INEGI’S STATISTICAL AND GEOGRAPHICAL INFORMATION STRATEGY

Manuel Cuéllar-Río, Deputy Director General of Information Integration
manuel.cuellar@inegi.org.mx

ModernStats World Workshop 2022

Belgrade, Serbia
June 2022
GOAL

To consolidate a data ecosystem that contributes to satisfy the needs of users of statistical and geographical information, and that complies with security, confidentiality, quality, and interoperability principles.
STRATEGIC FRAMEWORK

STATISTICAL AND GEOGRAPHICAL INFORMATION

PRODUCTION → INTEGRATION → DISSEMINATION

GOVERNANCE + ARCHITECTURE + METADATA

SECURITY + CONFIDENTIALITY + QUALITY + INTEROPERABILITY

TECHNOLOGICAL INFRASTRUCTURE
CONVERGENCE OF FOUNDATIONS

- Governance
- Architecture
- Metadata
- Security
- Confidentiality
- Quality
- Interoperability
- Technological Infrastructure

Data and Information sets
STATISTICAL AND GEOGRAPHICAL INFORMATION

PRODUCTION

MPEG

Modelo del Proceso
Estadístico y Geográfico

NEEDS/DATA

1. Documentación de necesidades
2. Diseño
3. Construcción
4. Captación
5. Procesamiento
6. Análisis de la producción
7. Difusión
8. Evaluación del proceso

IP/PROCESSES

GOVERNANCE + ARCHITECTURE + METADATA

SECURITY

CONFIDENTIALITY

QUALITY

INTEROPERABILITY

TECHNOLOGICAL INFRASTRUCTURE
STATISTICAL AND GEOGRAPHICAL INFORMATION

INTEGRATION

Sources
- IP/Processes
- Ministries
- Non traditional

Integration
- Extract
- Transform
- Load

Repositories
- DW
- Data Lake

GOVERNANCE + ARCHITECTURE + METADATA

SECURITY
- CONFIDENTIALITY
- QUALITY
- INTEROPERABILITY

TECHNOLOGICAL INFRASTRUCTURE
DISSEMINATION

Repositories

Data Lake

Applications

Queries and visualization

Data mining

Analytics

Services

Portals

Devices

Web services

GOVERNANCE + ARCHITECTURE + METADATA

SECURITY

CONFIDENTIALITY

QUALITY

INTEROPERABILITY

TECHNOLOGICAL INFRASTRUCTURE
CONSOLIDATION OF STRATEGIC FRAMEWORK

PRODUCTION
- Internal sources
- External and non-traditional sources

INTEGRATION
- Data Lake
- Raw data
- Data query and transformation at run-time
- ETL

DISSEMINATION
- Data virtualization
- Reports and analysis
- Download and exchange
- Data Science

GOVERNANCE + ARCHITECTURE + METADATA

SECURITY
CONFIDENTIALITY
QUALITY
INTEROPERABILITY

TECHNOLOGICAL INFRASTRUCTURE
# Roadmap

<table>
<thead>
<tr>
<th>Foundation</th>
<th>Baseline</th>
<th>Medium Term</th>
<th>Long Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>Governance and Architecture</td>
<td>Regulatory framework; guidelines for surveys conceptual and sample design; operational systems (processes documentation, information programs inventory, processes’ costs); general and specialized training; first set of processes architectures.</td>
<td>Consolidation of regulatory framework, with guidelines for design phase and tools for standardized documentation; interoperability between production processes and programming, budgeting, and costs; training and communication integral strategy; architecture models by production method.</td>
<td>Production process and overarching activities are harmonized through an NSS-wise regulatory framework; consolidated data ecosystem that integrates non traditional data sources; architecture models that optimize data flows, exchange, consumption, and technological infrastructure.</td>
</tr>
<tr>
<td>Metadata</td>
<td>Metadata management strategy; regulatory reform proposal; initial development of new integrated metadata system.</td>
<td>Updated regulatory framework; integrated metadata system in operation; interoperable metadata initiatives for all production methods; training and communication integral strategy.</td>
<td>Metadata are consistently managed in production, integration, and dissemination; the integrated metadata system operated as the core of the ecosystem, ensuring that data and metadata go hand-in-hand as part of the same information service.</td>
</tr>
<tr>
<td>Security and Confidentiality</td>
<td>Regulatory framework; information security inventory system in operation; specialized training; initial practices in line with MPEG (GSBPM).</td>
<td>Training and communication integral strategy; practices fully compliant with MPEG (GSBPM).</td>
<td>Interoperable practices and systems that contribute to strengthen information security and statistical confidentiality.</td>
</tr>
<tr>
<td>Quality</td>
<td>Regulatory framework; system in operation; product quality indicators for some production methods; general and specialized training.</td>
<td>System consolidated and interoperable with operational systems; quality indicators for all production methods and dimensions; training and communication integral strategy; evaluations catalogue.</td>
<td>Quality standardization, evaluation and improvement are done based on consistent and transparent evidence.</td>
</tr>
<tr>
<td>Interoperability</td>
<td>Interoperability model; first interoperability maturity assessment; information infrastructure usage analysis.</td>
<td>Regulatory framework; regular interoperability maturity assessment; information infrastructure inventory aligned with information programs inventory; data and metadata integration models; training and communication integral strategy.</td>
<td>Consolidation of regulatory framework; wide and consistent use of information infrastructure; standard-and-best-practices-based information services that ease interoperability.</td>
</tr>
<tr>
<td>Technological Infrastructure</td>
<td>Regulatory framework; platform-based integral strategy; ICT security, computer systems, data, computer services, and ICT infrastructure.</td>
<td>Regulatory framework harmonized with data governance that Transversal platforms react agile to production, integration, and dissemination, and facilitates application and services development institutionally-wise.</td>
<td></td>
</tr>
</tbody>
</table>