



StatCities

17-18 settembre 2015, Palermo

Nuovi paradigmi per l'uso di dati di fonte amministrativa a fini statistici e censuari a livello microterritoriale

Manlio Calzaroni – Istat, Direttore centrale delle rilevazioni censuarie e dei registri statistici

Nuove esigenze

In tutti i settori della statistica continua ad aumentare la necessità di informazioni. Con l'aumento della complessità e dell'interrelazione dei dati rilevati, crescono anche le esigenze degli utenti di disporre di dati **Integrati, coerenti e con il massimo dettaglio territoriale**.

su tematiche che riflettono diversi **fenomeni** di base **correlati e interdipendenti**.

Quindi il modello "stovepipe" in cui le statistiche nei diversi settori vengono prodotte in modo indipendente non è adatto a soddisfare le esigenze politiche di insiemi di dati integrati.

Nuovo modo di produrre

... le statistiche per settori specifici non sarebbero più prodotte in modo indipendente; sarebbero invece prodotte come **parti integrate in sistemi di produzione completi** [impostazione delle statistiche basata sull'idea di un magazzino di dati (*data warehouse*)] per gruppi di statistiche.

Questi sistemi sarebbero basati su una comune infrastruttura (tecnica); applicherebbero nella misura del possibile software standardizzato e utilizzerebbero tutte le fonti di dati disponibili (*statistiche e, soprattutto, amministrative*).

Nuovi problemi da affrontare

A tal fine occorre individuare come le informazioni da fonti diverse possono essere messe insieme e sfruttate per scopi diversi, ad es.:

- mediante l'eliminazione di differenze metodologiche,
- uniformando le classificazioni statistiche.

***Per ottenere questi risultati:
È indispensabile integrare microdati,
cioè dati di prevalente origine amministrativa***

Che consentono un dettaglio classificatorio e territoriale adeguato

Prospettive da Eurostat

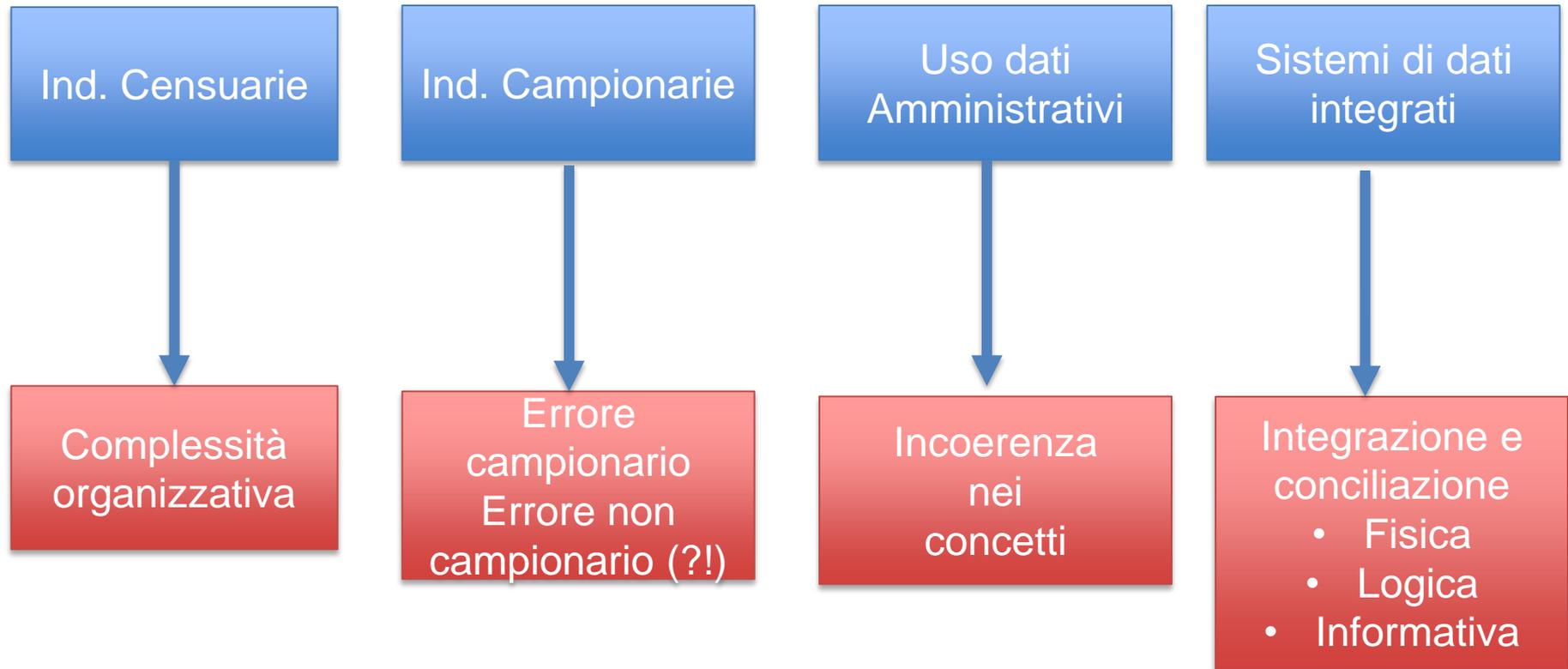
COMUNICAZIONE DELLA COMMISSIONE AL PARLAMENTO EUROPEO E AL CONSIGLIO

**sul metodo di produzione delle statistiche UE:
una visione per il prossimo decennio**

Bruxelles, 10.8.2009

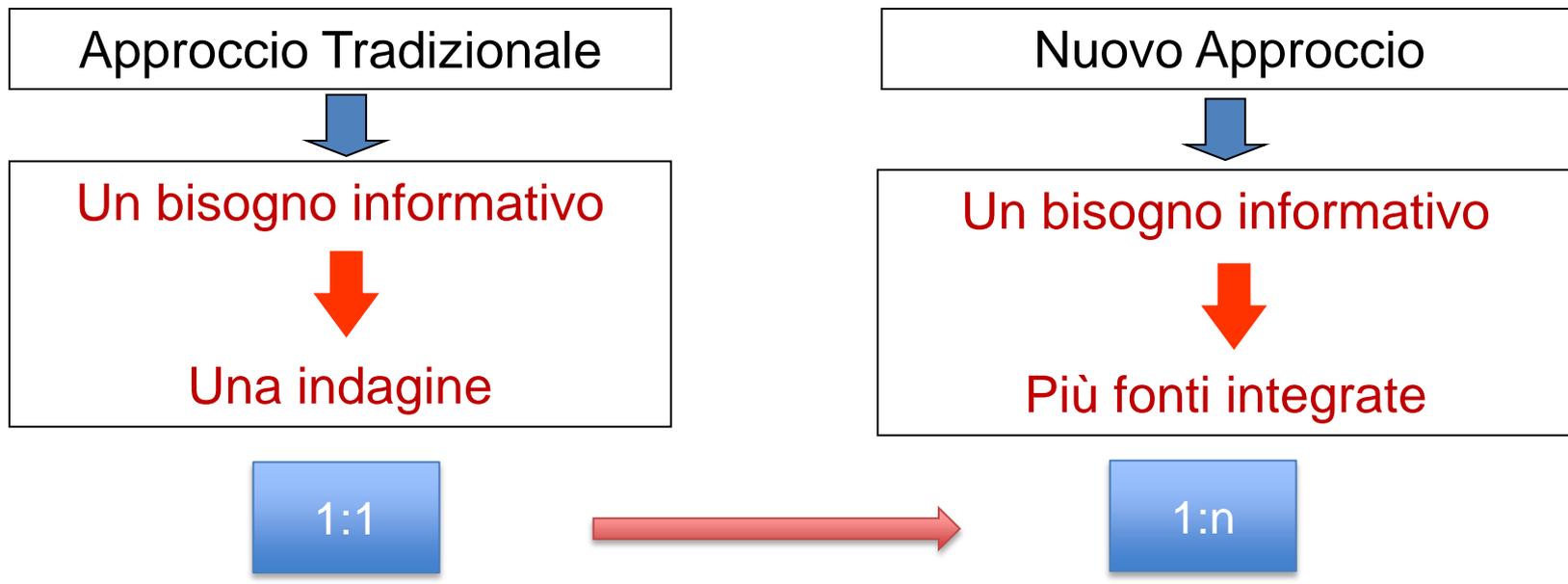
COM(2009) 404 definitivo

Evoluzione del processo di produzione della statistica

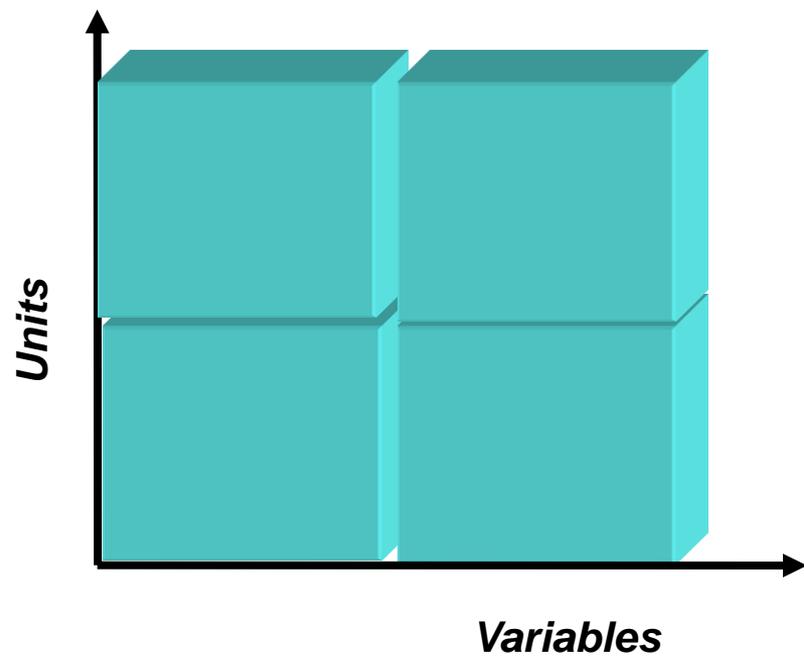


Sistema integrato di dati

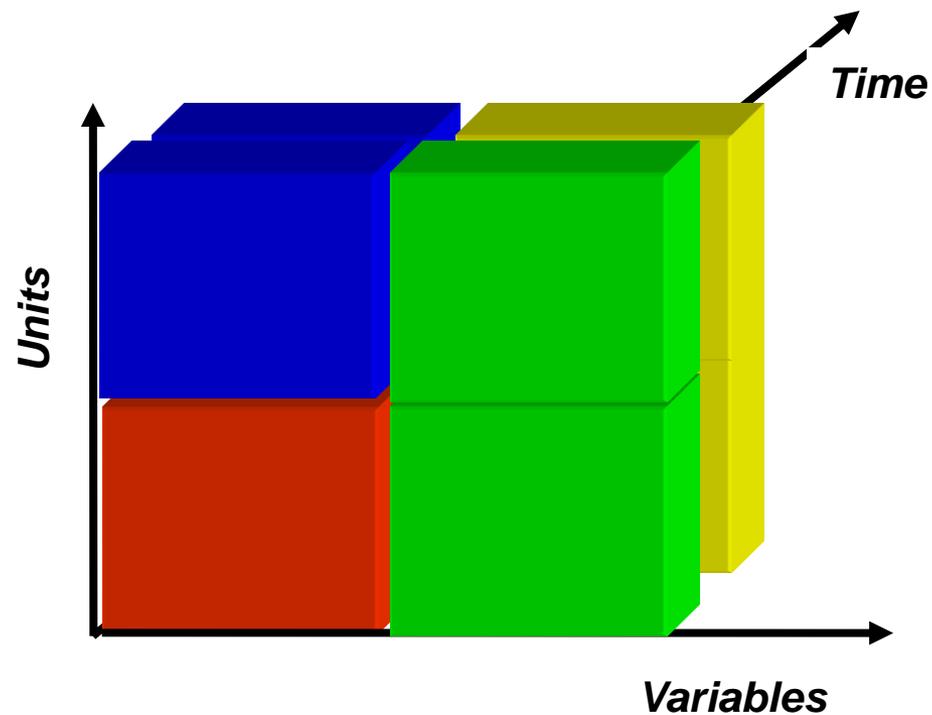
- Riduzione delle risorse finanziarie ed umane
- Riduzione del “fastidio statistico”
- Incremento, in quantità e qualità, delle informazioni statistiche richieste dagli utenti (nazionali e internazionali)
- Incremento della innovazione tecnologica e organizzativa
- Incremento di informazioni di natura differente (dichiarazioni, tracce digitali) disponibili.
- Nuove legislazioni, nazionali ed europee, che facilitano l’accesso da parte degli INS a dati non statistici



Statistical Survey



Multiple Integrated Collection



- Le diverse fonti possono non essere disponibili in tempi diversi
- Possono utilizzare concetti/classificazioni non coerenti fra loro
- Contengono differenti tipologie di errori (non campionari/campionari)
- Possono contenere differenti livelli di qualità intrinseca

Processo produttivo complesso :

INDUSTRIALIZZAZIONE/CENTRALIZZAZIONE

La risposta Istat

Sistema Integrato di Microdati - SIM

Def.: Archivio di microdati amministrativi e statistici integrati a supporto dei processi di produzione statistica

Obiettivi

- Comune pretrattamento di Dati Amministrativi (DA)
- Conformità con le leggi sulla confidenzialità ed il collegamento di dati
- Uniformità di accesso ai DA per i produttori di statistiche
- Evitare duplicazioni di lavoro
- Comune descrizione di metadati e qualità dei DA

Funzioni del SIM

Integrazione riferita al processo di collegamento tra unità identificate in fonti diverse: individui, unità economiche, **luoghi**. Ogni unità è identificata con un numero ID unico e stabile(nel tempo).

A. Analisi Formale dei Concetti/ Identificazione delle unità da DA

B. Caricamento dei dati su tavole

C. Registrazione

D. Integrazione

E. Diffusione ai produttori di statistiche in ISTAT

Processi statistici utilizzando DA

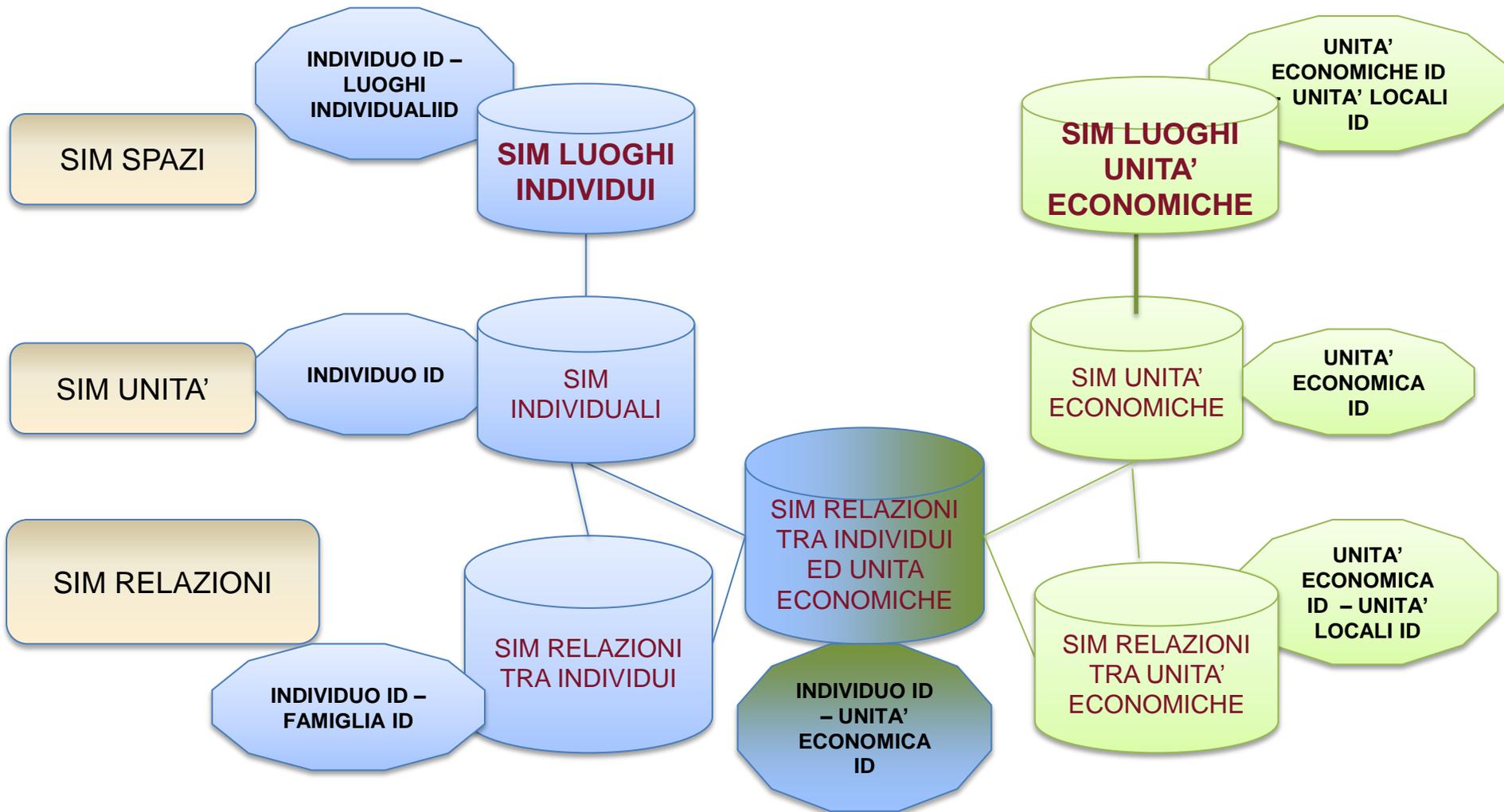
Diffusione per gli utenti statistici



L'integrazione alimenta lo sviluppo di Banche dati di integrazione di diversi sottosistemi di unità statistiche.

Le BD per l'integrazione sono contenitori di microdati utili per consentire una visione uniforme delle unità analizzate mostrando le informazioni disponibili nelle diverse fonti.

Sottosistemi SIM



Sottosistemi integrati: Luoghi

I luoghi costituiscono un autonomo oggetto di analisi

il relativo processo di integrazione porta alla realizzazione dei sottosistemi

- **SIM Luoghi Individui** cioè il sottosistema integrato dei luoghi che interessano le persone fisiche
- **SIM Luoghi Unità Economiche** cioè il sottosistema integrato delle localizzazioni presso le quali un'unità economica esercita le proprie attività

Sottosistemi sui luoghi: il processo di integrazione

Nei SIM luoghi confluiscono le fonti in cui sono presenti indirizzi.

A seguito del processo di integrazione, ad ogni indirizzo è associato un proprio codice identificativo.

La presenza delle chiavi primarie (codice individuo e codice unita) permette di collegare i microdati tra i sottosistemi di base e quelli dei luoghi.

Sottosistemi sui luoghi: il processo di integrazione

LUOGHI COME UNITA' DI RIFERIMENTO

Dove per luogo si intende il singolo indirizzo

Potenzialità:

Partendo dall'indirizzo è possibile aggregare:

- ambiti territoriali i più disparati, e non legati ad ambiti amministrativi predefiniti

- I luoghi del vivere quotidiano:

Lavoro, studio, divertimento, ad es.: i cinema, i ristoranti, gli impianti sportivi, ecc...

Cioè le singole attività economiche del registro luoghi

I luoghi associati alle unità economiche

SIM luoghi unità economiche: contiene le localizzazioni delle unità economiche presenti nelle fonti amministrative.

Fonti potenziali: 25

Tipo informazione	Fonti
Sede	17
UL	9
Altro*	6
Indirizzo non completo	3

*UL sono intese in senso stretto per la parte ASIA; in Altro ci sono indirizzi di Scuole, Università, allevamenti, sedi all'estero

I luoghi associati agli individui

SIM luoghi individui: sono presenti i luoghi che in qualche modo possono interessare le persone fisiche riconosciute dai *data set* amministrativi inseriti nel sistema: residenza anagrafica, domicilio fiscale, indirizzi delle utenze domestiche e così via.

Tipo informazione	Fonti
Indirizzo di residenza	8
Domicilio fiscale	2
Altro	6
Indirizzo non completo	7

Fonti: 15

Possibili sviluppi

Dati amministrativi sulla localizzazione

Per i censimenti permanenti

- censimento popolazione
- censimento edifici e abitazioni
- censimento agricoltura
- censimento delle unità economiche

Per Archimede

Aree tematiche del CP 2011

Questionario CP 2011	Sottosistemi SIM
Notizie Anagrafiche	Individui
Stato civile e matrimonio	Relazioni tra individui
Cittadinanza	Individui
Istruzione e formazione	Individui Relazioni Individui Unità
Condizione professionale	Individui Relazioni Individui Unità
Spostamenti pendolari	Individui Unità Relazioni Individui Unità Luoghi Individui Luoghi Unità

Censimento Non-Profit

Sono state integrate nel SIM Unità 6 fonti amministrative specifiche del settore

Oltre 450mila record utili ogni anno

Contributo per

- dati anagrafici e stato di attività
- forma giuridica
- attività svolte
- luoghi (parzialmente)

Censimento Istituzioni Pubbliche

Integrate nel SIM 3 fonti amministrative

Contributo per

- dati anagrafici e stato di attività
- forma giuridica
- attività svolte
- luoghi
- occupazione (parzialmente)

L'informazione è completata da fonti statistiche

Lista S13 + Indagine RIDDCUE

Censimento Agricoltura

Dati amministrativi; 3 fonti specifiche

Contributo per

- soggetti; titolari; rappresentanti legali
- conduttori; tipo conduzione
- porzioni; particelle; fogli
- attività svolte
- superficie; destinazione
- occupazione (parzialmente)

Alcuni esempi del funzionamenti del SIM

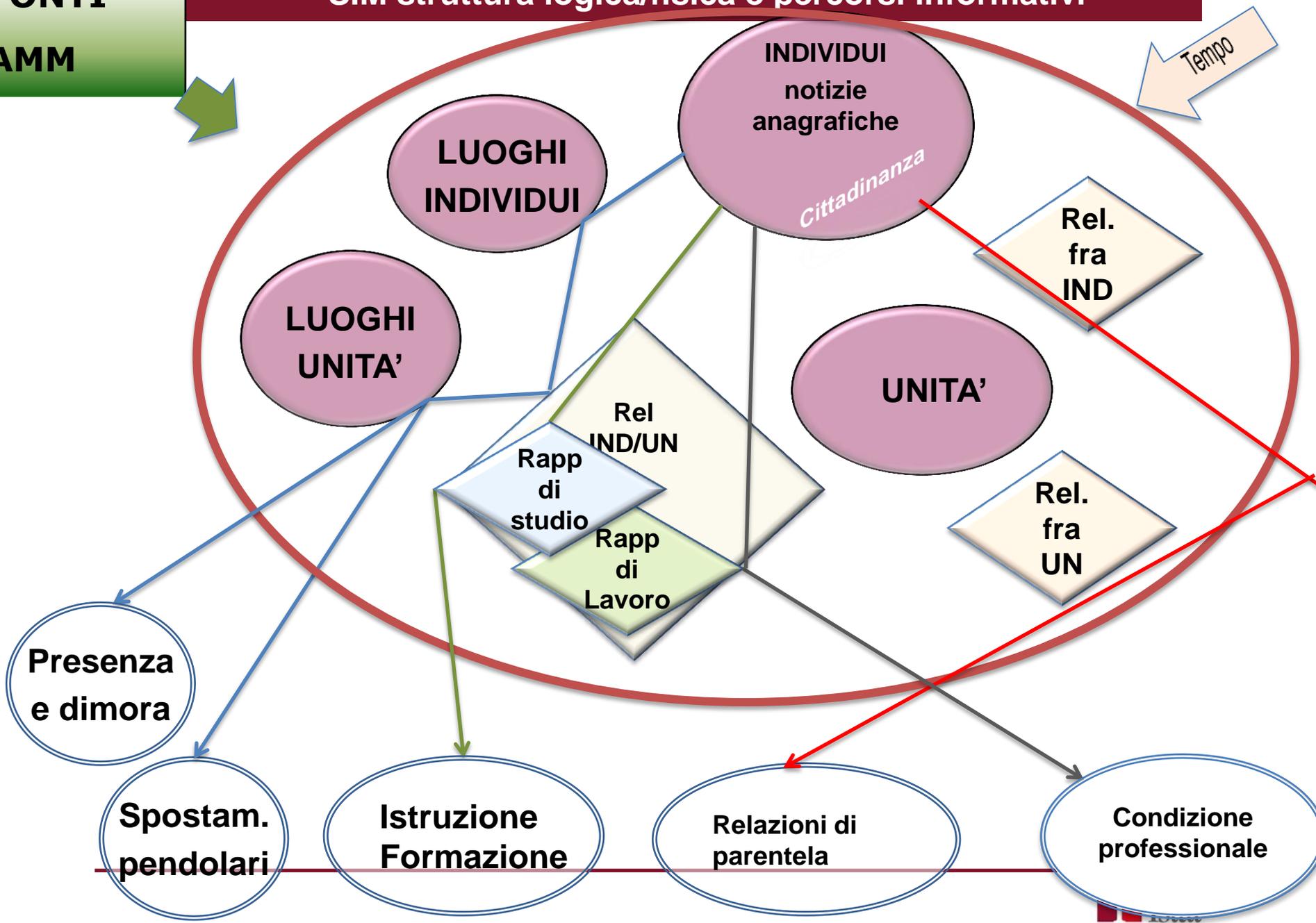
Possibile grazie

alla definizione di un codice identificativo unico per i sottosistemi delle unità di analisi

alla costruzione dei sottosistemi di relazioni

**FONTI
AMM**

SIM struttura logica/fisica e percorsi informativi



Grazie per l'attenzione