

# **Best International Practices Translate into Local Context - Microdata Laboratory of the National Institute of Statistics and Geography of Mexico**

Natalia Volkow\*

\* National Institute of Statistics and Geography, natalia.volkow@inegi.org.mx

**Abstract:** The paper presents the definition and operation of developing microdata access services of the National Institute of Statistics and Geography following international best practices. The local context has special characteristics that had to be taken in consideration to make the Microdata Laboratory operational and safeguard the legitimate use of statistical information for research purpose or to support the definition, operation or evaluation of public policy.

## **1 Background**

The National of Institute of Statistics and Geography (INEGI) has strived to harness the potential that offers information and communication technology (ICT) in dissemination of data to provide access to statistical information to different types of users. The institutional web site was developed in 1995 and through it, the Institute has offered various tools to access aggregate statistical information, always striving to provide as much geographical disaggregation as it was possible, without violating the confidentiality and taking advantage of the state of the art of ICT. In the past, the access to microdata was offered in an ad hoc way by the areas that generate information, on a personal base relation with the research community. The Institute has also offered the service of special tables that are generated upon the request of users. The access to this service is through the web site, a call centre, an email service and through the network of information centres that the Institute has in all main cities of the country.

The Law of the National System of Statistical and Geographic Information was published in April 16, 2008 provided autonomy to the Institute and created a Board of Governors to regulate the National Statistical and Geographic Information System. The article 100 of this Law establishes that the Institute, following best international practices, will provide access to microdata. In the first session of the Board of Governors of the Institute, they authorized the terms and conditions in which the Institute will provide access to microdata of all statistical projects carried out by the Institute.

This agreement established that access to household surveys, government censuses and the samples of the household and population censuses will be provided direct from the Institute web site by download of the files of the household surveys that contain raw microdata. The confidentiality is kept by limiting to municipality level the geographic location of each unit of observation from the random sample.

This agreement also established that access to microdata from surveys and censuses of establishments and agricultural units and from the population and household census will be provided through a research data centre and remote processing only for research purpose or to public servants that are defining or evaluating public policies. The Institute also kept in operation the regular public service of processing specific tabulates, upon request with charge.

## **2 Microdata Laboratory Project**

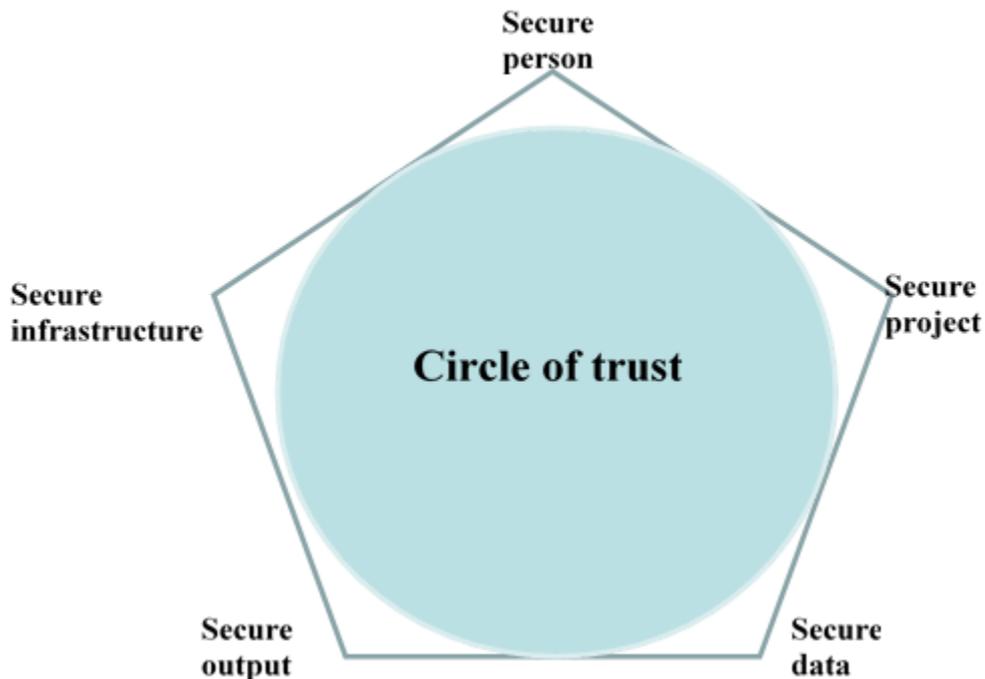
INEGI appointed a leader in charge of defining an implementing an integral strategy to institutionalize the service of microdata access that comply the following activities:

- Develop a microdata laboratory
- Integration of the legal framework for the operation of the service of microdata access
- Develop a section in INEGI web site dedicated to microdata access with information architecture perspective
- Develop an information system to support the follow up of

The person in charge of the project attended three international meetings on this subject, did a bibliographic research, interviewed experts, analysed web sites of other National Statistics Offices (NSO) and international organizations that offer access to microdata for research purpose and visit research data centres in the United Kingdom and the United States to come to know best international practices in access to microdata.

INEGI decided to develop a microdata laboratory base on the risk management model developed by the Office of National Statistics of the United Kingdom (Ritchie, 2008). This model considers, with a holistic perspective, five elements that should be secure altogether, to create what is call a circle of trust, where authorized users will have direct access to the microdata that are confidential information. The challenge was to put in operation this best international practice into the local context of INEGI with scarce resources.

## **Risk management model**



INEGI developed the technical infrastructure of the laboratory with the same type of technological platform that the Secure Data Service of the UK operates in its Secure Data Service. The platform allows the creation of a virtual space in which only authorized users with specific privileges can enter the secure enclave through doom machines. The personnel of the areas that generate information can access to input the data need for a session of a specific user or extract the outputs of the processing that the user has done to check them for confidentiality, before they are deliver to user. The laboratory has twelve terminals, they are in secure space that has biometric systems control for the access and only the personnel of the laboratory can enter, it also has CCTV hours 24 surveillance.

As for the secure persons, only public servants, postgraduate students and researchers of an academic or research institutions, or personnel from international organizations can apply to the service of the Laboratory. Each institution must sign an agreement with INEGI to provide the accreditation of the people affiliate to it. Following the practice carried by the Secure Data Service and the UK Economic and Social Research Council, INEGI signed an agreement with the National Council of Science and

Technology (CONACYT) by which the researchers and students that are receiving benefits from the Council will have its accreditation, and will be able to use the Laboratory.

The service is provided through Internet. Users must download an application format, fill it and send it with their Curriculum Vitae, document of proof of institutional affiliation and official identification. In the format, they must describe the aim of its project, methodology and justify the need for the access to microdata. The research project must serve public good, by supporting the definition or evaluation of public policy or else academic research to be published, so it is publicly available. The data required must be justified in terms of the research project aim. The output of the processing is checked by the personnel of the areas that generate information to validate they do not breach confidentiality.

The procedure was defined like a star, having one single point of contact with users that manages the relation with all the areas of INEGI that generate statistical information. It is the experts that download the microdata in the virtual desk and them who extract the output and check it and if it does not breach confidentiality cleared it and send it to the single point of entry that delivers it to the users.

As for the legal framework, the Institute developed two norms, one that regulates the statistical information it generates and another for the National Statistical and Geographic Information System (NSGIS). By law it is compulsory for the Institute to provide access to the microdata of all its statistical projects. Each of the institutions that integrate the NSGIS must define if they provide access to the microdata of its statistical projects. If they decide to do so, they can sign an agreement with the Institute that will allow them to use the laboratory and avoid in this way the duplication of the infrastructure.

INEGI develop a section in its web site dedicated to Microdata, it offers to users all the data to which they provide access to microdata, the terms and conditions in which they could have access to it. For each statistical project to which it provides access to its microdata.

INEGI INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA

México

Inicio | Contacto | Síguenos: RSS, Twitter, Facebook, YouTube, LinkedIn

Estadística ▾ Geografía ▾ Investigación ▾ Productos y Servicios ▾ Acerca del INEGI ▾

Inicio > Estadística >

## Microdatos

El INEGI, en cumplimiento con lo dispuesto en el marco regulatorio, proporciona conforme al Calendario de Difusión de Información Estadística y Geográfica dos formas de acceso a microdatos:

- Directo mediante descarga gratuita de archivos de microdatos a través de esta página.
- Indirecto, a través de los siguientes mecanismos:
  - Laboratorio de microdatos para servidores públicos del Estado Mexicano, funcionarios de organismos internacionales e investigadores y estudiantes de posgrado.
  - Servicio de procesamiento para el resto de la sociedad, el cual se puede gestionar a través de esta página o en los Centros de información INEGI.

Con la finalidad de facilitar la definición de su requerimiento, se recomienda consultar los aspectos metodológicos del proyecto de su interés.

### Censos y conteos

- ▶ Agropecuarios y ejidales
- ▶ Características de las localidades y del entorno urbano
- ▶ Económicos
- ▶ Gobierno\*
- ▶ Población y vivienda

### Encuesta Nacional Agropecuaria

- 2014
- 2012

### Encuestas en establecimientos

- ▶ Anuales
- ▶ Mensuales
- ▶ Especiales

### Encuestas en hogares\*

- ▶ Encuestas y muestras vinculadas a censos o conteos
- ▶ Regulares
- ▶ Especiales
- ▶ Módulos
- ▶ Históricas

### Otras encuestas\*

- Población Privada de la Libertad (ENPOL) 2016
- Prácticas de lectura - 2006

### Proyectos de vinculación de datos

- Perfil de las empresas manufactureras de exportación EAIM - Comercio Exterior (SCIAN 2007 - 21 subsectores) - 2007 en adelante

### Registros administrativos - estadísticas

- ▶ Económicas
- ▶ Sociales
- ▶ Vitales\*

\* El acceso a los microdatos de estos operativos es mediante descarga del archivo, pero se puede ofrecer Laboratorio de microdatos si así lo requiere el usuario.

Continuar

Each statistical project has a microdata web page in this section except for those that the access is by download. The microdata pages of the statistical project have standardized navigation and common look and feel. Each page includes two subsections. The first is for all the methodological aspects available, ideally in this subsection each project should have:

- Presentation with short description of the main attributes of the statistical project that provides users with a rapid overview of it.
- Methodological document
- Glossary
- Questionnaire
- File descriptor

- Example of the data base (with no representative data)
- Catalogues
- DDI Metadata if it is available

INEGI INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA

Inicio | Contacto | Síguenos: RSS, Twitter, Facebook, YouTube

Estadística | Geografía | Investigación | Productos y Servicios | Acerca del INEGI

Inicio > Estadística > Encuestas en establecimientos >

## Encuestas en Establecimientos

### Encuesta Anual de la Industria Manufacturera (EAIM) - 2009 en adelante

#### Microdatos

##### Aspectos metodológicos

- Presentación
- ▶ Catálogos
- Cuestionario
- Estructura de archivos (FD)
- Glosario
- ▶ Metadato (DDI)
- Síntesis metodológica
- Ejemplo de la base de datos con valores alterados, solo permite mostrar las características de ésta y probar algoritmos.

DBF | STATA | CSV

[Ver todas las series disponibles](#)

#### Acceso a microdatos

- Elija el tipo de servicio al cual puede tener acceso de acuerdo a su adscripción institucional:

**Laboratorio de microdatos/Procesamiento remoto**

- Servidores públicos de instituciones del Estado mexicano
- Funcionarios de organismos internacionales y organismos públicos de estadística
- Investigadores de instituciones académicas
- Estudiantes de instituciones académicas de nivel superior

[Continuar](#)

**Servicio de procesamiento**

Público en general

[Continuar](#)

#### Otras formas de consulta

- ▶ Consulta en línea
  - Serie histórica BIE
- Publicaciones

The second subsection includes the terms and conditions in which users can have access to the microdata of the statistical project. They are two ways to have access to microdata the laboratory and the requirement of specific tables, the possibility to choose one or the other depends on the institutional affiliation of the user, as it was mentioned before.

The main achievements of the project were to create institutional services, before users could have access if they had personal connections. Nowadays all the relation is done through the web page and the institutional email [microdatos@inegi.org.mx](mailto:microdatos@inegi.org.mx)

Both services Remote processing and the Laboratory, that was inaugurated in April 2013, are provided for academic research on a free border basis and free of charge. To enter the Laboratory, users need to have an accreditation of their institution that must sign an agreement with INEGI, and he or she must sign some terms and conditions and undertake a training session. We have signed 39 agreements for accreditation, 7 with international organizations, 14 with foreign academic institutions and 18 Mexican academic institutions.

The number of users being serviced have increased steadily since then.

<b>Year</b>	<b>Research Projects</b>	<b>Remote Processing</b>	<b>Microdata Laboratory</b>
2011	6	6	-
2012	41	41	-
2013	138	127	11
2014	164	143	21
2015	147	102	45
2016	158	88	70
2017	84	37	47
<b>Total</b>	<b>738</b>	<b>544</b>	<b>194</b>

The services have make evident the need to work data architecture and to try to recuperate historical data to harness the potential of research. The areas that generate the information that at the beginning were reticent to the project, have support the services and even modify certain aspects to support users. We have the challenge to create a sole technical platform both for users and the technical people of the areas that generate information to facilitate the process of servicing users working from the Data warehouse. As with the original project of the Laboratory this new stage faces not only technical but organizational challenges.

## References

Ley del Sistema Nacional de Información Estadística y Geográfica, (2008) (Diario Oficial de la Federación México).

Ritchie, Felix. *Secure access to confidential microdata: four years of the Virtual Microdata Laboratory*, Economic & Labour Market Review May 2008, Volume 2, Issue 5, pp 29–34.

Organisation for Economic Co-operation and Development (OECD) (2014). *Microdata Access for OECD Research and Policy Analysis*, OECD: Paris France.