



Technical Workshop on Time, People, Places and Relationships: A New Lens to Organize Social and Demographic Outcomes

POPULATION DATA SYSTEM

Inga Masiulaitytė-Šukevič, State Data Agency (Statistics Lithuania)
Federico Segui, National Institute of Statistics (INE), Uruguay

Discussant

Mauricio Rodriguez, INEGI, Mexico

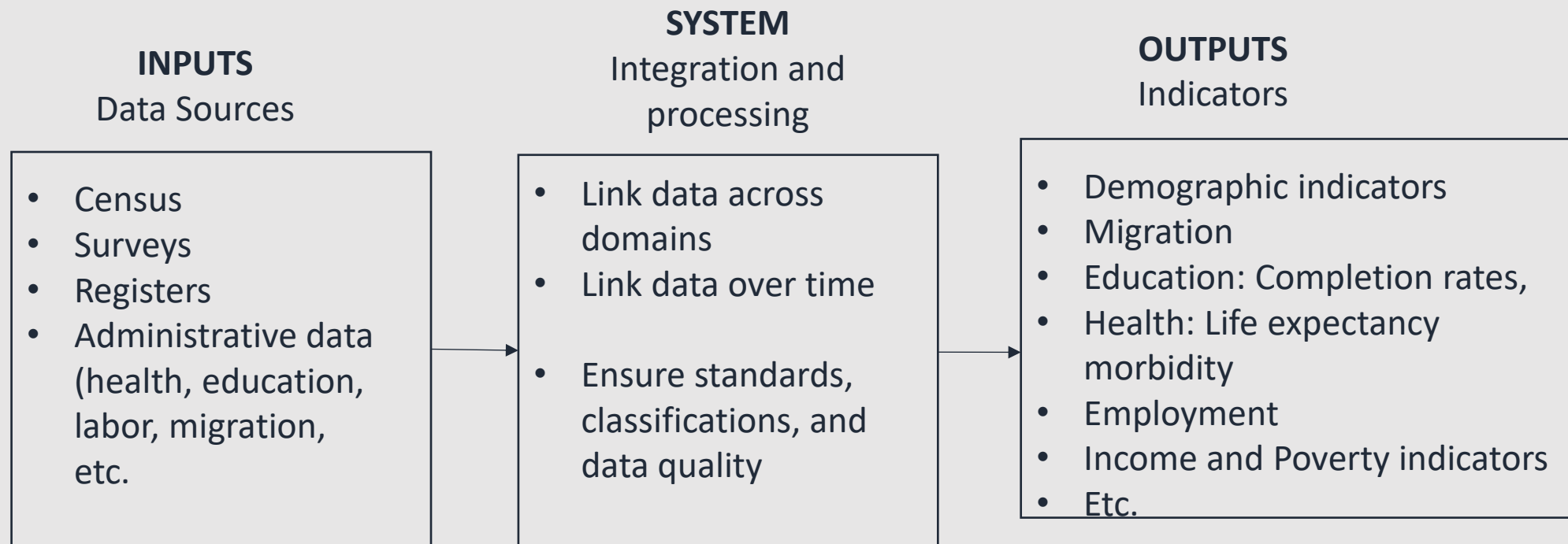


Overview

- Population Data System (formalization, components)
- Methodological maturity
- Methodological Maturity Assessment Tool



Population data system





Rationale for Formalizing a Population Data System

- **Integrated and strategic data management:** Moves from fragmented, ad hoc data compilation to a coordinated approach.
- **Enhanced policymaking:** Improves usability of population data for government planning and decision-making.
- **Structured governance:** Establishes clear roles, data standards, access protocols, and privacy safeguards.
- **Collaboration and partnerships:** Institutionalizes cooperation between NSOs and key data providers through:
 - Data-sharing agreements
 - Common metadata standards
 - Technical infrastructure investments
- **Public trust and communication:** Highlights benefits for government efficiency, citizen services, and accountability to build political and financial support.
- **Interoperability and harmonization:** Breaks down silos between censuses, population registers, and household surveys.
- **Improved data quality and access:** Ensures reliable, harmonized, and accessible population information for policy and planning.



Establishing a population base (stock)

- **Population Stock:** Up-to-date list of all usual residents at a specific point in time
- **Update Methods**
 - Continuous: Population registers
 - Periodic: Censuses adjusted with administrative data (civil registration, migration)
- **Structure & Identifiers**
 - Unique identifiers or proxy linkages for data integration
 - Core demographic attributes: sex, age, place of residence, citizenship
 - Regular updates via civil registration and migration records
- **Household & Social Attributes**
 - Household size, marital status, relationships
 - Parent/child identifiers
 - Enables standardized household/population characteristics
- **Purpose / Uses**
 - Denominators for vital rates, labor force metrics, service coverage indicators
 - Population frames for household surveys and statistical analyses

Classifying population sub-groups (segments)

Category	What it covers	Examples
Individual Attributes	Basic personal characteristics	Age, Sex, Ethnicity
Dynamic Groupings	Groupings that change based on life situation	Labor market segment (unemployed, informal, formal), Disability category, Family type
Interoperability	Using international standards + national alignment	ISCED (Education), ISCO (Occupation), ICF (Disability)
Functionality	How the system handles classification	Point-in-time status (currently employed), Transitions (student → employed → retired)
Adaptability	Flexibility to include new priorities	Digital inclusion, Care responsibilities, Environmental exposure
Standard Sub-Groups	Common categories used for policy and equity analysis	<ul style="list-style-type: none"> - Age, sex, marital status - Disability & health - Ethnicity & migration - Employment & education - Household type & income level



Establishing an alternative population base

- Population systems should be developed and constructed to meet the diverse needs of users in our rapidly changing world.
- Creating an alternative population base means using different locations, time frames, or data sources to define and count people instead of relying on the traditional “usual residence” method used in censuses.

Comparison of alternative population bases

Population base	Definition	Main data sources	Key uses
Usual residence population	People who usually live in a given area (traditional census concept).	Census, registers, household surveys	Demographic indicators, service coverage, official denominators
Workplace population	People who work in an area, regardless of where they live.	Employer records, census, CDR	Urban planning, transport, labor market analysis
Workday population	People present in an area during working hours (residents + commuters).	Travel surveys, CDR, admin data	Service demand (food, health, transport), emergency preparedness
Study place population	People attending educational institutions in an area, regardless of home.	School records, census, CDR	Education planning, transport for students, urban development
Study day population	People present in an area during school/university hours.	School records, CDR, surveys	Service demand near schools, safety, health, public transport
Second address population	People with a second residence in an area (seasonal/dual).	Registers, surveys, utility data	Seasonal planning, tourism, housing, local services



Measuring changes in size and composition (flow)

- **Dynamic changes in population** come from:
 - Births
 - Deaths
 - Migration (internal & international)
 - Household formation & transitions
- **Data sources for flow tracking:**
 - Civil Registration and Vital Statistics (**CRVS**)
 - Border movement records
 - Residence change declarations
 - Household surveys & updates
- **Importance of longitudinal data:**
 - Link events over time for better analysis
 - Enable demographic modeling & population projections
 - Support **urban planning** and **school enrollment forecasting**
 - Strengthen **emergency response** during crises
- **Countries with near-universal registration:**
 - Can continuously update population base
 - Enable **real-time monitoring** of shocks (pandemics, mass migration, disasters)
- **Policy relevance:**
 - Flow data help with **resource allocation, equity monitoring, and crisis management.**



Methodological maturity

Advancements

- Improved censuses, household surveys, and civil registration/vital statistics (CRVS).
- Wider use of international standards and administrative data.
- Data linkage using unique IDs or probabilistic matching.
- Better anonymization, machine learning, and record linkage methods.

Challenges

- Integration still project-based, not real-time.
- Metadata inconsistencies reduce comparability.
- Complex household definitions (non-traditional, multigenerational, shared custody).
- Limited ethical governance of sensitive linked data.
- Restricted access to longitudinal microdata.

Way forward

- Invest in harmonization, secure access systems, and ethical frameworks.
- Strengthen international cooperation and capacity building.



Levels of maturity of Population Data System

(almost) Fully register-based (census) population statistics system	Combined system (registers and household surveys)	Survey-based statistics system
Strengths		
<ul style="list-style-type: none"> • Near real-time population updates • Reduced respondent burden • Reduced costs • Produce longitudinal microdata for life-course and policy analysis • Mature governance frameworks for data sharing, privacy, and ethical oversight. 	<ul style="list-style-type: none"> • Use of both administrative and survey data provides flexibility • Surveys allow measurement of domains not covered in registers (e.g., informal labor, perceptions, income sources) • Registers enhance timeliness and support small-area estimates • Some progress in producing disaggregated indicators consistent with SDG reporting 	<ul style="list-style-type: none"> • Household surveys remain versatile for capturing complex social indicators (poverty, labor force, health, education) • Strong adherence to international survey methodologies ensures comparability • Capacity development programs (e.g., by UN agencies) have strengthened survey design, sampling, and indicator construction
Implementation gaps		
<ul style="list-style-type: none"> • Real-time integration • Household/family definitions • Governance and trust • International comparability 	<ul style="list-style-type: none"> • Metadata harmonization • Sustainability of integration • Coverage and quality • Ethical governance • Longitudinal analysis 	<ul style="list-style-type: none"> • Timeliness • Cost and sustainability • Integration capacity • Coverage gaps • Governance • Confidentiality and trust

Methodological Maturity Assessment Tool (MMAT) for population and social statistics systems

6 Core dimensions

- A. Legal and institutional framework
 - Existence of legal provisions for data sharing and protection.
 - Degree of institutional coordination among data providers.
 - Independence and authority of the NSO.
- B. Data sources coverage
 - Completeness of CRVS.
 - Availability of population and housing censuses.
 - Existence and quality of administrative registers (civil, tax, education, health, migration).
 - Use of alternative sources (geospatial, private sector data).
- C. Integration and linkage capacity
 - Presence of unique identifiers or standardized linkage protocols.
 - Frequency and sustainability of data integration across sources.
 - Availability of longitudinal datasets.
- D. Methodological standards and metadata
 - Adoption of international standards (e.g., SDMX, ISCED, ISCO, ICF).
 - Harmonization of classifications across domains.
 - Metadata repositories and interoperability frameworks.
- E. Data accessibility and governance
 - Existence of secure access environments (data enclaves, remote access).
 - Policies for anonymization, confidentiality, and ethical oversight.
 - Stakeholder engagement and communication of data use.
- F. Timeliness and responsiveness
 - Frequency of updates (real-time, annual, periodic).
 - Ability to capture population flows (births, deaths, migration, household transitions).
 - Resilience in crisis situations (pandemics, disasters, conflict).



MMAT: Indicators and scoring

Each dimension can be assessed with scoring indicators from 0 to 2, where 0 = not developed, 1 = partially developed, 2 = fully developed.

A. Legal and institutional framework

- Legal basis for CRVS and registers (0–2).
- Data sharing agreements across institutions (0–2).
- Independence and resources of NSO (0–2).

B. Data sources coverage

- Completeness of CRVS (births, deaths registration) (0–2).
- Existence of continuous population register (0–2).
- Coverage of migration registers (0–2).
- Regular censuses and household surveys (0–2).
- Use of alternative/big data sources (0–2).

C. Integration and linkage capacity

- Unique identifiers in registers (0–2).
- Integration of migration data with CRVS and residence registers (0–2).
- Routine integration mechanisms (not project-based) (0–2).
- Existence of longitudinal microdata (0–2).

D. Methodological standards and metadata

- Compliance with international standards (SDMX, ISCED, ISCO, ICF) (0–2).
- Consistency of definitions across domains (0–2).
- National metadata repository (0–2).

E. Data accessibility and governance

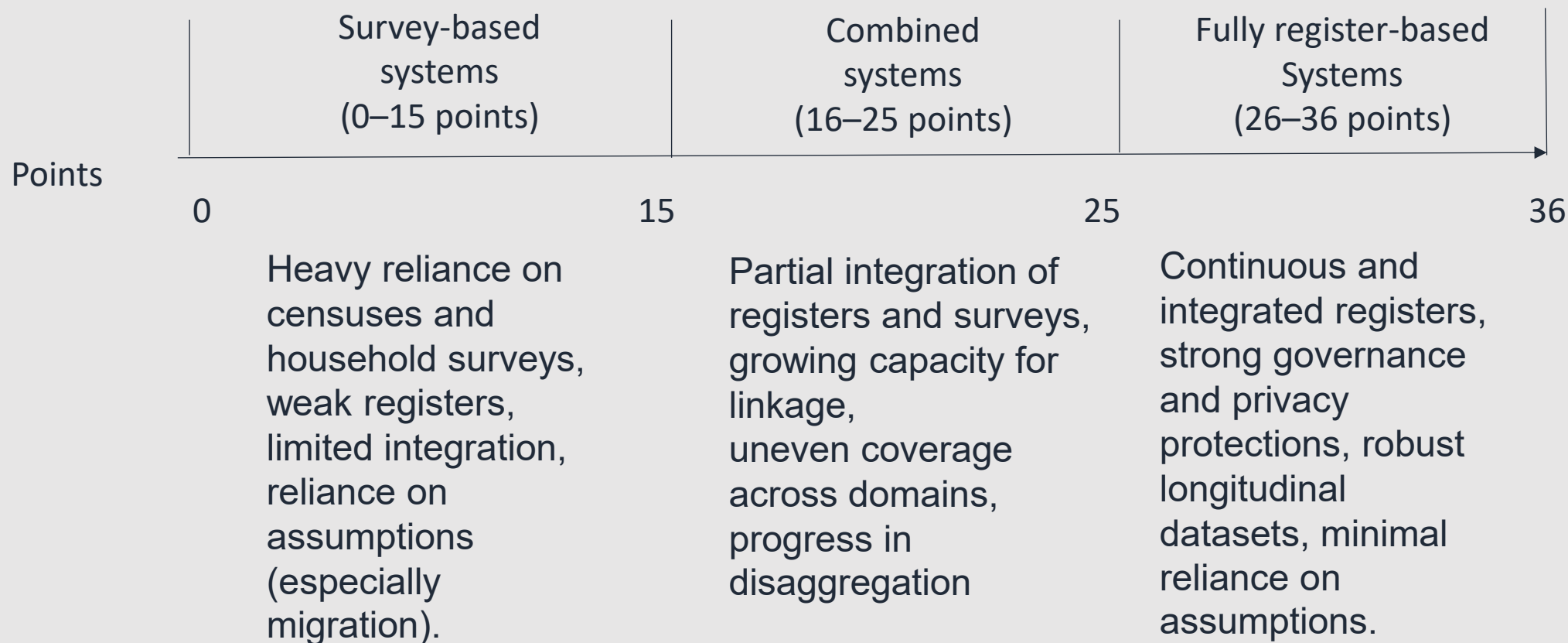
- Secure access platforms (enclaves, remote access) (0–2).
- Anonymization and disclosure control protocols (0–2).
- Transparency and stakeholder trust mechanisms (0–2).

F. Timeliness and responsiveness

- Frequency of updates to population base (0–2).
- Capacity to measure population flows (births, deaths, migration) (0–2).
- Crisis responsiveness (COVID-19, migration surges, disasters) (0–2).



MMAT: Maturity classification





Areas requiring further research

- Real-time or near-real-time integration mechanisms
- Metadata harmonization across statistical domains
- Challenges in measuring household and family structures
- Governance of linked data systems
- Migration statistics and registers
- Access to longitudinal microdata for policy research
- Ontology and domain-specific analysis
- Harmonisation of the population bases
- Usage of privately held data for official statistics



Time for discussion