

3 - 4 LUGLIO 2024

Efficientamento delle indagini: il progetto Sistema Integrato Censimento e Indagini Sociali

SILVIA LORIGA

Primo ricercatore | Istat



Conferenza Nazionale di **Statistica**

**La statistica ufficiale
nel tempo
dell'Intelligenza
Artificiale**

#CNStatistica15

Progetto Istat che prevede il coinvolgimento di molte colleghe e colleghi

DCME/MEB *Stefano Falorsi, Claudia De Vitiis, Davide Di Laurea, Alessio Guandalini, Francesca Inglese, Federica Piersimoni, Monica Russo, Marco Dionisio Terribili, Roberto Benedetti (esterno)*

DCRD *Matteo Mazziotta, Rita Ranaldi*

DCDC, DCSW *in fase di coinvolgimento*

OBIETTIVO

analisi e formulazione di **proposte**
per introdurre nel medio termine
innovazioni per migliorare la qualità e
risolvere le criticità riscontrate

OUTPUT ATTESO

concept note che fornisca elementi utili
per le decisioni strategiche

- **aspetti metodologici**
- **strategie tecniche di indagine**

Perché avviare uno studio per migliorare l'efficienza delle indagini sociali?

I disegni campionari e gli stimatori in uso non sono già scelti massimizzando l'efficienza?

Riguardo alle tecniche di indagine perché dovrebbe essere necessario un ripensamento?

Superare l'approccio stove-pipe verso una **logica di sistema**

- Tra indagini
- Tra fasi di indagine

Valutare possibili miglioramenti relativamente a **disegno campionario, metodologia di stima e tecnica di indagine**

- Considerando questi aspetti congiuntamente
- Per migliorare l'efficienza
- Correggere le distorsioni

Sfruttamento di **informazioni ausiliarie**

- Coordinando indagini sociali e Master Sample (Censimento)
- Sfruttando informazioni del SIR

Analisi e soluzione di alcune **criticità** riscontrate

- Sottocopertura selezionata liste telefoniche
- Aumento mancate risposte

Sempre **minore disponibilità di recapiti telefonici** e peggioramento della loro qualità

➡ Problemi per la conduzione di (prime interviste) Cati

Possibili soluzioni:

- **Campionamento a due fasi**
 - Massimizzando la raccolta di recapiti telefonici nella prima fase
- **Tecniche miste e Responsive-Adaptive Survey Designs** (*Metron 2024*)
- Strategie **Multiple Frame** in presenza di liste a copertura incompleta
- Eventuale acquisizione di liste di **numeri cellulari**
 - Campionamento di individui o metodi per la ricostruzione della famiglia

**Archivio Consodata
datato 2018**

**Recapiti telefonici
dal Master Sample
da 30% a 25%**

Riflessione sui questionari in relazione alle tecniche, valutando la molestia statistica

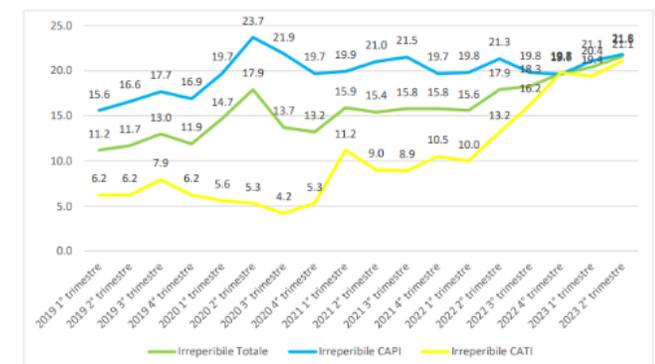
Calo dei tassi di risposta

- ➡ Diminuzione della dimensione campionaria effettiva
- ➡ Selezione del campione

- ➡ aumento degli **errori campionari**
- ➡ **distorsione** nelle stime

Possibili soluzioni:

- **Tecniche miste e Responsive-Adaptive Survey Designs**
- In caso di campionamento a due fasi, **ritorni mirati sui non rispondenti** nella prima fase
 - Eleggibilità (non residenti, seconde case...)
 - Irreperibilità
- **Uso di informazioni ausiliarie per correggere la distorsione del campione**
 - Dal Master Sample in caso di campione a due fasi
 - Da fonti amministrative



Da un contesto in cui gli output statistici erano basati quasi esclusivamente su **dati di indagine**

- Principio della rappresentatività dei campioni
- Approccio inferenziale *design-based*

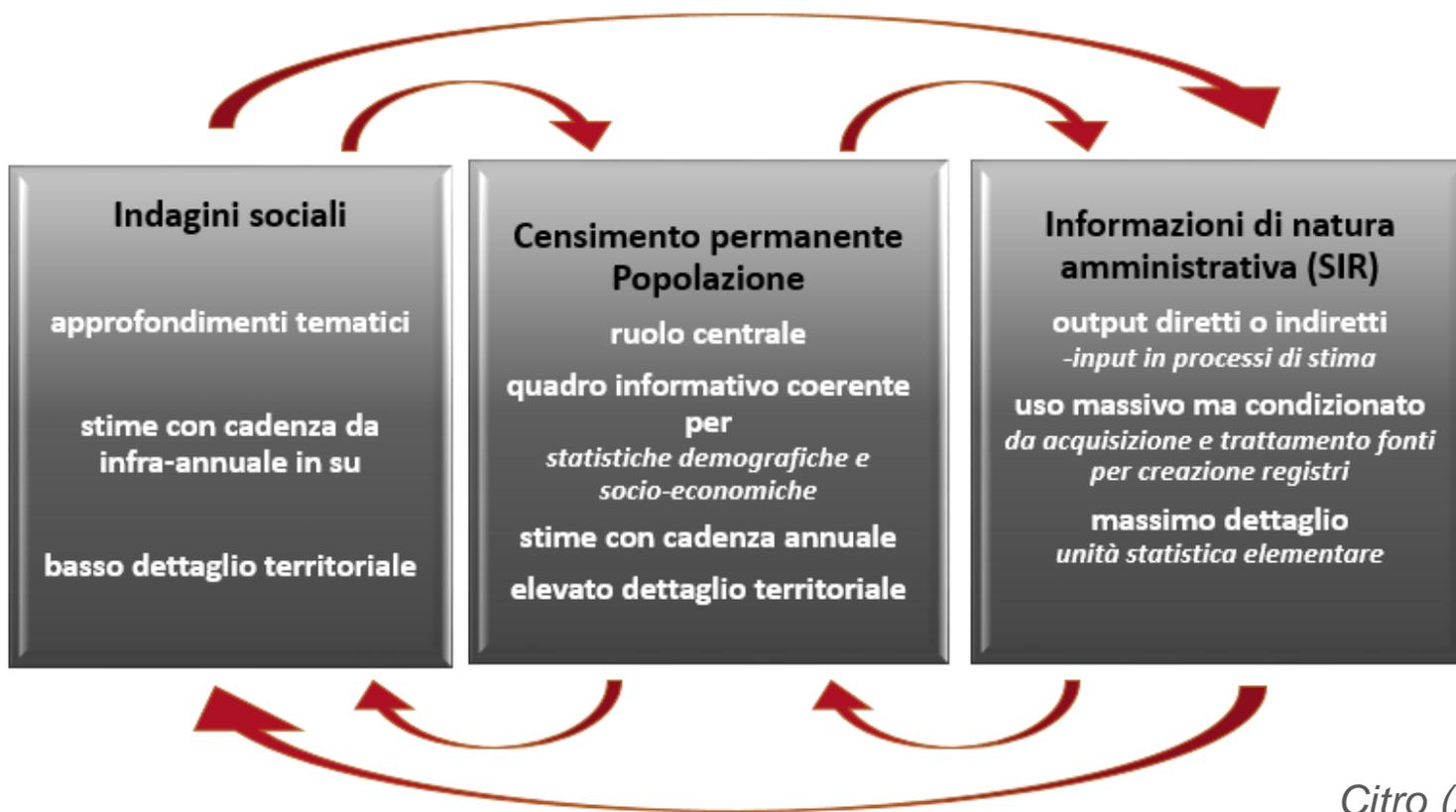
Ad un nuovo framework fondato su **processi di integrazione di dati**:

Tra **diverse indagini**

- Campionamento a due fasi
- Pooling
- Statistical matching, Projection

Tra **indagini e dati amministrativi (SIR)**

- Disegno campionario
- Imputazioni, correzione errori misura
- Correzione mancate risposte
- Metodologia di stima (diretta o indiretta)



Citro (2014), Lohr (2017), Rao (2021)

Campionamento a due fasi

- Sperimentazioni su disegni campionari **Master Sample, Aspetti Vita Quotidiana, Forze Lavoro**
- Valutazione comparativa dell'efficienza di **diversi scenari di integrazione**
 - Solo primo stadio: comuni
 - Anche secondo stadio: famiglie
 - Con/senza sfruttamento delle variabili di prima fase nella stima in seconda fase



Potenziati **miglioramenti di efficienza** o **riduzioni campionarie** a strategia data

Possibili **cambiamenti nella strategia** relativamente all'integrazione con il Master Sample

Elementi di valutazione:

- Trade-off tra integrazione e molestia statistica
- Sfruttamento della prima fase per raccogliere
 - Recapiti telefonici
 - Variabili funzionali all'integrazione con indagine di seconda fase

Campionamento spaziale

- Sperimentazioni
 - Già condotta sul **Master Sample** Indagine L
 - In corso sul primo stadio di una indagine «tipo Forze Lavoro»
- Valutazione comparativa dell'efficienza di un disegno con **selezione di comuni spazialmente bilanciati**, rispetto a **variabili con diversi gradi di correlazione spaziale**
 - Sfruttamento di stime censuarie e dati amministrativi **a livello comunale**



Potenziati **miglioramenti di efficienza** o **riduzioni campionarie** a strategia data

Miglioramento della **copertura** dei **domini non pianificati**, funzionale anche alla produzione di stime per piccole aree

Elemento di valutazione:

Analisi dell'impatto per la raccolta dei dati:
maggiore capillarità della rete di rilevazione

- **Armonizzazione** (es. stratificazione, popolazione di riferimento per calibrazione...)
- **Standardizzazione** dei processi e delle fonti (es. fonti ausiliarie, liste di estrazione...)
- **Sfruttamento integrato delle fonti:**
 - Analisi e trattamento errore misura
 - Miglioramento efficienza
 - Correzione distorsione delle stime



**COERENZA
BY DESIGN**

Coordinamento con progetto COERES (Sistema Integrato per la Coerenza dei dati dei Registri e delle Survey)

- Analisi della coerenza delle stime: domini, definizioni...
- Indicatori per la misurazione della coerenza (da monitorare nel tempo)

3 - 4 LUGLIO 2024

grazie
per l'attenzione

SILVIA LORIGA

Primo ricercatore | Istat



Conferenza Nazionale di **Statistica**

**La statistica ufficiale
nel tempo
dell'Intelligenza
Artificiale**

#CNStatistica15