

IBGE

Instituto
Brasileiro de
Geografia e
Estatística

*Brazilian Institute of
Geography and Statistics*

Privacy Enhancing Technologies

Augusto Fadel
augusto.fadel@ibge.gov.br

June 11, 2024

Enabling Private Data Analysis for External Users
The Brazilian NSO Perspective on PETs



Context

In the production of official statistics.

Demand for

Increase **timeliness** and **granularity**
reduce **costs** and maintain **quality**

Technological revolution

Alternative **data sources**
Guarantee **quality** and **confidentiality**

Democratic society

Formulation and monitoring of **public policies**
Fight **misinformation**

—

Data access as an obstacle to
scientific development
statistics production
public policies

...

–
Data is often **private**
for a very **good reason.**

—

Trade-off between
privacy and **utility**.

UN Fundamental Principles of Official Statistics

1. Relevance, Impartiality, and Equal Access
2. Professional Standards, Scientific Principles, and Professional Ethics
3. Accountability and Transparency
4. Prevention of Misuse
5. Sources of Official Statistics
6. Confidentiality
7. Legislation
8. National Coordination
9. Use of International Standards
10. International Cooperation

UN Fundamental Principles of Official Statistics

- 1. Relevance, Impartiality, and Equal Access**
2. Professional Standards, Scientific Principles, and Professional Ethics
3. Accountability and Transparency
4. Prevention of Misuse
5. Sources of Official Statistics
- 6. Confidentiality**
7. Legislation
8. National Coordination
9. Use of International Standards
10. International Cooperation

UN Fundamental Principles of Official Statistics

6. Confidentiality

// Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be **strictly confidential and used exclusively for statistical purposes.**

UN Fundamental Principles of Official Statistics

1. Relevance, Impartiality, and Equal Access

// Official statistics provide an **indispensable element in the information system of a democratic society**, serving the government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honor citizens' entitlement to public information.

— Legislation **does not prevent** **access** to sensitive data.

Encourages the development of **safe methods**.



2018 Lei Geral de
Proteção de Dados
Pessoais
(data protection regulation)

2022 Personal data
protection as a
fundamental right

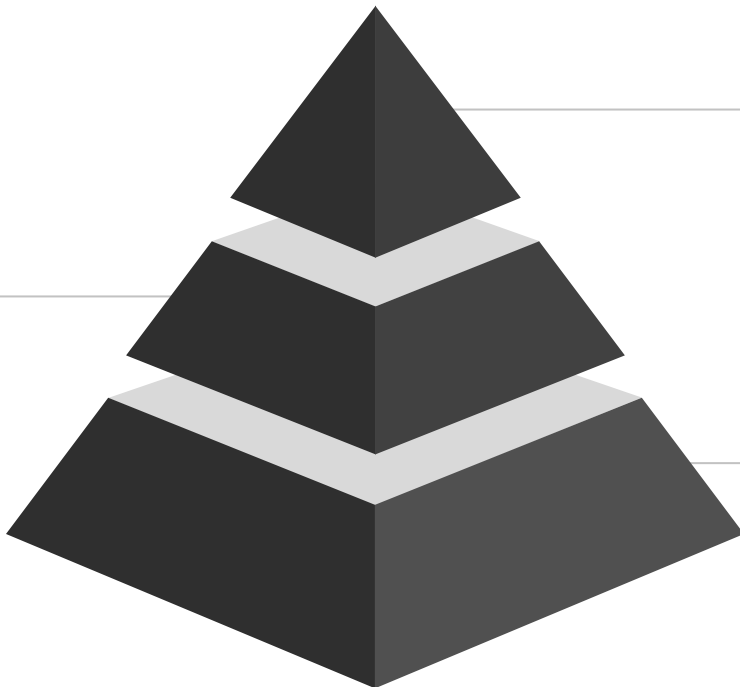
Responsible use of data.

Not just **legally** compliant, but **ethical**, **legitimate** and **fair**.

Data Privacy, Ethics and Protection: Guidance Note
on Big Data for Achievement of the 2030 Agenda

unsdg.un.org/resources/data-privacy-ethics-and-protection-guidance-note-big-data-achievement-2030-agenda





Secure Access Facility

1

- _Local and highly restricted access
- _Knowledge of specific analysis tools is required
- _Identified data
- _Maximum utility

Microdata

2

- _Remote access
- _Publicly available
- _Knowledge of an analysis tool is required
- _Traditional SDC methods
- _Intermediate utility

Cross-tabulations

3

- _Remote access
- _Publicly available
- _No knowledge of analysis tools is required
- _Traditional SDC methods
- _Limited utility

Secure Access Facility

- _Local and highly restricted access
- _Knowledge of specific analysis tools is required
- _Identified data
- _Maximum utility

1

Microdata

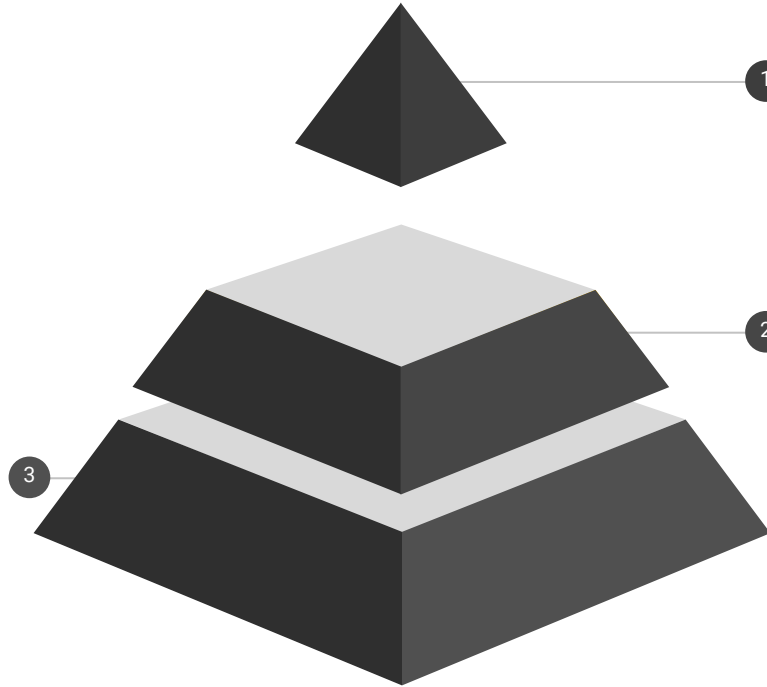
- _Remote access
- _Publicly available
- _Knowledge of an analysis tool is required
- _Traditional SDC methods
- _Intermediate utility

2

Cross-tabulations

- _Remote access
- _Publicly available
- _No knowledge of analysis tools is required
- _Traditional SDC methods
- _Limited utility

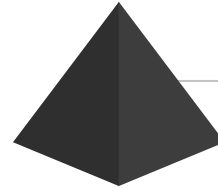
3



Secure Access Facility

- _Local and highly restricted access
- _Knowledge of specific analysis tools is required
- _Identified data
- _Maximum utility

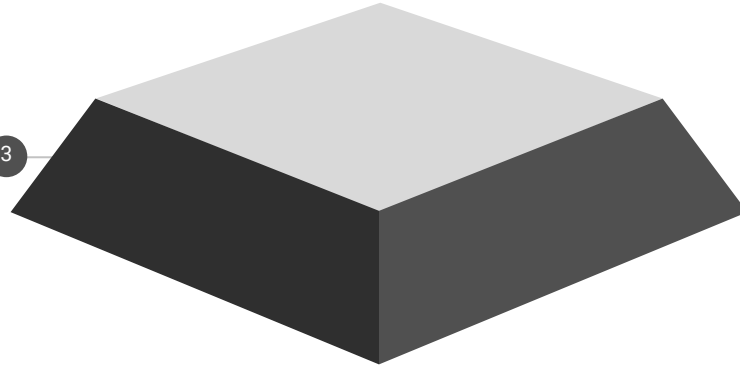
1



Cross-tabulations

- _Remote access
- _Publicly available
- _No knowledge of analysis tools is required
- _Traditional SDC methods
- _Limited utility

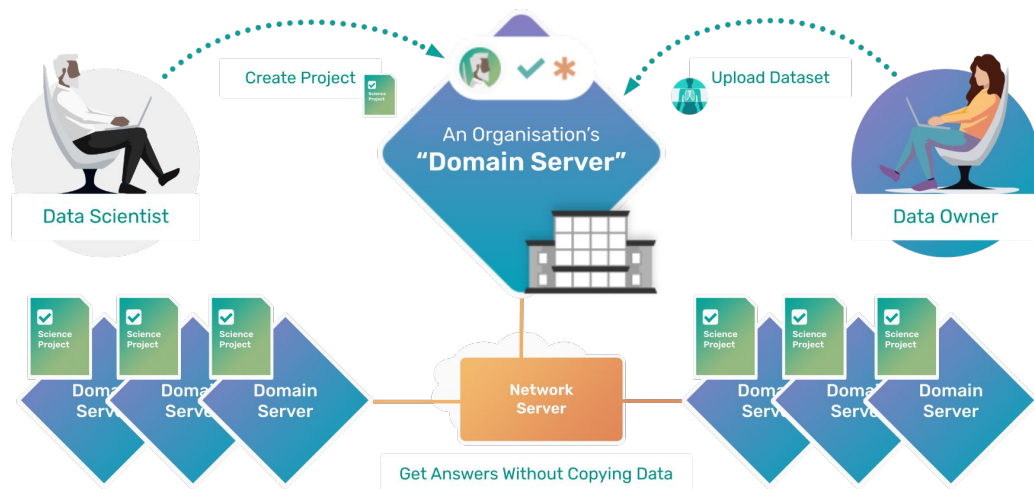
3



A close-up photograph of a silver metal padlock attached to a dark, vertically-grained wooden door. The padlock is the central focus, with its shackle and body clearly visible. The lighting is dramatic, highlighting the metallic texture of the padlock against the dark wood.

**Extract relevant
information from sensitive
data, preserving the
privacy of the individuals
represented therein.**

Privacy-Preserving Data Science



Secure Access Facility

- _Local and highly restricted access
- _Knowledge of specific analysis tools is required
- _Identified data
- _Maximum utility

1

Microdata

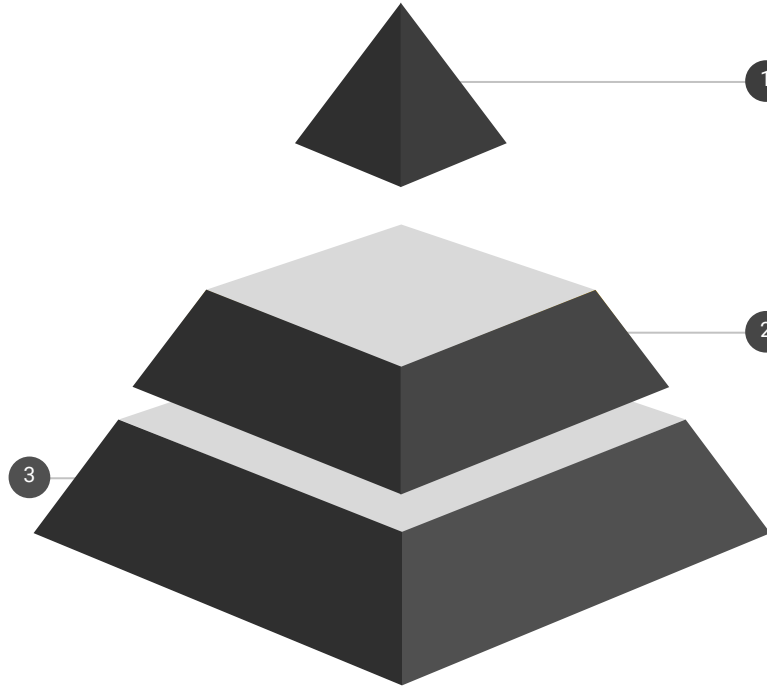
- _Remote access
- _Publicly available
- _Knowledge of an analysis tool is required
- _Traditional SDC methods
- _Intermediate utility

2

Cross-tabulations

- _Remote access
- _Publicly available
- _No knowledge of analysis tools is required
- _Traditional SDC methods
- _Limited utility

3



Secure Access Facility

- _Local and highly restricted access
- _Knowledge of specific analysis tools is required
- _Identified data
- _Maximum utility

Microdata

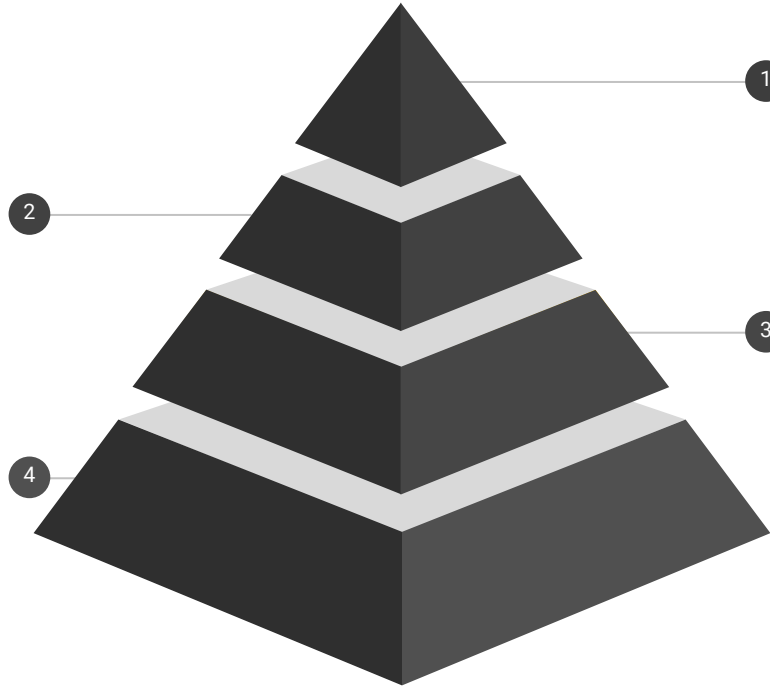
- _Remote access
- _Publicly available
- _Knowledge of an analysis tool is required
- _Traditional SDC methods
- _Intermediate utility

Privacy-Preserving Data Science

- _Remote access
- _Knowledge of specific analysis tools is required
- _SDC + PETs
- _High utility
- _Multiple data sources integration

Cross-tabulations

- _Remote access
- _Publicly available
- _No knowledge of analysis tools is required
- _Traditional SDC methods
- _Limited utility



Level 0

no risk setup

- public data
- manual audit
- isolated server

Level 1

::

- **private** data
- manual audit
- isolated server

Level 2

::

- **private** data
- **automatic** audit
(there may be exceptions)
- isolated server

Level 3

ready to scale

- **private** data
- **automatic** audit
(there may be exceptions)
- **online** server

Proof of Concept

2017 Brazilian Agricultural Census Data

censoagro2017.ibge.gov.br

- Conducted **since 1920**, it is the most comprehensive statistical and territorial investigation of **Brazilian agricultural production**
- Gross value around **US\$215 billion** (R\$1.14 trillion) in 2023
- **Environmental and climate** issues
- no publicly available **microdata**

Proof of Concept

Women from rural areas

www.embrapa.br/busca-de-noticias/-/noticia/50779965/mapa-embrapa-e-ibge-apresentam-os-dados-sobre-mulheres-rurais

Brazilian Agricultural Research Corporation (Embrapa)

Ministry of Agriculture and Livestock (MAPA)

Brazilian Institute of Geography and Statistics (IBGE)

- SDG 5 - gender equality
- Term of commitment
- Special cross-tabulations
- 2+ months



Highlights

Traditional SDC + PETs

Flexibility on utility/privacy trade-off

Benefits for external users

Unburden internal teams

Foster research/fact checking

Python knowledge

More complex infrastructure

Adapt current internal processes

UN Fundamental Principles of Official Statistics

- 1. Relevance, Impartiality, and Equal Access**
2. Professional Standards, Scientific Principles, and Professional Ethics
3. Accountability and Transparency
4. Prevention of Misuse
5. Sources of Official Statistics
- 6. Confidentiality**
7. Legislation
8. National Coordination
9. Use of International Standards
10. International Cooperation

UN Fundamental Principles of Official Statistics

- 1. Relevance, Impartiality, and Equal Access**
2. Professional Standards, Scientific Principles, and Professional Ethics
3. Accountability and Transparency
4. Prevention of Misuse
5. Sources of Official Statistics
- 6. Confidentiality**
7. Legislation
- 8. National Coordination**
9. Use of International Standards
- 10. International Cooperation**



**“When you can
measure what you are
speaking about, and
express it in numbers,
you know something
about it.”**

William Thomson (1824-1907)

Thank you!

Augusto Fadel
augusto.fadel@ibge.gov.br