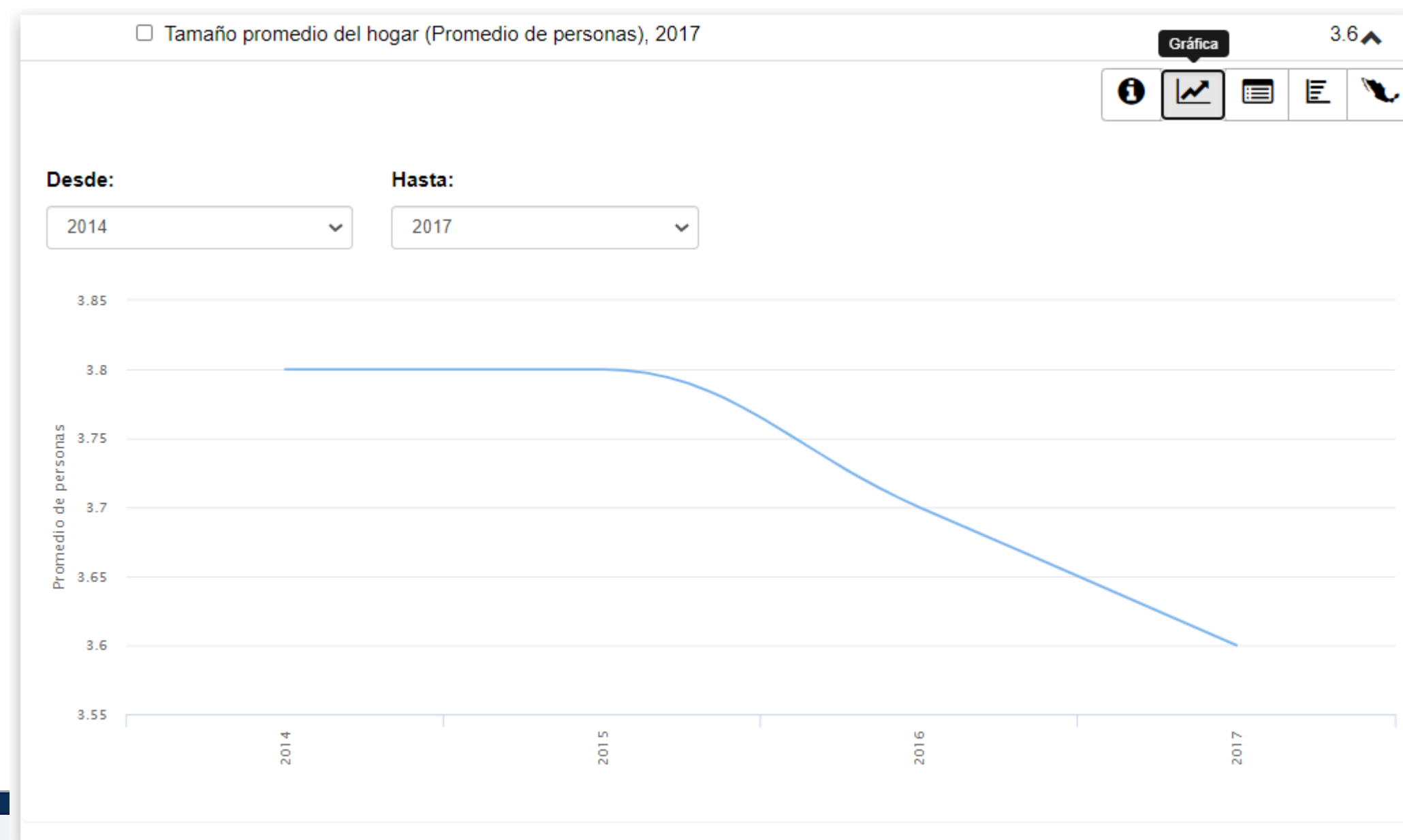
Services

- › Online services
 - Indicators bank
 - Dynamic graphs
- › Mobile Apps
 - Mexico in figures
- › For developers
 - APIs
 - Widgets

Características de los hogares



Características de los hogares



Access by statistical programme

Home / Programs / 2010 Census of Population and Housing Units

Home / Programs / 2010 Census of Population and Housing Units

Population and Housing Unit

Census and intercensal historical series (1990-2010)

2020

2010

2005 (Count)

2000

1995 (Count)

1990

1980

1970

1960


1950

1940

1930

1921

2010 Census of Population and Housing Units

 From May 31st to June 25th, the 2010 Population and Housing Census took place. More than 106,000 interviewers went over the country with the purpose of visiting each home and ask some questions about them and their occupants.

Note: Derived from the judgment of the Plenum of the Supreme Court of Justice of the Nation regarding Constitutional Controversy 41/2011, published on June 26th, 2013. All results of the 2010 Population and Housing Census for the Tultepec, Nextlalpan and Tultitlan municipalities, State of Mexico.

[Read more](#)

[Related topics](#)

Documentation | [Tabular data](#) | [Microdata](#) | [Open data](#) | [Publications](#) | [Tools](#)

- Results presentation
- Methodology
- Questionnaires
- Glossary
- Geostatistical framework
- Classifieds
- Metadata (DDI standard)
- Support material for dissemination

Tabular data

- Users can also create their own tables for the main indicators of each statistical programme.
- Predefined (by the producer) tabular data is also available.

Filters

Groups: Geographical area: [Consult](#)

Search in title [Q](#) Results found: **227** [Download all files](#)

Title	Period	Formats
- Interactive tabular data		
Total population	2010	
Population aged 3 years and over	2010	
Population aged 12 years and over	2010	
Female population aged 12 years and over	2010	
Population in households and its housing units	2010	
Census households	2010	
Housing units	2010	
- Predefined tabular data		
Population pyramid	2010	PDF ↓ 25.0 KB
- Basic questionnaire		
- Population		
Localities and population by state depending on the size of the town	2010	PDF ↓ 357 KB XLS ↓ 37.0 KB

Microdata

- Whenever confidentiality allows, we provide an anonymized version of the data in several formats.
- If confidentiality applies, we provide sample tables with the data structure so that users can create their scripts and either:
 - Run them in our data enclave
 - Have us run their scripts to get the results.
- A detailed description of the structural metadata is necessary so that the user understands what the data has.

Download Microdata Lab / Remote processing Processing service

Filters
Groups:
All ▼ Consult

Search in title 🔍 Results found: 185 Download all files

Title	Period	Formats
- Census (basic questionnaire)		
- Databases (examples)		
Households	2010	DBF ↓ 3.77 MB SAS ↓ 3.81 MB DTA ↓ 3.76 MB SAV ↓ 10.1 MB
People	2010	DBF ↓ 24.9 MB SAS ↓ 25.6 MB DTA ↓ 25.3 MB SAV ↓ 43.4 MB
+ Documentation of the database		
- Sample (extended questionnaire)		
+ Databases		
+ Documentation of the database		
- Results on infrastructure and characteristics of the urban environment		
+ Databases		
+ Documentation of the database		
- Results on localities with less than 5,000 inhabitants		
+ Databases		

Microdata

- We have defined a standard for the documentation of the database.
- The data dictionary is its main component.
- The data producer has to provide an Excel for which each sheet is a table. Each variable in a table is defined by adding: description, mnemonic, question, valid values.

- Documentation of the database			
Descripción de la base de datos	2010		XLS ↓ 80.5 KB
Classification	2010		
General Integration of Localities (CIGEL)	2010		DBF ↓ 3.61 MB

CENSO DE POBLACIÓN Y VIVIENDA 2010					
DICCIONARIO EXPLOTACIÓN - CARACTERÍSTICAS DE LAS VIVIENDAS					
CUESTIONARIO BÁSICO					
VIVIENDAS_EE					
Cons.	Descripción	Mnemónico	Pregunta y categoría	Rango Válido	Longitud
IDENTIFICACIÓN GEOGRÁFICA					
1	Entidad Federativa	ENT		{01..32}	2
2	Municipio o Delegación	MUN		{001..570}	3
3	AGEB	AGEB	Area geoestadística básica		4
4	Localidad	LOC		{0001..9999}	4
5	Manzana	MZA	Manzana	{001...999}	3
					16
LLAVE ÚNICA DE VIVIENDA					
6	Identificación de vivienda	idvivienda		{1..9999999}	7
CARACTERÍSTICAS DE LAS VIVIENDAS					
7	Clase de vivienda particular	CLAVIVP	Clase de vivienda particular	{1..7,9,b}	1
			Casa independiente	1	
			Departamento en edificio	2	
			Vivienda en vecindad	3	
			Vivienda en cuarto de azotea	4	
			Local no construido para habitación	5	
			Vivienda móvil	6	
			Refugio	7	
			No especificado	9	
			Blanco por pase	b	
8	Pisos	PISOS	¿De qué material es la mayor parte del piso de esta vivienda?	{1..3,9,b}	1
			Tierra	1	

Open data

- In 2012 we adopted a Norm for open data dissemination.
- All producers have to abide by it and that format is made available by statistical programme.
- We rolled out a detailed manual so that producers know what metadata and data has to be included.

2010 Census of Population and Housing Units



From May 31st to June 25th, the 2010 Population and Housing Census took place. More than 106,000 interviewers went over the country with the purpose of visiting each home and ask some questions about them and their occupants.

Note: Derived from the judgment of the Plenum of the Supreme Court of Justice of the Nation regarding Constitutional Controversy 41/2011, published on June 26th, 2013. All results of the 2010 Population and Housing Census for the Tultepec, Nextlalpan and Tultitlan municipalities, State of Mexico.

[Read more](#)

[Related topics](#)

[Documentation](#) [Tabular data](#) [Microdata](#) **[Open data](#)** [Publications](#) [Tools](#)

Search in title



Results found: **65**

[Download all files](#)

Open data standard

Title	Period	Formats
- Download files		
- Main results by location (ITER)		
Estados Unidos Mexicanos	2010	i o CSV ↓ 23.4 MB
Aguascalientes	2010	i o CSV ↓ 216 KB
Baja California	2010	i o CSV ↓ 372 KB
Baja California Sur	2010	i o CSV ↓ 172 KB
Campeche	2010	i o CSV ↓ 222 KB
Coahuila de Zaragoza	2010	i o CSV ↓ 341 KB

Open data

- Each “package” contains a zip file with the same structure for all statistical programmes.
- Producers have to provide the catalogues used in the data in an open format.
- Data and its data dictionary as well as its metadata have to be made available also in an open format.

Name	Size	Packed	Type	Modified	CRC32
..			File folder		
catalogos	459	169	File folder	8/29/2016 1:24 ...	
conjunto_de_datos	105,569,538	24,476,201	File folder	8/29/2016 1:24 ...	
diccionario_de_datos	43,506	6,271	File folder	8/29/2016 1:24 ...	
metadatos	1,330	680	File folder	8/29/2016 1:24 ...	

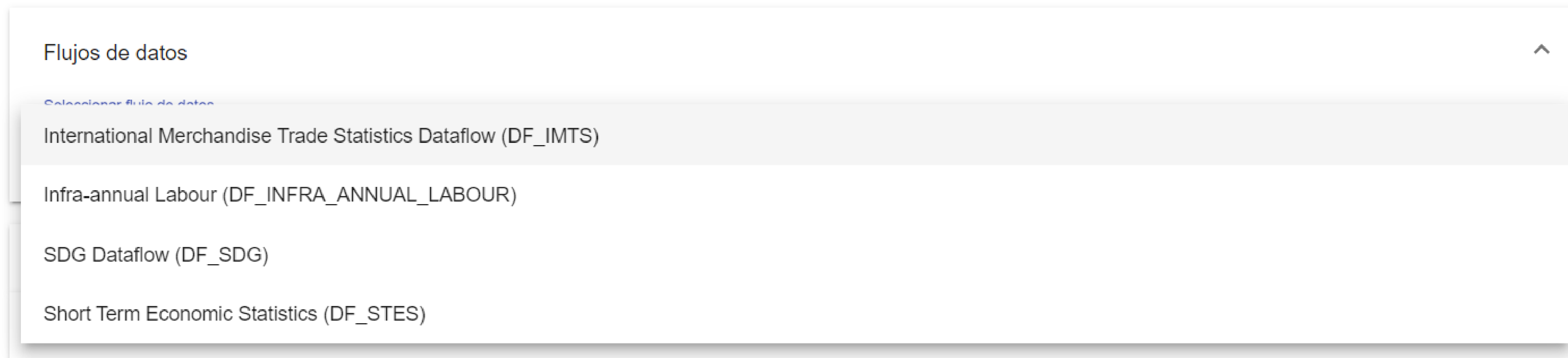
Other metadata standards

- For survey and census data, we have also adopted the DDI 2.0 standard.
- We have a “national network of metadata” where the user can access the metadata in the DDI standard by statistical programme.

The screenshot displays the INEGI website's 'Red Nacional de Metadatos' (National Network of Metadata) interface. The page features a search bar at the top right with the text 'Buscar...' and a search icon. The main navigation menu includes 'Inicio', 'Datos', 'Servicios', 'Transparencia', and 'Investigación'. The page title is 'Inicio > Catálogo de proyectos'. The main content area is titled 'Red Nacional de Metadatos' and includes a description: 'La Red Nacional de Metadatos (RNM) es la herramienta implementada por el INEGI, como plataforma de difusión y consulta de los metadatos de los proyectos de información estadística básica; constituye un sistema de catalogación que permite a los usuarios realizar búsquedas y consultas generales sobre las operaciones estadísticas o específicas a nivel de variable de los proyectos estadísticos que han sido documentados de manera exhaustiva en la Iniciativa de Documentación de Datos (DDI, por sus siglas en inglés).' Below the description is a search section with a 'Buscar por palabra clave' field, a 'Buscar' button, and a 'Restablecer' button. There are also filters for 'Filtrar por Año' (showing operations from 1985 to 2020) and 'Filtro por tipo de acceso a datos' (with options for 'ninguna', 'Microdatos disponibles', and 'Datos no disponibles'). The main content area is titled 'Catálogo de proyectos' and shows a list of projects. The first project is 'Balanza Comercial de Mercancías de México 2020, Información revisada al mes de septiembre. México, 2020'. The second project is 'Encuesta Mensual de la Industria Manufacturera 2020, Datos al mes de agosto (2013=100) México, 2020'. The page also includes a pagination bar showing 'Mostrando 1-15 de 398 operaciones estadísticas' and a 'Siguiete' button.

Other metadata standards

- Some international agencies require data in an SDMX format.
- Specific flows are made available in an SDMX-compliant format.



The screenshot shows a web interface with a header "Flujos de datos" and a dropdown menu. The dropdown menu is open, displaying a list of data flows. The first item is "International Merchandise Trade Statistics Dataflow (DF_IMTS)", which is highlighted. Other items include "Infra-annual Labour (DF_INFRA_ANNUAL_LABOUR)", "SDG Dataflow (DF_SDG)", and "Short Term Economic Statistics (DF_STES)".

Flujos de datos ^

[Seleccionar flujo de datos](#)

- International Merchandise Trade Statistics Dataflow (DF_IMTS)
- Infra-annual Labour (DF_INFRA_ANNUAL_LABOUR)
- SDG Dataflow (DF_SDG)
- Short Term Economic Statistics (DF_STES)

How can we improve
usability /
discoverability /
searchability?

Usability

- For usability improvement we have created a team whose purpose is to “defend” user’s needs.
- Industry practices such as user center design, segmentation, personas, A/B testing, etc are part of the toolbox for this team.
- Our main conceptual tool is to have a relatively fixed segmentation of our users.
- We define segments for which we design tools and products to consume data and metadata based upon their intended usage, needs and technical savvyness.
- Example of our segments:
 - Policymaker
 - Policy evaluator
 - Data producer
 - Researcher
 - Analyst

Accessibility (NQAF) is defined by the user




They do not know the data exists

They do not know in what way the information can be used.

The products we make for them to access the data are not useful for their purposes/technical expertise.

A **user** has an accessibility gap when:



We seek to
“speak” in each
user’s terms

We create user profiles or “personas” for each segment so that we are able to understand their intended uses for the data.

For each “persona” we characterise their intended usage, technical expertise, the formats to which they are accustomed.

For each persona we define:

- 1.A set of products (compilation of data presentations) (GSIM definition of producto)**
- 2.A set of promotion activities (webinars, tutorials, etc)**

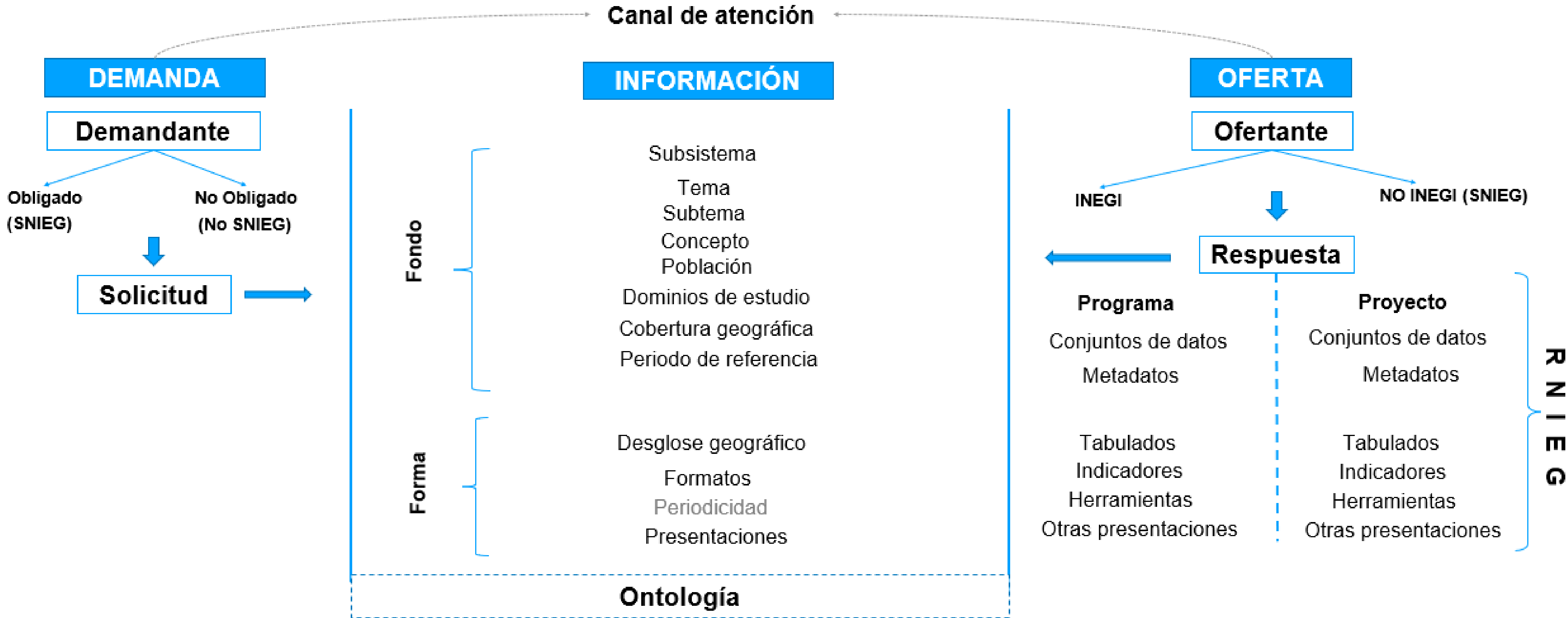
Data discoverability

- Most of the usage in our website is based on referrals (from search tools such as Google and from social networks such as Facebook).
- Most of our work so far to improve discoverability has been based on improving our SEO.
- The main tool we use for this is Google search console (<https://search.google.com/>)
- This allows us to include an enriched search maps in a way that users can better find what they need while searching the web in their preferred search engine.



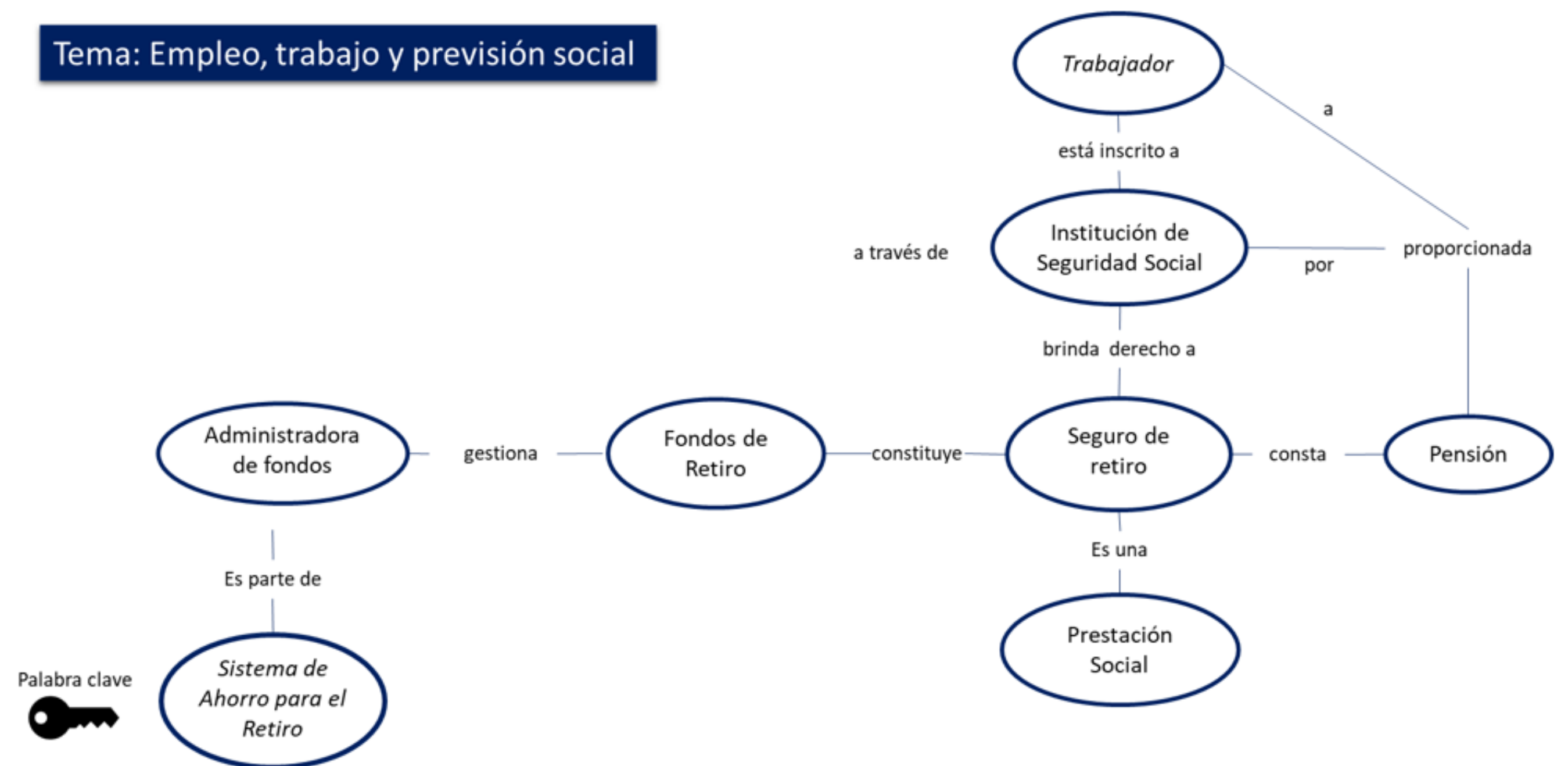
Data discoverability

- We are currently developing an algorithm with our data science team to improve our website's search engine.



Data discoverability

- We are using graphs to create an enrich semantic versión of the way different concepts interconnect.
- In that way, we are “breaking” a user’s question into different pieces and do a probabilistic match to our information inventory.



What features should
DYB's future data
dissemination have?

Some recommendations

- Adopt a metadata standard.
 - Based on the type of data the DYB has, probably SDMX would be the best bet.
 - That would require defining a DSD.
 - For strategic input and technical support, I would recommend contacting the SDMX group.
- Characterize and segment your userbase.
 - Trying to attend to all users needs is imposible.
 - Segment your core users and charaterise them (via semi structured interviews of a representative group of 5-6 users per segment is usually enough to get a sense of their intended usage and pain)
 - Based your strategy for new features on your core user's needs, pain points and technical savvyneess.
 - Always design with the intended user's perspective and needs in mind.

Conociendo México

01 800 111 46 34

www.inegi.org.mx

atencion.usuarios@inegi.org.mx



INEGI Informa